ANATOMY OF AN RNA VIRUS: DISSECTING THE HOST-VIRUS INTERACTIONS THAT GOVERN DICISTROVIRUS GENE EXPRESSION AND TRANSMISSION

by

Craig Howard Kerr

B.Sc., The University of Guelph, 2011

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biochemistry and Molecular Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2018

Abstract

Viruses exist as obligate intracellular parasites, with one of the largest classifications of viruses being the positive single stranded RNA viruses ((+)ssRNA). Viral families in this group are incredibly diverse in their replication schemes and host tropisms. Despite this, there exist fundamental principles between them. Unravelling these common mechanisms can give rise to a greater understanding of virus biology and lead to the development of novel antiviral therapies and biotechnology. Members of the *Dicistroviridae* contain monopartite, (+)ssRNA genomes, between 8 to 10 kilobases in size. Infectious to agriculturally and economically important arthropods, these viruses have served as model systems to study fundamental cellular processes such as translation and innate immunity. Dicistroviruses contain two open reading frames (ORFs), which are translated by two distinct internal ribosome entry sites (IRESs). The 5' untranslated region IRES drives translation of the viral non-structural proteins encoded in ORF1, whereas the intergenic region (IGR) IRES directs translation of the viral structural proteins of ORF2. The scheme by which these viruses replicate is poorly described. Here, we develop the first infectious clones of the dicistrovirus type species, Cricket paralysis virus (CrPV), termed CrPV-2 and -3. We demonstrate that this clone is fully infectious in *Drosophila* S2 cells and causes mortality when injected into adult flies. Utilizing this clone, we examined how specific mutations in the IGR IRES affect viral gene expression *in vivo.* Moreover, we demonstrate that the CrPV IGR IRES uses an unusual mechanism for +1-frame translation of a hidden overlapping ORF, which is important for viral pathogenesis. Finally, using a combination of biochemical and mass spectrometry based approaches we show that CrPV may usurp cellular pathways to obtain an envelope. This thesis offers insights into the complex replication scheme of dicistroviruses and provides a foundation for future studies into the life cycle of these viruses.

Lay Summary

Viruses rely on host cells to replicate. By studying the ways in which different viruses replicate we gain a greater understanding of the core principles that govern them. This can lead to the development of new antiviral therapies for humans and animals alike. Insect viruses, such as dicistroviruses, are commonly used as models to study fundamental processes of life. However, the life cycle of these viruses is poorly understood. This thesis provides additional knowledge into how thesis viruses replicate in an effort to develop new tools and information to combat emerging viral threats. Specifically, we developed the first genetic tool that allows researchers to mutate and manipulate sequences in the viral genome. We then utilized this tool to study the strategies implemented by dicistroviruses to synthesize their proteins in the host cell. Finally, we investigated how these viruses hijack pathways in the cell to transfer new viruses between cells.

Preface

A portion of the **Introduction** has been published and adapted for use in this thesis: Kerr, C.H., and E. Jan. (2016). Commandeering the ribosome: lessons learned from Dicistroviruses about translation. *Journal of Virology,* 90(12):5538-5540. The review was written by me under the guidance of Dr. E. Jan.

A version of **Chapter 2** has been published: Kerr, C.H., Wang, Q.S., Keatings, K., Allan, D., Yip, C.K., Foster, L.J., and E. Jan. (2015). The 5' untranslated region of a novel molecular infectious clone of the dicistrovirus Cricket paralysis virus modulates infection. *Journal of Virology,* 89(11):5919-5934. I conducted all the experiments and data analysis apart from Figure 2.4 done by Dr. Q.S. Wang. Dr. E. Jan helped interpret data and write the manuscript.

Chapter 3 is based on the original research article: Kerr, C.H., Ma, Z.W., Jang, C.J., Thompson, S.R., and E. Jan. (2016). Molecular analysis of the factorless internal ribosome entry site in Cricket paralysis virus infection. *Scientific Reports*, 6:37319. Experiments for part of Figures 3.2 and 3.3 were performed by Ma, Z.W. otherwise I conducted and analyzed all experiments. E. Jan assisted in editing the manuscript and experimental design.

In **Chapter 4** Dr. Q.S. Wang contributed equally to the work. Dr. Q.S. Wang performed experiments and analyzed data for Figures 4.1, 4.2A and Appendix A Figures A.1, A.2, A.3, A.4, and A.5. I wrote the manuscript with help from Dr. E. Jan.

The reuse and reprint of all published work is with permission from all journals referenced.

Table of Contents

Abstract ..... **ii**

Lay Summary ..... **iii**

Preface ..... **iv**

Table of Contents ..... **v**

List of Tables ..... **xi**

List of Figures ..... **xii**

List of Symbols ..... **xv**

List of Abbreviations ..... **xvi**

Acknowledgements ..... **xxi**

Chapter 1: Introduction **1**

1.1 Overview of RNA viruses 1

1.1.1 Positive-sense single stranded RNA viruses 4

1.1.1.1 Picornavirus life cycle – Poliovirus as a model 5

1.2 Translation initiation 10

1.2.1 Cap-dependent eukaryotic translation initiation 11

1.3 Translation mechanisms in RNA viruses 15

1.3.1 Leaky scanning and re-initiation of translation 17

1.3.2 Cap-independent translation initiation by RNA viruses 20

1.3.2.1 Internal ribosome entry sites 20

1.3.2.2 3’ cap-independent translation elements 23

1.4 Recoding mechanisms 24

1.4.1 Programmed ribosomal frameshifts 25

1.4.2 Translational bypass 28
1.5 Mechanisms of non-enveloped RNA virus release 30
1.5.1 Lytic release 30
1.5.2 Non-lytic release 32
1.6 The *Dicistrovirdae* family of viruses 37
1.6.1 Characteristics and Genome organization 39
1.6.2 *Dicistrovirus* virus host interactions 43
1.6.2.1 The 5' untranslated region internal ribosome entry site 44
1.6.2.2 The intergenic region internal ribosome entry site 45
1.6.2.3 IGR IRES-mediated alternative reading frame selection 52
1.6.2.4 Significance of the IGR IRES 52
1.7 Thesis investigation 54
Chapter 2: The 5'UTR of a *Cricket paralysis virus* infectious clone modulates viral infection 56
2.1 INTRODUCTION 56
2.2 MATERIALS & METHODS 58
2.2.1 Cell culture and virus 58
2.2.2 CrPV cDNA construction 58
2.2.3 Bicistronic and minigenome constructs 60
2.2.4 *In vitro* transcription and RNA transfection 61
2.2.5 RT-PCR and Northern blots 61
2.2.6 Western blots 62
2.2.7 Electron microscopy 63
2.2.8 Transwell assay 63

2.2.9 *In vitro* and *in vivo* translation assays 63
2.2.10 Fly stocks and viral injections..... 64
2.2.11 Nucleotide sequence accession numbers 65
2.3 RESULTS 65
2.3.1 Selection of a 196-nucleotide element within the CrPV 5'UTR that augments translation..... 65
2.3.2 Construction and characterization of a CrPV infectious clone 70
2.3.3 Transfection of *in vitro* transcribed CrPV-2 and CrPV-3 RNA resembles CrPV infection 75
2.3.4 S2 cells transfected with CrPV-2 or CrPV-3 RNA produce infectious virions 80
2.3.5 Infection by CrPV-2 and CrPV-3 clones is not dependent on helper viruses..... 81
2.3.6 CrPV-2 and CrPV-3 virions are infectious in *Drosophila melanogaster* flies 84
2.3.7 The duplication in the 5'UTR of CrPV-2 is stable and reduces viral fitness 87
2.4 DISCUSSION 90
Chapter 3: Analysis of CrPV IGR IRES-mediated translation during viral infection 96
3.1 INTRODUCTION 96
3.2 MATERIALS & METHODS 99
3.2.1 Cell culture and virus. 99
3.2.2 Construction of CrPV-3 mutants. 100
3.2.3 *In vitro* transcription and RNA transfection. 100
3.2.4 Northern blot analysis. 101
3.2.5 Western blots. 101
3.2.6 *In vitro* and *in vivo* translation assays. 101

3.3 RESULTS 102
3.3.1 IGR IRES mutants of CrPV-3 can synthesize ORF2 proteins in Sf-21 extracts 102
3.3.2 IGR IRES perturbations in the CrPV infectious clone still produce viable virus106
3.3.3 Viable viruses with mutated IGR IRESs are impaired in viral growth at the translation level 110
3.4 DISCUSSION 112
Chapter 4: IGR IRES-mediated translation of a +1 overlapping reading frame enhances CrPV infection in flies 115
4.1 INTRODUCTION 115
4.2 MATERIALS & METHODS 117
4.2.1 Cell culture and virus. 117
4.2.2 Plasmids and bicistronic reporter constructs. 118
4.2.3 *In vitro* transcription and RNA transfection. 119
4.2.4 *In vitro* translation assays. 120
4.2.5 Northern blot analysis. 120
4.2.6 Western blots. 121
4.2.7 LC-MS/MS Analysis. 121
4.2.8 Subcellular fractionation by differential centrifugation. 122
4.2.9 Immunofluorescence. 122
4.2.10 Fly stocks and viral injections. 123
4.3 RESULTS 123
4.3.1 The CrPV IGR IRES-dependent supports +1-frame translation, is IGR IRES-dependent, and initiates downstream 123

4.3.2 The spacer region downstream of the IGR IRES is necessary for +1 frame translation 128
4.3.3 Molecular determinants of CrPV +1-frame translation 131
4.3.4 The integrity of the VLR and pseudo-translocation of the IGR IRES through the ribosome is required for ORFx expression 132
4.3.5 +2 frame translation mediated by the CrPV IRES 134
4.3.6 CrPV +1-frame ORFx is expressed yet not required for infection in *Drosophila* S2 cells 137
4.3.7 ORFx contributes to CrPV pathogenesis in adult flies and associates with membranes 140
4.4 DISCUSSION 146
Chapter 5: CrPV may hijack exosomes to facilitate cell-to-cell transmission in *Drosophila* cell culture.....153
5.1 INTRODUCTION 153
5.2 MATERIALS & METHODS 155
5.2.1 Cell culture and virus. 155
5.2.2 Exosome isolation. 156
5.2.3 Sucrose and iodixanol gradients. 156
5.2.4 Protein digestion and duplex demethylation labeling. 157
5.2.5 Liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis. 157
5.2.6 RT-PCR. 159
5.2.7 Western blots. 159
5.3 RESULTS 160

5.3.1 Isolation of exosome-like vesicles from S2 cells during CrPV infection 160

5.3.2 Proteomic analyses of CrPV-infected cells and derivative ELVs. 162

5.3.3 CrPV may hijack ELVs during infection 166

5.4 DISCUSSION 172

Chapter 6: Summary and future directions 176

REFERENCES 186

Appendices 211

Appendix A : CHAPTER 4 SUPPLEMENTARY DATA 211

Appendix B : CHAPTER 5 SUPPLEMENTARY DATA 216

List of Tables

Table 2.1 Nucleotide and amino acid sequence differences between CrPV-2 and CrPV Sequences..... 74

Table 2.2 Titre of CrPV-2, CrPV-3, and CrPV-ORF1-STOP in Invitrogen S2 cells and Kc167 cells. 86

Table 3.1 List of mutations generated in the CrPV IGR IRES..... 104

List of Figures

Figure 1.1 Baltimore classification of viruses. 2

Figure 1.2 Poliovirus genome organization and polyprotein processing. 7

Figure 1.3 Overview of the poliovirus replication cycle. 8

Figure 1.4 Pathway of canonical eukaryotic translation. 13

Figure 1.5 Known non-canonical translation mechanisms utilized by RNA viruses. 18

Figure 1.6 The classes of viral internal ribosome entry site (IRES) RNA elements. 21

Figure 1.7 Example mechanisms of recoding. 26

Figure 1.8 Exosome biogenesis and HAV hijacking. 34

Figure 1.9 ESCRT-mediated vesicle biogenesis. 36

Figure 1.10 Genome organization of dicistroviruses. 40

Figure 1.11 Morphology and structure of the dicistrovirus virions. 41

Figure 1.12 The *Dicistroviridae* intergenic region internal ribosome entry site. 46

Figure 1.13 Translation initiation pathway of the IGR IRES. 49

Figure 2.1 RT-PCR and sequence of the CrPV 5'UTR from different viral stocks. 66

Figure 2.2 Translational control mediated by the duplication in the CrPV-2 5'UTR. 69

Figure 2.3 Construction and map of the full-length cDNA clones CrPV-2 and CrPV-3. 73

Figure 2.4 CrPV-2 and CrPV-3 RNA produces viral proteins *in vitro*. 76

Figure 2.5 Transfection of *in vitro* transcribed CrPV-2 and CrPV-3 RNA in S2 cells. 80

Figure 2.6 Transfected CrPV RNA clones produce infectious viral particles. 82

Figure 2.7 S2 Cells transfected with *in vitro* transcribed CrPV-2 and CrPV-3 RNA accumulate infectious virions over time. 83

Figure 2.8 CrPV infectious clones can infect Drosophila S2 cells devoid of Drosophila C virus, Flock house virus and Drosophila X virus. 85

Figure 2.9 Injection of CrPV-2 and CrPV-3 virions into adult *Drosophila melanogaster* flies. 88

Figure 2.10 The CrPV-2 5'UTR reduces viral fitness. 89

Figure 3.1 Secondary structure schematic of the CrPV Intergenic region IRES. 103

Figure 3.2 IGR IRES mutants can synthesize viral structural proteins in the context of the CrPV-3 genome. 107

Figure 3.3 Viruses harboring mutations in the IGR IRES produce viable virus. 109

Figure 3.4 Viable viruses with mutated IGR IRESs are impaired in viral growth at the translational level. 111

Figure 4.1. The translation of +1 frame CrPV ORFx is IGR IRES-dependent and initiates downstream of the PKI region 128

Figure 4.2. The downstream sequence of CrPV IGR IRES promotes +1-frame translation. 130

Figure 4.3. Pseudotranslocation is necessary for +1-frame translation. 133

Figure 4.4. CrPV IRES-mediated +2 frame translation. 135

Figure 4.5. ORFx is expressed in CrPV infected S2 cells. 136

Figure 4.6. Pathogenesis of a CrPV mutant lacking ORFx is attenuated in adult *Drosophila melanogaster* flies. 139

Figure 4.7. Cytotoxicity of CrPV ORFx expression in Drosophila S2 cells. 142

Figure 4.8. Subcellular localization of CrPV ORFx. 145

Figure 4.9. Models of CrPV IGR IRES mediated +1-frame ORFx translation initiation 147

Figure 5.1 Validation of exosome-like vesicle isolation from *Drosophila* S2 cells. 163

Figure 5.2 Exosome-like vesicle can be isolated from *Drosophila* S2 cells during CrPV infection. 164

Figure 5.3 Quantitative proteomic analysis of ELVs isolated from mock- and CrPV-infected cells at 6 hpi..... 165

Figure 5.4 MS/MS spectra of one peptide identified from the CrPV ORF2 enriched in ELVs derived from CrPV-infected cells. 167

Figure 5.5 CrPV may hijack ELVs during infection of *Drosophilia* S2 cells. 170

Figure 5.6 ELVs from infected S2 cells elicit subsequent infection and release from cells without lysis. 171

List of Symbols

α	alpha
β	beta
γ	gamma
μ	micron
%	percent
$^\circ$	degree
Δ	delta

List of Abbreviations

(-)	Negative sense
(+)	Positive sense
3C	3C protease
4E-BP	eIF4E-binding protein
ABPV	Acute bee paralysis virus
ALIX	ALG2-interacting protein X
ALPV	Aphid lethal paralysis virus
ATP	Adenosine triphosphate
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
cDNA	Complementary deoxyribonucleic acid
CITE	Cap-independent translation element
CrPV	Cricket Paralysis virus
Cryo-EM	Cryo-electron microscopy
CSFV	Classical swine fever virus
CVB3	Coxsackievirus B3
DCV	Drosophila C virus
DNA	Deoxyribonucleic acid
dsRNA	Double stranded RNA
DXV	Drosophila X virus
EBV	Epstein Barr virus
eEF	Eukaryotic elongation factor

eIF	Eukaryotic initiation factor
EMCV	Encephalomycarditis virus
eRF	Eukaryotic release factor
ESCRT	Endosomal sorting complex required for transport
FFU	Flourescent foci forming units
FHV	Flock house virus
Fluc	Firefly luciferase
FMDV	Foot-and-mouth disease virus
GCN2	general control non-derepressible-2
GTP	Guanine triphosphate
h.p.i.	Hours post infection
h.p.t.	Hours post transfection
HA	Hemagluttinin
HaIV	Halastavi arva virus
HAV	hepatitis A virus
HBV	hepatitis B virus
HCV	hepatitis C virus
HiPV	Himetobi P virus
HIV-1	Human immunodeficiency virus-1
HoCV-1	Homalodisca coagulata virus-1
HRI	heme-regulated inhibitor kinase
IAPV	Israeli actue paralysis virus
IGR	Intergenic region

Imd	Immunodeficiency
IP	Immuno-precipitation
IRES	Internal ribosome entry site
ITAF	IRES trans-activating factor
Kb	Kilobase
KBV	Kashmir bee virus
LC-MS/MS	Liquid chromatography tandem mass spectrometry
MOI	Multiplicity of infection
mRNA	Messenger RNA
mTOR	Mammalian target of rapamycin
MVB	Multi-vesicular body
nt	Nucleotide
ORF	Open reading frame
PABP	Poly(A) binding protein
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PERK	PKR-like endoplasmic reticulum kinase
PIC	Pre-initiation complex
PK	Pseudoknot
PKR	Protein kinase R
PRF	Programmed ribosomal frameshift
PSIV	Plautia stali intestine virus
PV	Poliovirus

RdRp	RNA-dependent RNA polymerase
RhPV	Rhopalosiphum padi virus
Rluc	Renilla luciferase
RNA	Ribonucleic acid
RNAi	RNA interference
rRNA	Ribosomal ribonucleic acid
RT-PCR	Reverse transcription polymerase chain reaction
Sf21	Spodoptera frugiperda
SG	Stress granule
SHAPE	Selective 2' hydroxyl acylation analyzed by primer extension
SINV-1	Solenopsis invicta virus-1
siRNA	Small interfering ribonucleic acid
SL	Stem-loop
ssRNA	Single-stranded RNA
TBST	Tris buffered saline with tween
TC	Ternary complex
TLR	Toll-like receptor
tRNA	Transfer ribonucleic acid
TrV	Triatoma virus
TSG101	Tumor susceptibility gene 101
TSV	Taura syndrome virus
UTR	Untranslated region
UV	Ultraviolet

VLR	Variable loop region
VP	Viral protein
VPg	Viral protein genome-linked
WT	Wild-type

Acknowledgements

Reflecting on this journey through graduate school, I have been fortunate to have been surrounded by so many incredible people. From my friends and family in Ontario to the new relationships I've made in Vancouver, I am enormously thankful for their support in my work, extracurricular activities, and overall endeavors. Moving to Vancouver was one of the biggest decisions of my life and those who I have had the pleasure of spending the past five(ish) years have made it all worthwhile.

First and foremost, I want to say thank you to my two senior supervisors, Drs. Eric Jan and Leonard Foster. It has been an honour to work in your labs and have the both of you as mentors. Both of you have been enablers that has allowed me to have a unique experience in graduate school. This experience has helped me develop critical thinking, analytical, and professional skills that give me the confidence to achieve my future goals in academia.

Eric – Working in your lab has been an incredible experience. Your dedication and enthusiasm for science is infectious (pun intended) and I whole heartedly believe that your trust in me as an independent researcher is what has allowed me to succeed in graduate school. I have appreciated the discussions we've had in the lab, your office, or over drinks about new experiments and exciting research. Hopefully we can continue them in the future! Go Oilers!

Leonard – Over the past five years I've learned so much from working in your lab. I never thought I'd enjoy mass spectrometry as much as I would – who knew it could do so much! Seeing you as a successful scientist, father, and friend has given me the confidence to pursue a career in academia. I hope in the future we will continue to go on amazing hiking trips and chat over ceasars or scotches. Next one is on me!

To the Jan Labbers. I feel as though you guys have essentially been my adoptive family while I've been in Vancouver. I've seen all of you more in the past five years than anyone else in my life, and I wouldn't change a thing. Thank you all for your guidance, criticisms, conversations, laughs, and in general keeping me in check. Dr. Hilda Au, you're one of the most intelligent women I've met and by far the most meticulous. I've always held your opinion in high regard when asking for help and hope that you're able to continue with your love of research. Dr. Julienne Jagdeo, thanks for always keeping me in line in the lab and for supporting me in times of conflict. I know that you will succeed in whatever path you choose – I trust you will become one of the best damn policy makers out there. Dr. Anthony Khong, your light-hearted attitude always generated a great atmosphere in the lab. Since you've become a big-time researcher I've missed our discussions about hockey and experiments – here's hoping we can both make it as profs in the future so we can continue them. Dr. Qing Wang aka Sunny aka 'Mom,' thank you for supporting me in my research and in my personal life. Your excitement for everything and smile always lighted up the lab! To the newer members of the lab – Jibin Sadasivan, Nichalle Brito, Mathew Kirby, Reid Warsaba, Jasmine Li-Brubacher, and Daniel Andrews – you are an incredible group of people and your enthusiasm for science has helped to re-invigorate my own. You're in good hands and I wish you all the best of luck in the future. Lastly, I can't forget the numerous undergraduate students who have helped me accomplish the goals of my PhD. Helen Tran, Erik Tammpere, and Julian Stegmuller, you're alright, thanks.

To the Fosterites, although I haven't always been present in the lab, you have continuously treated me as one of your own. I would not have been able to learn as much as I have over the past years if it weren't for all of you. Special thanks go out to Jenny Moon, Nik Stoynov, Tom Clark, Jason Rogalski, and Greg Stacey – if it weren't for the four of you I would

be lost in the waters of mass spectrometry and bioinformatics. Nichollas Scott, I appreciate your mentorship and teaching me the ins and outs of sample prep protocols and allowing me to look over your shoulder as you ran samples. Alison McAfee, you're an incredible person and scientist. You have one of the brightest futures and are seriously making a name for yourself. I'm grateful that I have gotten the chance to get to know you over the past years. I wish you the best in your future career path!

I've always held the principle that graduate school is what you make of it and that who you surround yourself will largely affect the outcome. I am fortunate that in my first few months living in Vancouver I could create some of the best friendships I've ever had. I couldn't imagine what my PhD would have been like if it weren't for our most tremendous karaoke/ dodgeball team, Microballology. This team has the best dodgeballers and scientists with huge smarts. To everyone who was a part of that team and now our newly formed chat, including Eric "Hollywood" Brown, Sean "All-star" Workman, Nathaneal "Gumby / Na'el / BART" Caveney, Keith "Shark Lyfe" Mewis, Erik "The Red" Nielson, Jan "Grinder" Burian, Mike "Jonesy" Jones, and Bill "Wild Billy" Rees, I'll miss you guys. Finding a group of individuals like these is hard to find as this group of gentlemen is unique; they're a support group, a scientific and political forum, party people. Spending time chatting over drinks about science and life, playing video games or board games is something that will truly be missed. In addition, I spent a large part of my time playing hockey at UBC with the Euler's/Breadwings and hiking mountains in BC, thank you to those who I played alongside of and those who spent time with me in the mountains, these were truly enriching experiences. As well, I need to thank Michael Carlson. You have essentially been my partner in crime during our PhDs. From working on papers

together, forming the biochemistry student society, to creating our own company, it's been great to have you along for the ride.

Finally, as some of the most important people in my life, I need to thank my friends and family in Ontario. To Adam MacArthur and Zac Baszak, thank you for always keeping me line and making sure I know I'm a nerd. Love you guys. To my sister, thank you for your continued support, help when I visit, sisterly encouragement, and general awesomeness. Mom, you are an incredible woman and a true inspiration. Thank you for supporting me in whatever I do making sure to tell me "don't worry it'll work out" when I get stressed and attempting to understand what I do. You've raised me to be passionate, level headed, and to be kind to others. I hope that someday I can repay all that you've done for me. Lastly, I need to thank my fiancé and future wife, Emily. You are my rock. Thank you SO much for your unending support, love, and laughter. Without you the last few years would not have been the same and I am super excited for what our future holds. I can't wait.

For my Mother and Father. Enjoy.

Chapter 1: Introduction

1.1 Overview of RNA viruses

Among its many roles, ribonucleic acid (RNA) can serve as the genetic material of viruses. Viruses have been studied and characterized since the late 19th century. However, it was not until the work of Bawden and Pirie in 1937 that RNA was found to be contained within the viral particles of Tobacco mosaic virus (Bawden and Pirie, 1937). Roughly twenty years later, through the combined efforts of multiple researchers it was revealed that RNA is the *bona fide* genetic material of Tobacco mosaic virus (Fraenkel-Conrat and Singer, 1999). Fast-forward sixty years, the field of RNA virus biology has expanded to include 70 distinct viral families encompassing over 1500 viral species that exist as pathogens across all six kingdoms of life (Adams et al., 2017).

All known viruses are classified into one of seven groups (Group I, II, III, IV, V, VI, or VII) based upon the type of genetic material it bears (Figure 1.1)(Baltimore, 1971). Genuine RNA viruses (i.e. those lacking a DNA intermediate) fall within Groups III, IV, and V. Group III RNA viruses include those with double stranded RNA genomes such as the *Reoviridae* family (e.g. Rotavirus). Group V encompasses viruses that contain a single stranded negative sense RNA genome. Notable viruses from this family include Influenza A virus (*Orthomyxoviridae*) and Ebola virus (*Filoviridae*). Finally, the largest group is the single stranded positive sense RNA viruses comprising Group IV such as the *Picornaviridae* and *Flaviviridae* families.

Regardless of these classifications, RNA viruses continue to serve as a repository for advancing our understanding of fundamental biological questions in cell biology and human health. As the cause of several diseases that have reached epidemic proportions, such as the 1918

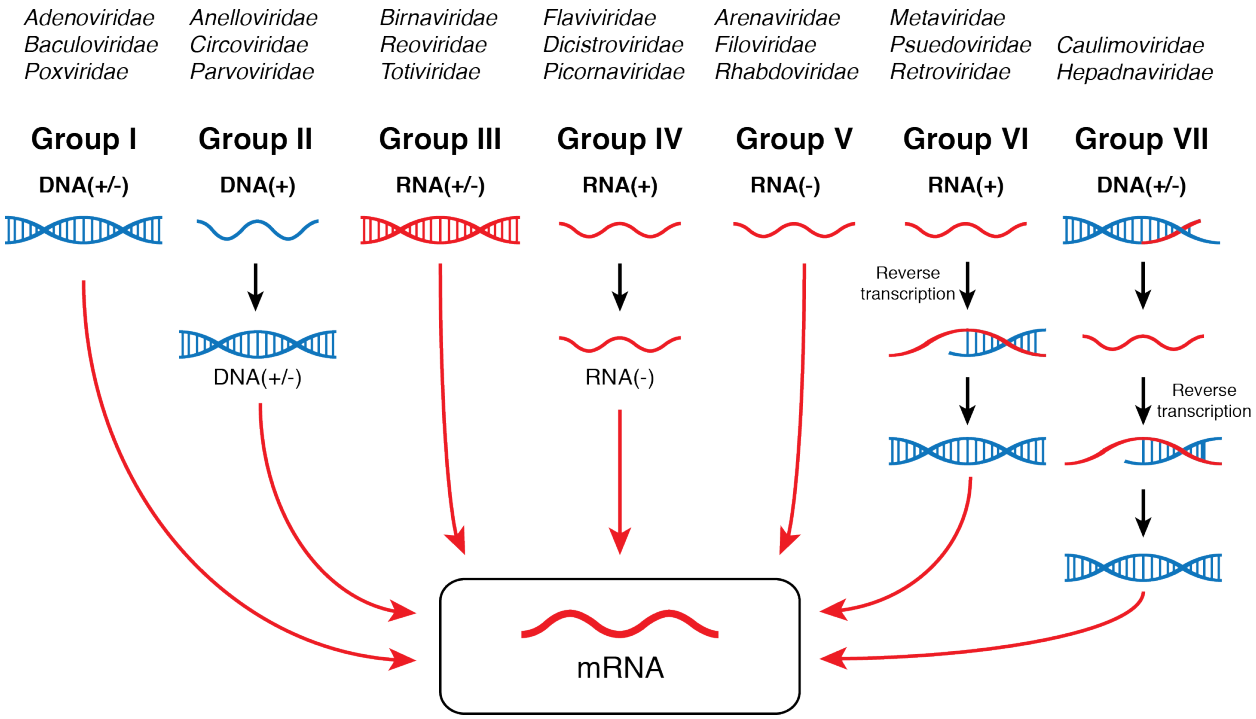


Figure 1.1 Baltimore classification of viruses.

Devised by David Baltimore in 1971, this classification scheme groups viral families based on the genetic material and replicative mechanism. Group I viruses consists of those with dsDNA genomes. The genetic material of Group II viruses is ssDNA. Group III viruses harbor dsRNA genomes. Groups IV and V consist of positive (+) ad negative (-) sense ssRNA genomes, respectively. Group VI viruses contain a positive sense ssRNA that is reverse transcribed into DNA. Finally, Group VII viruses harbour dsDNA genomes that become single stranded RNA before being reverse transcribed and finally transcribed into mRNA.

Spanish flu and more recently the 2013 Ebola virus outbreaks, this class of viruses has plagued human health throughout history.

The challenges surrounding RNA viruses lie in the unique properties they exhibit as the most rapidly evolving entities currently known. During infection, mutations are incorporated into the viral RNA at a high frequency (10^{-4} mutations per nucleotide)(Sanjuan et al., 2010). The incredible plasticity in their genomes leads to quasi-species in any given population of RNA viruses (Lauring and Andino, 2010). As a result, RNA viruses offer a challenging obstacle for the development of vaccines and antiviral treatments. This underscores the need for novel drug targets and by understanding fundamental mechanisms common to RNA viruses it would be possible to develop broad range antiviral therapies. Despite the destructive nature of these viruses, they have contributed substantially to our understanding of modern molecular biology and advances in biotechnology. For example, reoviruses (a Group III dsRNA virus) were used to elucidate the mechanism of TLR3 (Toll-like Receptor 3), which was a crucial step towards our understanding of the human innate immune system (Alexopoulou et al., 2001). Additionally, genomic RNA from phages served as some of the first mRNA templates for *in vitro* translation systems, expanding our knowledge of protein synthesis. Finally, multiple tools we use in modern biology to investigate cellular processes are derived from RNA viruses, such as internal ribosome entry sites and 'stop-go' sequences (Jang et al., 1988; Lee et al., 2005; Pelletier and Sonenberg, 1988; Takahashi et al., 2007; Telpalo-Carpio et al., 2015).

Overall, RNA viruses continue to play a major role in science and society. It is of utmost importance to fully grasp how these viruses replicate and transmit between hosts. Understanding the common underlying principles that are central to the replication processes of RNA viruses

not only impacts our abilities to limit the continuing toll and emerging threat of viral diseases, but can lead to new innovative tools for use in modern molecular biology and industry.

1.1.1 Positive-sense single stranded RNA viruses

Infecting all kingdoms of life except possibly archaea, positive sense single stranded RNA [(+)ssRNA] viruses are the largest group within the RNA virus classifications. In general, (+)ssRNA viruses contain compact genomes ranging from 5 to 32 kilobases (kb) in size. The genomes typically encode relatively few proteins, usually ranging from three to ten, one of which is always an RNA-dependent RNA polymerase (RdRp). Although each virus tends to have a unique replication cycle, there remain some parallels between all of these viruses: (1) the genomic RNA functions as both genetic material and messenger RNA, allowing it to be translated by host ribosomes immediately upon infection; (2) they replicate through a negative sense RNA intermediate where the positive strand RNA is produced in substantial excess compared to the negative strand; (3) (+)ssRNA viruses form membrane associated replication complexes.

Due to the broad host range within this class of viruses, they impact all aspects of life globally including agriculture (discussed more in Section 1.4), ecological activity, and human health. Studies in recent years have demonstrated the (+)ssRNA virosphere greatly influences both the marine and terrestrial ecosystems (Culley et al., 2003; Shi et al., 2016; Suttle, 2007). These viruses contribute to the large turnover of biomass that occurs daily in oceans around the world, and help to shape the biological network of invertebrates globally. It cannot be understated how significant of an impact (+)ssRNA viruses have had on human health and disease. Of note, members of the *flaviviridae* (e.g. dengue virus) and *picornaviridae* (e.g.

Poliovirus) have caused substantial burdens on human health and healthcare systems over the past few hundred years (Murray et al., 2013; Nathanson and Kew, 2010). To understand these viruses, model systems have been developed to investigate their mechanisms of replication and unearth common principles among the (+)ssRNA viruses. In particular, poliovirus has served as a model for the past 50 years to study the life cycle of (+)ssRNA viruses and has provided fundamental insights into the replication of picornaviruses (Leveque and Semler, 2015). Additionally, the global emergence of arthropod-borne viruses (e.g. Zika virus, dengue virus) has thrust insect RNA viruses into the spotlight as more relevant models of the viral life cycle in an insect host.

1.1.1.1 Picornavirus life cycle – Poliovirus as a model

As this thesis focuses primarily on characterizing the life cycle of a (+)ssRNA virus, it is prudent to understand the replication of one of the most well-understood viruses in this group, Poliovirus. Discovered in 1908, Poliovirus (PV) is a member of the *picornaviridae* family and is the etiological agent of poliomyelitis in humans. It possesses a genome size of ~7.5 kb that containing a single open reading frame (ORF). The single ORF is flanked by highly structured 5' and 3' untranslated regions (UTR) followed by a poly(A) tail at the 3' end of the genome. Unlike traditional eukaryotic mRNAs, PV does not harbour a 5' cap on its genome; instead it encodes a viral protein genome-linked (VPg) that binds the 5' end and functions in replication (Jang et al., 1988; Lee et al., 1977; Pelletier and Sonenberg, 1988; Sarnow, 1989; Spector and Baltimore, 1974). In addition to VPg, PV encodes ten additional proteins all of which are expressed as a single polypeptide. Translation of the polypeptide occurs via an internal ribosomes entry site (IRES) in the 5'UTR, after which it is proteolytically processed into seven non-structural and

four structural proteins by a virally encoded proteinase. The viral structural proteins VP1, VP2, VP3, and VP4 assemble to form the icosahedral capsid of PV. Meanwhile the non-structural proteins include those vital for viral replication and disruption of the host cell processes such as VPg, 2C (a putative helicase), 3C (the viral proteinase), and the RdRp ($3D^{pol}$; Figure 1.2)(Hanecak et al., 1982; Lee and Wimmer, 1988; Pfister and Wimmer, 1999; Toyoda et al., 1994).

Mature poliovirus particles (160S particle) attach to the host cell through their cognate receptor, glycoprotein CD155 (Figure 1.3)(Mendelsohn et al., 1989). Attachment triggers receptor mediated endocytosis and causes the virion to undergo dramatic rearrangements of the capsid proteins leading to the porous A or 135S particle. The virus-to-135S transition externalizes VP4 and the amino-terminal extension of VP1 (De Sena and Mandel, 1977; Fricks and Hogle, 1990). These extensions insert into the host cell membrane, distorting it and leading to the formation of an ion channel (Danthi et al., 2003; Fricks and Hogle, 1990). The RNA is then thought to exit the viral capsid into the cytoplasm at the axis of two-fold symmetry in a 3'-to-5' manner; however, the exact trigger of RNA release remains unclear (Harutyunyan et al., 2013). Upon delivery of the viral RNA into the cytoplasm, VPg is removed and host ribosomes, translation initiation factors, and IRES *trans*-acting factors (ITAFs) are recruited to the 5'UTR IRES to facilitate translation (discussed more in Section 1.2)(Andreev et al., 2012; de Breyne et al., 2009; Svitkin et al., 2001). Following translation initiation, the nascent polypeptide is translated and concurrently processed by the viral proteinases to generate the individual viral proteins (Figures 1.2 & 1.3).

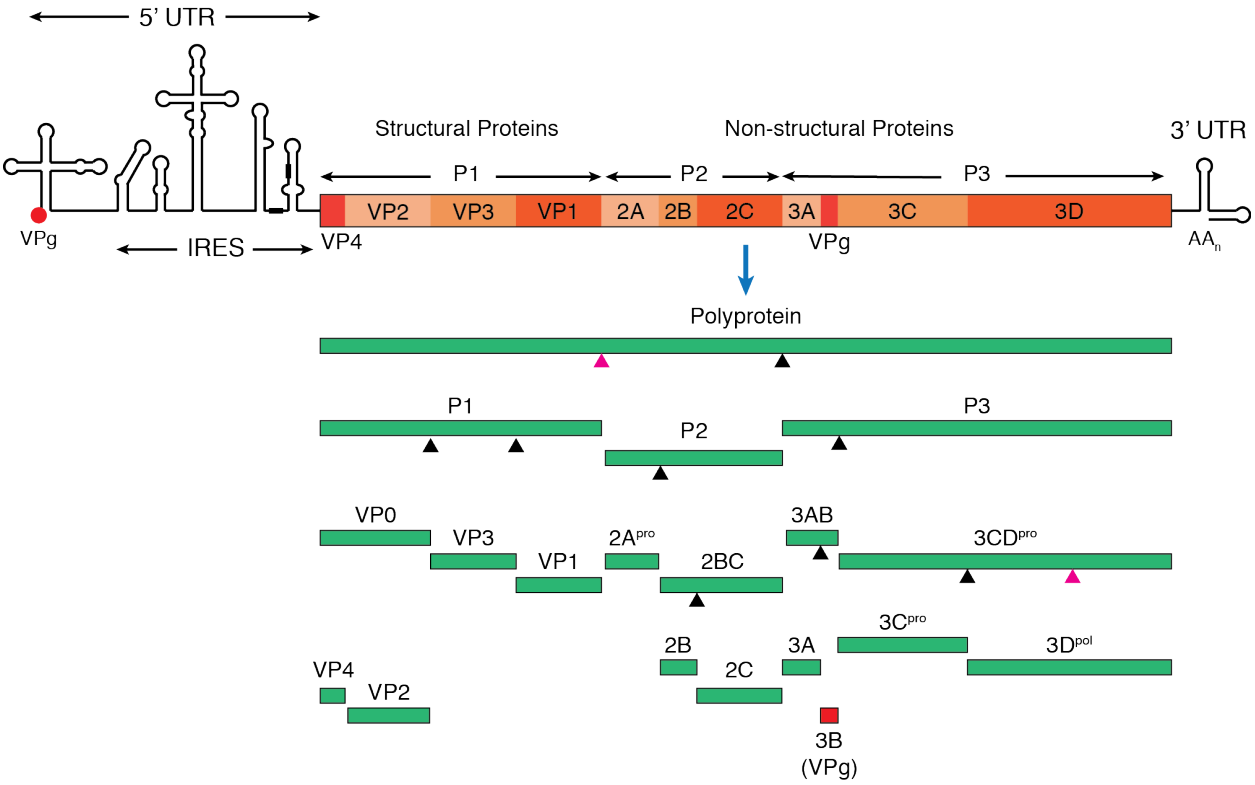


Figure 1.2 Poliovirus genome organization and polyprotein processing.

The viral RNA genome a VPg protein covalently linked to the 5' end and a 3' poly(A) tail. An IRES element in the 5'UTR facilitates translation of the viral polyprotein. The polyprotein is organized into structural and non-structural proteins located at the N- and C-termini, respectively. The viral 2A and 3C/3CD proteases cleave the polyprotein into 3 main precursors P1-3. Black triangles represent 3C cleavage sites, while magenta triangles are 2A cleavage sites. The processing of VP0 to VP4 and VP2 occurs through an unidentified protease, but is required for maturation of the virus.

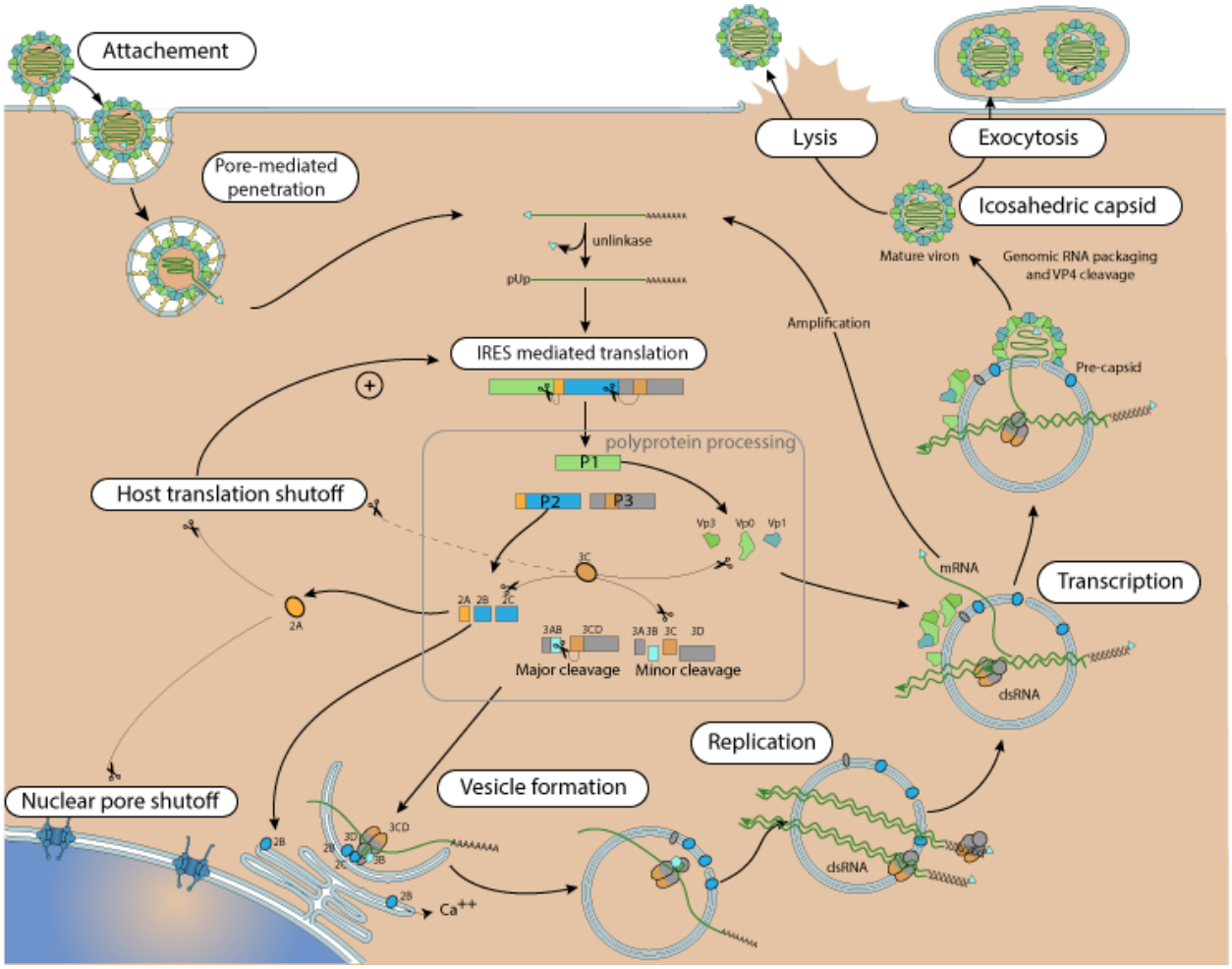


Figure 1.3 Overview of the poliovirus replication cycle.

The life cycle begins as the virus attaches to its cognate receptor, CD155. After receptor-mediated endocytosis, the virion undergoes conformational changes that facilitates export of the genomic RNA. The genome is then translated by host ribosomes and the polyprotein is cleaved into the individual viral proteins. These proteins act to shut off nuclear export and translation (2A), remodel membranes (2B, 2C, 3A, 3B), and replicate the viral RNA (3C/3CD). Replication of the viral RNA occurs at replication complexes on membrane vesicles. Negative strand RNA serves as a template for positive strand synthesis. Newly synthesized genomic RNA is co-packaged into capsids and the virus is released via lysis or non-lytic means. Adapted from ViralZone (www.expasy.orf/viralzone; Swiss Institute of Bioinformatics).

As infection progresses, two hallmark activities occur: (1) there is a massive rearrangement of host cellular membranes into a 'membranous web' like structure and (2) host cell protein synthesis is shutdown. Hijacking of the host cell components to form a membranous web acts to block nuclear transport, release host proteins for viral use, and to serve as a platform for viral RNA replication complexes (Belov et al., 2007; Belov et al., 2000; Gustin and Sarnow, 2001; Hsu et al., 2010). PV replication complexes are formed through the action of proteins 2B and 2C to co-opt Golgi-associated protein complexes and distort the cellular membranes (Cho et al., 1994). The exact origin of the membranes anchoring PV replication is not completely understood; however, it is thought to come from the endoplasmic reticulum (ER)(Jackson, 2014). As host proteins are commandeered for PV replication, viral RNA and proteins are simultaneously targeted to replication complexes (e.g. $3D^{pol}$). Replication commences with negative (-) strand synthesis. Here, the (+) strand viral RNA is first bound by proteins 3AB and anchored to the membrane to serve as a template for replication. Following this, VPg is uridylylated at a Tyr residue (Y3) near the N-terminus by the precursor protein 3CD while bound to a stem loop structure located within the PV genome (known as the cis-acting replication element). Uridylylated VPg (VPg-pUpU) relocates to the 3' poly(A) tail of the (+) strand RNA where it is used as a primer to synthesize the (-) strand viral RNA by $3D^{pol}$ (Flint and American Society for Microbiology., 2009; Paul et al., 2000). Newly synthesized (-) strand is then used to produce (+) strand RNA that is used as both a template for protein synthesis and genomic RNA (Figure 1.3). While RNA replication occurs, PV (+) strand RNAs are utilized to produce large amounts of viral proteins. This is achieved by shutting down host protein synthesis, allowing for preferential translation of viral RNA. In the case of PV, this is achieved in part by cleavage of eukaryotic initiation factor 4G (eIF4G) by the 2A protease (Etchison et al., 1982; Liebig et al.,

1993). This causes cap-dependent translation of the host cell to be inhibited while the C-terminal fragment of eIF4G, along with other factors, binds the 5'UTR of the viral RNA allowing for IRES-dependent translation to occur (discussed in more detail below).

The highly-orchestrated process of poliovirus RNA replication and protein synthesis culminates in association of the structural proteins with the (+) strand RNA into a newly assembled viral capsid. Following maturation of progeny virions, they are primarily released from the cell through lysis of the membranes; however, recent reports have shown that viral release is not exclusive to lysis and can occur through a non-lytic pathway whereby viruses can be packaged *en bloc* and released in extracellular vesicles (discussed more in Section 1.3)(Bird and Kirkegaard, 2015; Chen et al., 2015b).

The past 50 years of study on poliovirus has revealed substantial insights into the life cycle of (+)ssRNA viruses. Being the first animal RNA virus to have its genome sequenced, the first for which an infectious clone was made, and the first to have its X-ray crystal structure solved, PV continues to enable researchers to generate hypotheses about the other (+)ssRNA viruses and use it as a model to understand their replicative mechanisms within the host cell.

1.2 Translation initiation

Protein synthesis is principally regulated at the initiation step to allow for rapid, reversible, and spatial control of gene expression. In eukaryotes, initiation is characterized by the assembly of elongation-competent 80S ribosomes where the initiation codon base-pairs with the anticodon loop of an initiator Met-tRNA in the ribosomal P site. Moreover, initiation serves as the most prominent step that RNA viruses exploit to translate their genomes. Since this thesis addresses an unusual mechanism of viral translation it is requisite to understand general

eukaryotic translation mechanisms. Thus, the following section will outline the process of cap-dependent eukaryotic translation initiation and how viruses have circumvented this pathway using internal ribosome entry sites to achieve their required protein synthesis.

1.2.1 Cap-dependent eukaryotic translation initiation

Eukaryotic translation initiation occurs via a cap-dependent scanning mechanism that requires at least nine eukaryotic initiation factors (eIFs)(reviewed in (Hinnebusch et al., 2016; Hinnebusch and Lorsch, 2012)). Eukaryotic messenger RNA is comprised of two hallmark features that play an intrinsic role during initiation: the 5' cap and poly(A) tail (Figure 1.4). The 5'cap is the defining structural feature of eukaryotic mRNAs; it consists of a 7-methyl-guanosine that is co-transcriptionally linked through an inverted 5'-5' triphosphate bridge to the initiating nucleotide of the transcript (reviewed in (Shuman, 2002)). The capping pathway is conserved in all eukaryotic organisms that have been examined. The cap moiety itself serves to: 1) protect the RNA message from degradation by exonucleases, 2) promote pre-mRNA splicing, and 3) support a pioneer round of translation for mRNA quality control (Edery and Sonenberg, 1985; Konarska et al., 1984; Lewis and Izaurralde, 1997; Maquat et al., 2010; Schwer et al., 1998).

Initiation begins with recognition of the 5' cap by the cap-binding complex, eIF4F, and assembly of the 43S pre-initiation complex (PIC; Figure 1.4). eIF4F is comprised of three factors: the cap-binding protein, eIF4E; the RNA helicase, eIF4A; and the adaptor protein, eIF4G. Meanwhile the 43S PIC consists of methinoyl-initiator tRNA (Met-$tRNA_i$) in a ternary complex (TC) with GTP-bound eIF2 and eIF1, eIF1A, eIF3, and eIF5 (Hinnebusch et al., 2016; Hinnebusch and Lorsch, 2012). The subsequent attachment of the 43S PIC to m^7G-capped RNA

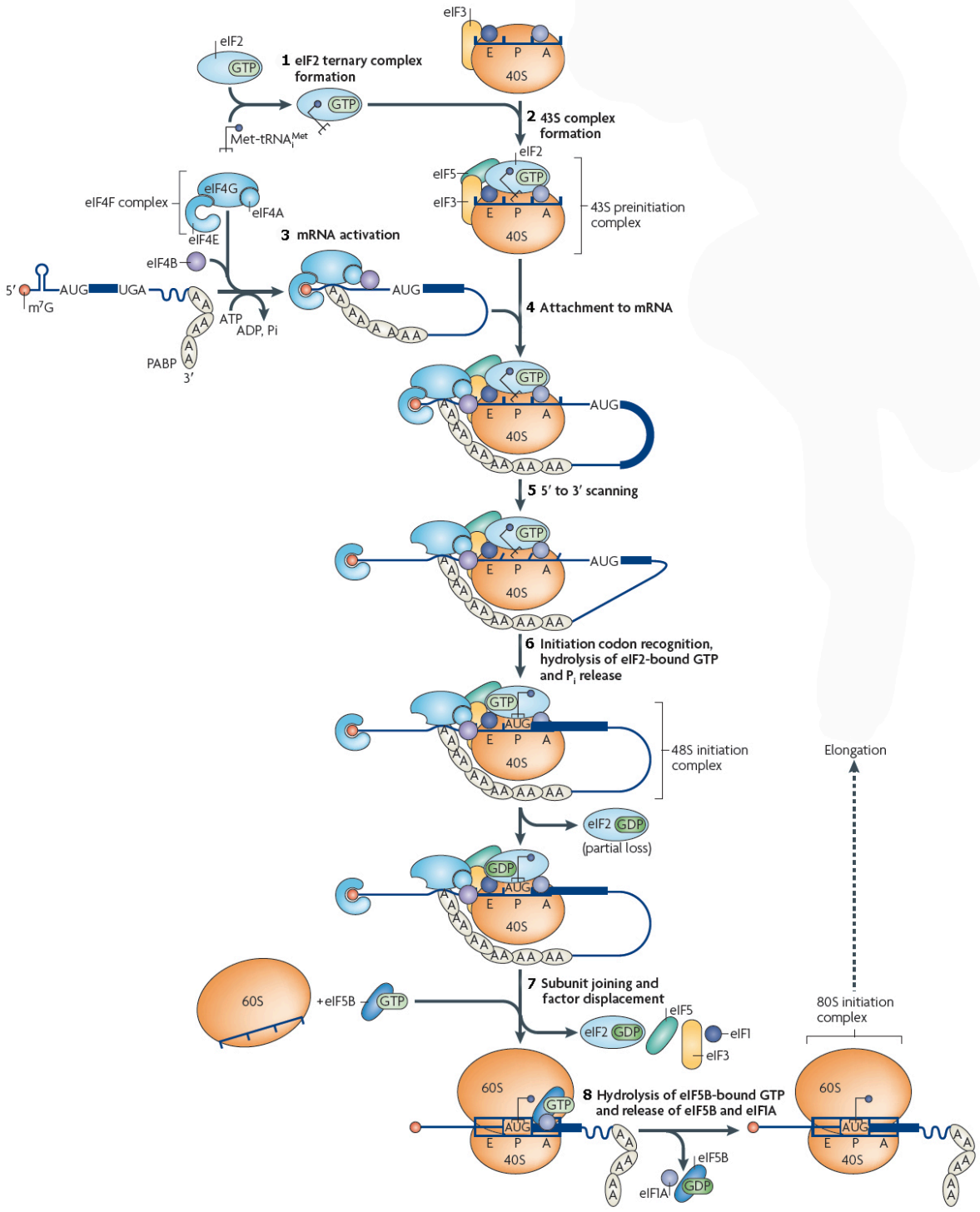

eIF2
GTP
1 eIF2 ternary complex formation
Met-tRNA$_i^{Met}$
eIF3
E P A
40S
2 43S complex formation
eIF2
eIF5
eIF3
43S preinitiation complex
eIF4F complex
eIF4G
eIF4A
eIF4E
3 mRNA activation
eIF4B
5′
m^{7}G
AUG
UGA
PABP
3′
ATP
ADP, Pi
4 Attachment to mRNA
5 5′ to 3′ scanning
6 Initiation codon recognition, hydrolysis of eIF2-bound GTP and P$_i$ release
48S initiation complex
eIF2 GDP
(partial loss)
Elongation
60S
+eIF5B
7 Subunit joining and factor displacement
eIF5
eIF1
eIF3
80S initiation complex
8 Hydrolysis of eIF5B-bound GTP and release of eIF5B and eIF1A
eIF5B
eIF1A

Figure 1.4 Pathway of canonical eukaryotic translation.

Canonical eukaryotic translations occurs in a series of eight steps **(1-8)**. The stages follows as: **(1)** formation of the ternary complex (TC) of eIF2·Met-tRNA$_i$·GTP; **(2)** binding of the TC to the 40S ribosomal subunit and eIF1, 1A, and 3 to form the 43S pre-initiation complex (PIC); **(3)** mRNA activation by binding of eIF4F; **(4)** attachment of the 43S PIC to the mRNA through the cap-binding complex; **(5)** ATP-dependent scanning in the 5' to 3' direction by the 43S complex; **(6)** start codon recognition and formation of the 48S initiation complex by switching the scanning complex to a 'closed' conformation; **(7)** 60S subunit joining and displacement of initiation factors; **(8)** hydrolysis of GTP by eIF5B and release of eIF1A to form elongation-competent 80S ribosomes. Adapted with permission from (Jackson et al., 2010).

---

is facilitated by interactions eIF3 with eIF4G. Additionally, eIF4G serves as an important scaffold to mediate circularization of the mRNA through interactions with the poly(A) binding protein (PABP) bound to the 3' poly(A) tail. Once binding and circularization are complete, the entire PIC scans 5' to 3' directionally along the 5'UTR in search of an AUG initiation codon. The helicase activity of eIF4A acts to unwind local RNA secondary structure while factors eIF1 and eIF1A induce a scanning-competent 'open' conformation of the ribosome (Passmore et al., 2007). Due to the nature of scanning, the initiation codon is frequently the AUG closest to the 5' end. However, some transcripts contain multiple AUGs in the 5'UTR. Thus, to ensure fidelity in translation initiation and proper usage of the authentic initiation codon, the scanning complex discriminates between AUG codons in favorable sequence context and those which are not. Termed the 'Kozak' sequence in mammals, an initiation codon that is within an optimal consensus sequence of GCC(A/G)CCAUGG, where a purine is at the -3 position and a guanine is at the +4 position, will allow for preferential assembly of the initiation complex (Kozak, 1986, 1991).

Upon recognition of the start codon through Watson-Crick base pairing between the Met-$tRNA_i$ and AUG codon, eIF1 is displaced, placing the 40S ribosomal complex into a 'closed' conformation that remains engaged on the mRNA (Passmore et al., 2007). Following stalling at the start codon, the GTPase-activating protein, eIF5, stimulates the GTP hydrolysis activity of GTP bound-eIF2 (Paulin et al., 2001). This reduces the affinity of eIF2 for Met-$tRNA_i$ and it dissociates from the 40S subunit (Kapp and Lorsch, 2004). eIF5B then induces the release of the remaining initiation factors (eIF1, eIF1A, and eIF3) and promotes 60S ribosomal subunit joining (Pestova et al., 2000). Subsequent hydrolysis of GTP from eIF5B alters the structure of the newly formed 80S ribosomes into an elongation-competent state. With the Met-$tRNA_i$ positioned at the P site of the ribosome, the adjacent A site is vacant for the delivery of the next aminoacyl-tRNA and elongation proceeds.

It is apparent now that translation initiation is a highly complex process requiring precise timing and execution. This attribute makes it amenable to tight regulation. Mechanisms of regulating initiation fall into two broad categories: those that impact the eIFs, and those that impact the mRNA. The latter category falls outside the scope of this discussion, but is well-reviewed elsewhere (Jackson et al., 2010). The most well characterized examples of the first type are the phospho-regulation of eIF2 and eIF4F. The availability of the ternary complex (TC; eIF2·Met-$tRNA_i$·GTP) is critical for delivery of the intiator Met-tRNA. eIF2 is a trimeric factor comprised of α-, β-, and γ-subunits. The γ-subunit is responsible for binding GTP and Met-$tRNA_i$ while the α-subunit is subject to regulation by phosphorylation at serine 51 during times of cellular stress. When phosphorylated, eIF2 sequesters eIF2B and effectively inhibits its ability to recycle GDP to GTP (Rowlands et al., 1988). Consequently, eIF2 levels in the ternary complex fall and most translation is halted. In mammalian cells, there are four known kinases

that post-translationally modify the eIF2 α-subunit upon cellular stress: PKR (protein kinase R) is stimulated by double stranded RNAs, PERK (PKR-like endoplasmic reticulum kinase) is activated by ER stress, GCN2 (general control non-derepressible-2) is activated upon amino acid depletion, and HRI (heme-regulated inhibitor kinase) activates during heme deficiency (Jackson et al., 2010). As expected, viral infection triggers the phosphorylation of eIF2 through PKR stimulation from virally-derived dsRNAs, ER stress from overloaded protein synthesis, or GCN2 due to consumption of cellular resources (Sonenberg and Hinnebusch, 2009).

Intracellular concentrations of eIF4F are also indirectly affected through post-translational modifications of the eIF4E-binding proteins (4E-BP). When hypo-phosphorylated, 4E-BP outcompetes eIF4G for the binding site on eIF4E. This impedes the binding of eIF4E with the eIF4G scaffold protein and disrupts recognition of the 5' m^7G cap by the 43S PIC. Activation of the kinase mTOR (mammalian target of rapamycin) leads to hyper-phosphorylation of 4E-BP, releasing eIF4E, and allowing it to assimilate into the cap-binding complex (Ma and Blenis, 2009; Richter and Sonenberg, 2005). Multiple cellular signals such as nutrient availability, cellular energy status, and growth factors or hormones activate mTOR signaling placing it at the nexus for several major regulatory networks. This allows cells to integrate multiple stimuli simultaneously, thereby triggering growth during favorable conditions or inhibiting translation when cell survival outweighs growth such as during times of viral infection.

1.3 Translation mechanisms in RNA viruses

As a main component in the Central Dogma of Molecular Biology, protein synthesis is an essential process in all forms of life. Including initiation, elongation, termination, and ribosome

recycling, translation is a highly-orchestrated process whereby the genetic information stored in messenger RNAs is transformed into a nascent polypeptide by ribosomes.

Viruses exist as obligate intracellular pathogens and therefore must repurpose at least some host cell machinery for their own use. No known virus encodes its own ribosome, leaving the hijacking of this cellular component as a unifying feature for the success of all viral propagation. As viruses co-evolve with their hosts, a molecular arms race arises between the two where the host evolves mechanisms to restrict viral gene expression and the virus counters through antagonizing factors or by circumventing the pathway altogether. Since translation is a critical step in the viral life cycle, viruses have evolved numerous mechanisms to circumvent cellular pathways of translation including molecular mimicry, appropriation of host components, and alternative means of translation (reviewed in (Jan et al., 2016)).

Alternative modes of translation or non-canonical translation is pronounced within RNA viruses (Au and Jan, 2014; Firth and Brierley, 2012). Canonical, 5'-end dependency in eukaryotic translation poses a major challenge for RNA viruses as it typically results in translation of a single mRNA to a single protein sequence. Due to constraints on genome size, RNA viruses, in general, must express multiple proteins from a single transcript to facilitate replication; thus, they have evolved a variety of solutions to overcome this obstacle. A subset of viruses simply conform to the 5' dependency; e.g., generating subgenomic RNAs containing a single ORF (e.g. coronaviruses) or harboring segmented genomes where each segment encodes one ORF (e.g. orthomyxoviruses). Conversely, a common strategy of (+)ssRNA viruses is to encode long polyproteins that are post-translationally cleaved into their individual parts via a viral protease (e.g. picornaviruses). Regardless, caveats exist with each of these strategies. Segmented genomes can be inappropriately packaged during viral assembly, leading to non-

infectious progeny virions; and polyprotein synthesis can be considered an inefficient use of the plentiful cellular resources as there is typically no control over stoichiometric quantities of each protein being expressed. Thus, non-canonical translation mechanisms present a unique way to exploit the host machinery to express multiple proteins from a single ORF.

Constraints in the size of RNA virus genomes due to limitations in particle size impart strong selective pressures, forcing RNA viruses to maximize coding capacity. The presence of overlapping open reading frames is a common mechanism to achieve multi-functionality from a limited RNA message. To express these ORFs, RNA viruses employ mechanisms of non-canonical translation at the initiation, elongation, and terminations steps of translation such as leaky scanning, ribosomal frameshifting (discussed in Section 1.2.2), and re-initiation, respectively (Figure 1.5). Although there exist multiple mechanisms for non-canonical translation, only a few will be briefly discussed here (all are reviewed in various places (Au and Jan, 2014; Firth and Brierley, 2012; Jan et al., 2016; Miras et al., 2017)).

1.3.1 Leaky scanning and re-initiation of translation

Leaky scanning is perhaps the most prominent translation mechanisms observed in RNA viruses (e.g. see *Tombusvirus* and *Hepevirus*). It allows for the expression of proteins with identical C-termini, overlapping ORFs, and even distinct sequential ORFs (Graff et al., 2006; Herzog et al., 1995; Turina et al., 2000). Leaky scanning is the process where the 40S ribosomal subunits bind near the 5' end, scan along the RNA, fail to initiate at the first AUG codon, allowing a significant portion to initiate translation at an alternative AUG codon downstream (Figure 1.5)(Kozak, 2002).

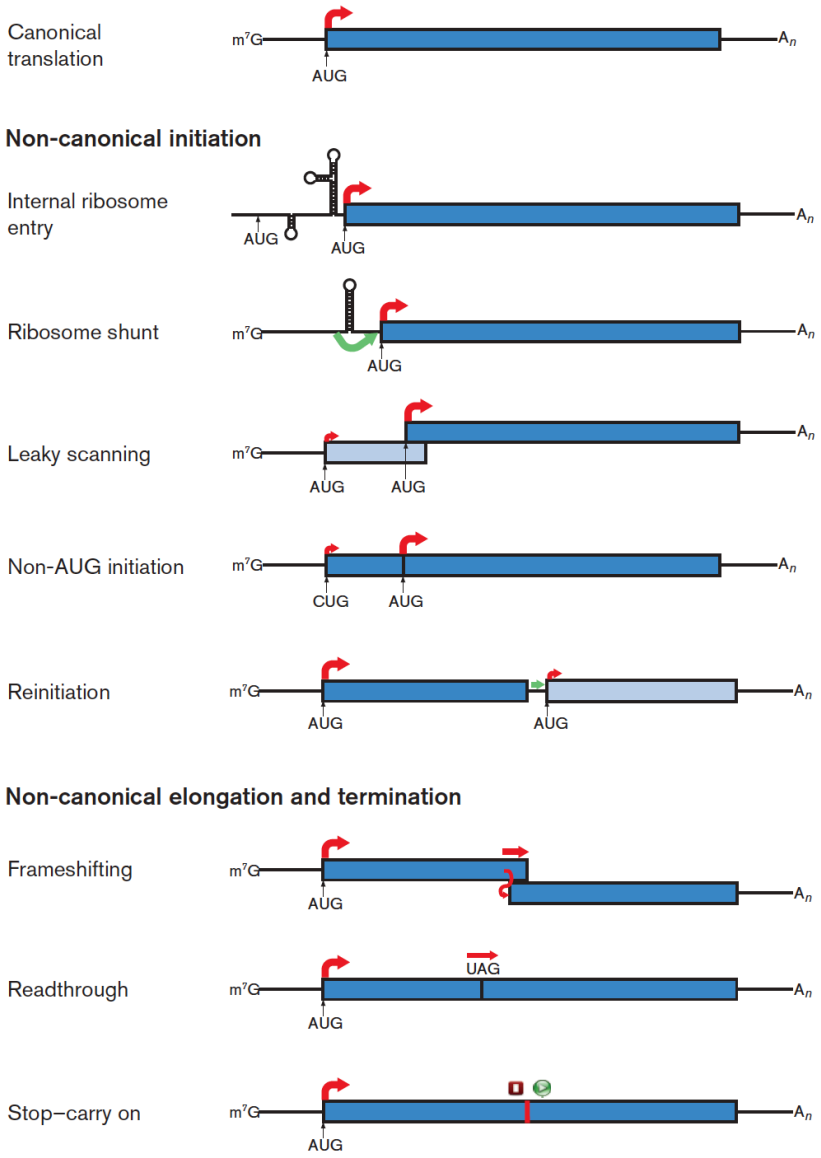


Figure 1.5 Known non-canonical translation mechanisms utilized by RNA viruses.

Top panel displays canonical eukaryotic translation. Red arrows indicate the translation initiation site at the beginning of the ORF or continuation of translation by 80S ribosomes. The thicker arrow shows the predominant path taken by 80S ribosomes. Green arrows display the suggest the non-canonical movement of 40S ribosomal subunits. Instances where two polypeptides are synthesized are indicated by different shades of blue. For frameshifting, since it generally leads to C-terminal extensions of the same polypeptide it is shown in the same colour. Stop-Carry on mechanisms are non-canonical at the termination and initiation steps as indicated by the red square and green circle. Reproduced with permission from (Firth and Brierley, 2012).

The frequency at which downstream initiation occurs is largely due to the nucleotide context at the initiation site (Kozak, 1986). If the context of the first AUG codon is sub-optimal there is potential for leaky scanning. However, some cases have been observed where initiation first occurs at an upstream non-AUG codons as well (Castano et al., 2009; Turina et al., 2000). Furthermore, leaky scanning can occur as a result of two initiation codons being within close proximity of one another (e.g. within ~10 nt). The data surrounding this model suggest that scanning may occur via alternating forward thrusts and backwards relaxations allowing downstream AUG codons to capture some ribosomes that would otherwise initiate at the first upstream AUG (Matsuda and Dreher, 2006; Williams and Lamb, 1989).

Re-initiation occurs at the termination step of translation elongation. Typically, upon termination of translation the 40S and 60S ribosomal subunits dissociate and are recycled for further use. Although in some cases with a short ORF (e.g. less than 30 codons) the 40S subunit may remain associated with the transcript, resume scanning, and re-initiate translation at a downstream AUG codon (Figure 1.5)(reviewed by (Jackson et al., 2012)). How the 40S subunit remains attached to the mRNA is unknown, but is thought to rely on two criteria: (1) specific initiation factors remain bound to the ribosome during translation of the short ORF and (2) the ribosome must scan for some distance before becoming fully competent to re-initiate (Poyry et al., 2007; Poyry et al., 2004). Use of short ORFs such as this is thought to be prominent in the cell for the regulation of downstream ORFs during times of stress (Ingolia et al., 2011; Morris and Geballe, 2000). As it is utilized by both viruses and eukaryotic cells, this mechanism of non-canonical translation appears to be widespread in nature.

Irrespective of the mechanism, viruses have evolved numerous strategies to bypass the need for canonical translation to express their proteins. Commonly RNA viruses will, in fact,

employ multiple non-canonical mechanisms at once to ensure proper translation of their genomes. By studying these mechanisms, we gain a deeper understanding of viral gene expression. This led to the development of extremely valuable tools for molecular biology and biotechnology applications such as viral internal ribosome entry sites that are often used to co-express reporter genes.

1.3.2 Cap-independent translation initiation by RNA viruses

1.3.2.1 Internal ribosome entry sites

Cap-dependent translation exists as the primary mode by which cellular transcripts recruit the translational machinery. Although some viruses adopted this mechanism, numerous viruses (especially (+)ssRNA viruses) utilize internal ribosome entry as an alternative form of translation initiation. Internal ribosome entry sites, or IRESs, are *cis*-acting structured RNA elements that can recruit the translational machinery independent of the 5' cap and typically reside within untranslated regions.

IRESs were first described in picornaviruses, specifically poliovirus (PV) and encephalomycoarditis virus (EMCV) (Jang et al., 1988; Pelletier and Sonenberg, 1988). Since picornaviral genomes lack the conventional 5' m^7G cap and have highly structured 5'UTRs containing multiple non-initiating AUG codons, they were initially studied for their ability to mediate cap-independent translation. Using a bicistronic vector system where pieces of the viral RNA were placed in the intergenic space of two reporter genes it was found that internal initiation of translation was possible (Jang et al., 1988; Pelletier and Sonenberg, 1988). This discovery was definitively proven after observing the expression of a reporter gene from a

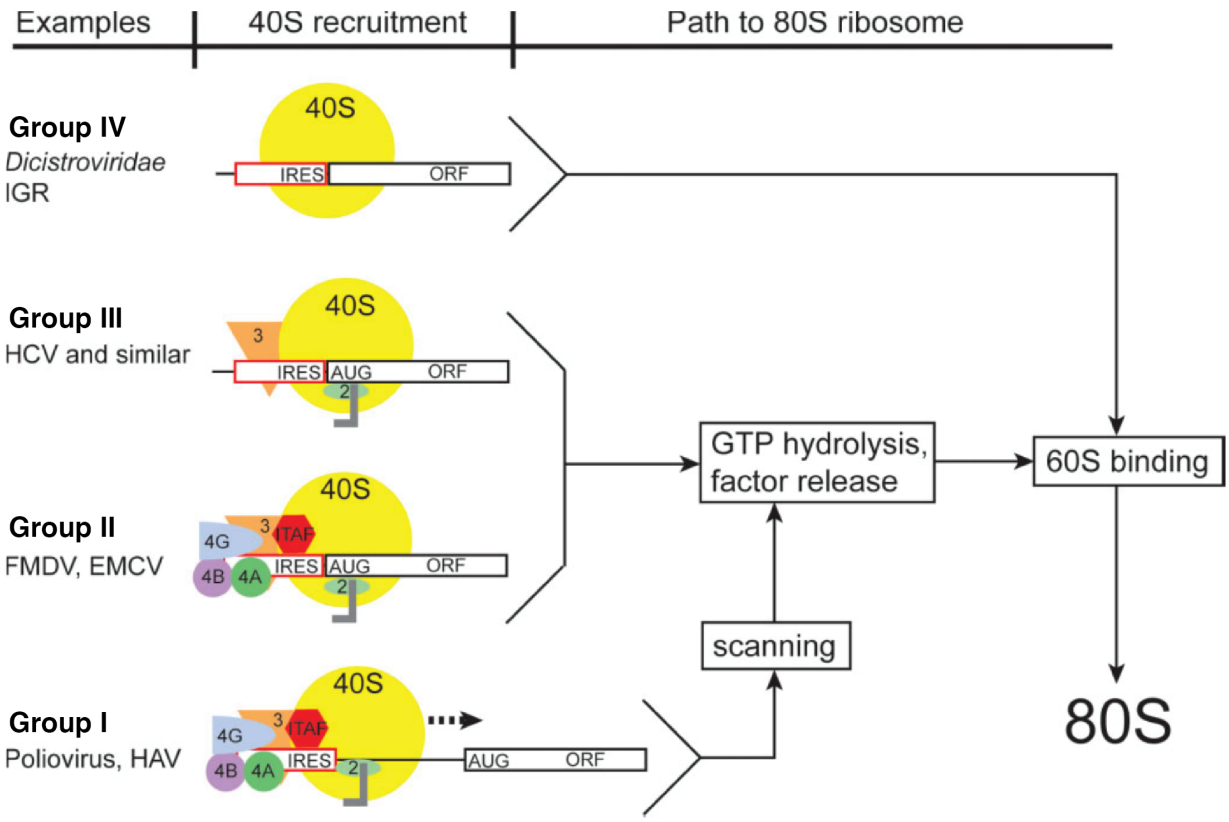


Figure 1.6 The classes of viral internal ribosome entry site (IRES) RNA elements.

Depending on the required initiation factors, IRES *trans*-activating factors, and phylogenetic features, viral IRESs can be classified into one of four groups (I, II, III, and IV). Group IV IRESs represent the most streamlined mechanism that dispenses the need for all canonical translation initiation factors and Met-tRNA$_i$. Group III IRESs directly bind the 40S ribosomal subunit and require eIF2 and 3, and tRNA$_i$. Group II and Group I IRESs require Met-tRNA$_i$ and all canonical initiation factors except eIF4E. The distinguishing feature between Group III and IV is that IIIs directly position the ribosome at the translational start site, while IVs must undergo 40S scanning before initiating downstream. Reproduced with permission from (Plank and Kieft, 2012).

circularized RNA harbouring the EMCV IRES (Chen and Sarnow, 1995). Since the discovery of these IRESs, they have been found to be encoded in numerous viral genomes and even a subset of cellular mRNAs (reviewed in (Kieft, 2008; Lacerda et al., 2017; Yamamoto et al., 2017). The field of cellular IRESs is beyond the scope of this thesis and as such will not be discussed.

Viral IRESs are classified into four groups: I, II, III, or IV (Figure 1.6). Classification is grounded on the requirements for translation initiation factors and IRES *trans*-activating factors (ITAFs), and phylogenetic similarities (Kieft, 2008). The Group I IRESs require Met-tRNA$_i$, ITAFs, and the full allotment of initiation factors except for eIF4E. These IRES recruit the 43S PIC to the IRES which then scans the mRNA for the downstream initiation site. The best described examples of these IRESs are within poliovirus and hepatitis A virus. Group II IRESs are functionally equivalent to Group I IRESs, with the exception that Group IIIs positon the initiation complex directly at the translation start site. EMCV and foot-and-mouth disease virus (FMDV) best exemplify this group of IRESs. Group III IRESs, such as those found in hepatitis C virus (HCV) and classical swine fever virus (CSFV), bind the 40S ribosomal subunit directly in conjunction with Met-tRNA$_i$ and a small subset of initiation factors (eIF2 and 3). With these IRESs, the 40S subunit binds and the initiation codon is positioned within the P site of the ribosome, removing the need for scanning (Pestova et al., 1998). Finally, IRESs belonging to Group IV have only been described within the *Dicistroviridae* family of viruses. These IRESs utilize the most streamlined mechanism known to date as they dispense of the need for all canonical translation initiation factors and Met-tRNA$_i$. Remarkably, these IRESs directly bind the ribosome and initiate translation from the ribosomal A site at a non-AUG codon (Figure 1.6; discussed more in Section 1.6.2.2)(Jan and Sarnow, 2002; Sasaki and Nakashima, 1999, 2000; Wilson et al., 2000a).

DNA-based bicistronic constructs are the classical tools for investigating viral or cellular IRESs. However, the use of these as research tools has been contentious. When using DNA-based constructs, potential IRES elements can contain cryptic promoters or splice sites that misconstrue experimental results (reviewed in (Kozak, 2003, 2005)). These artefacts can generate shorter or separate monocistronic RNAs that encode the second cistron, leaving it to be expressed through cap-dependent translation (Riley et al., 2010). Thus, rigorous control experiments must be conducted to verify a *bona fide* IRES element (reviewed in (Thompson, 2012)). To control for these potential pitfalls, Northern blotting or quantitative real-time PCR can be used to measure levels of each cistron and presence of aberrant transcripts. Additionally, co-transfecting siRNAs targeting the upstream reporter gene of the bicistronic construct can be used. A similar reduction in the expression of both cistrons should be observed if the transcript remains intact. To avoid these potential caveats altogether, an RNA-based reporter system can be used in place of a DNA-based construct (Thompson, 2012).

1.3.2.2 3' cap-independent translation elements

Although they are not considered IRES sequences, some viruses encode RNA elements that initiate translation from the 3' end of the genome (comprehensively reviewed in (Simon and Miller, 2013)). Many plant RNA viruses (e.g. *Tombusviridae* and *Luteoviridae*) contain sequences known as 3' cap-independent translation elements (3'-CITE) in their 3'UTR that take on different forms as I-shaped, Y-shaped, T-shaped, pseudoknotted structures, or radiating helices from a central RNA hub. Moreover, in most cases there is a long-distance RNA:RNA interaction that transfers bound factors at the 3' end to the 5' end. Implied by their name, 3'-CITEs functionally substitute for the 5' cap by recruiting translation initiation factors to the 5'

end of the viral genome, followed by scanning to the initiation codon. The most well-characterized CITE is found within Barley yellow dwarf virus (Guo et al., 2001; Wang and Miller, 1995). This CITE binds to eIF4E and 4G to directly recruit the 43S PIC along with eIF4A and 4B (Sharma et al., 2015). A long distance RNA:RNA kissing loop interaction then directs the initiation complex towards the 5'UTR of the viral genome allowing scanning to ensue. However, the exact mechanism by which 3'CITEs direct ribosomes to the 5' end of the RNA is unknown. Although 3'CITEs have only been characterized in plant viruses, it would seem unlikely that such an efficient 3' translational mechanism would not be utilized by mRNAs as well.

Overall, viruses have evolved a plethora of RNA structures to achieve the same goal: recruitment of ribosomes to the viral RNA for gene expression. Utilizing these unique cap-independent mechanisms have been co-opted as research tools for investigating fundamental biological questions. For example, the streamlined Group IV IRESs have been used to understand movement of the ribosome in relation to tRNAs (Butcher and Jan, 2016). Studying these mechanisms will continue to reveal unique viral gene expression strategies and uncover parallels between canonical and non-canonical translation processes.

1.4 Recoding mechanisms

Genetic decoding, the process of translating genetic material to protein sequences, is not static, but dynamic. Nature continues to surprise us with unconventional forms of decoding termed 'recoding' (REprogrammed deCODEing). Recoding refers to the active or passive change to the information stored in an organism's genetic code on a global (codon reassignments) or local (codon redefinition) scale. Moreover, recoding extends to translational

processes with mechanisms that disrupt decoding of nucleotide triplets (e.g. ribosomal frameshifting and translational bypassing). As this thesis focuses on an unusual viral mechanism for expanding its genomic capacity it is prudent to review some aspects of recoding. Multiple recoding mechanisms have been observed in nature (reviewed in (Baranov et al., 2015)), however, only ribosomal frameshifting and translational bypassing will be discussed here.

1.4.1 Programmed ribosomal frameshifts

Ribosomal frameshifting is an especially common form of viral decoding. As viruses have limited coding capacities due to constraints on genome size, frameshifting allows for the expression of multiple proteins giving rise to proteoforms with alternative C-termini and different functional properties. Often 'programmed' ribosomal frameshifts (PRF) occur at specific sites, implying that the distinct properties or sequences in the mRNA trigger the frameshift. PRFs occur during the elongation of a ribosome on the mRNA template whereby it is displaced in the translational reading frame either in the 5' (-1 frameshift) or 3' direction (+1 frameshift)(Figure 1.7). Frameshift elements are critical in regulating protein translation in the life cycle of numerous viruses. Specifically, in retroviruses such as HIV-1, the expression of the RNA-dependent RNA polymerase is reliant on a frameshift signal in the viral genome (Jacks et al., 1988a; Jacks et al., 1988b; Jacks and Varmus, 1985). Through a -1 PRF, the upstream Gag protein is translationally fused to the overlapping downstream Pol sequence generating Gag-Pol that functions in viral assembly (Brierley and Dos Ramos, 2006). The incidence at which the frameshifting occurs largely impacts the ratio of viral structural and non-structural proteins: any deviation from the optimal ratio results in decreased viral loads (Shehu-Xhilaga et al., 2001). In-depth characterization of retroviral and other frameshift elements has helped to define the

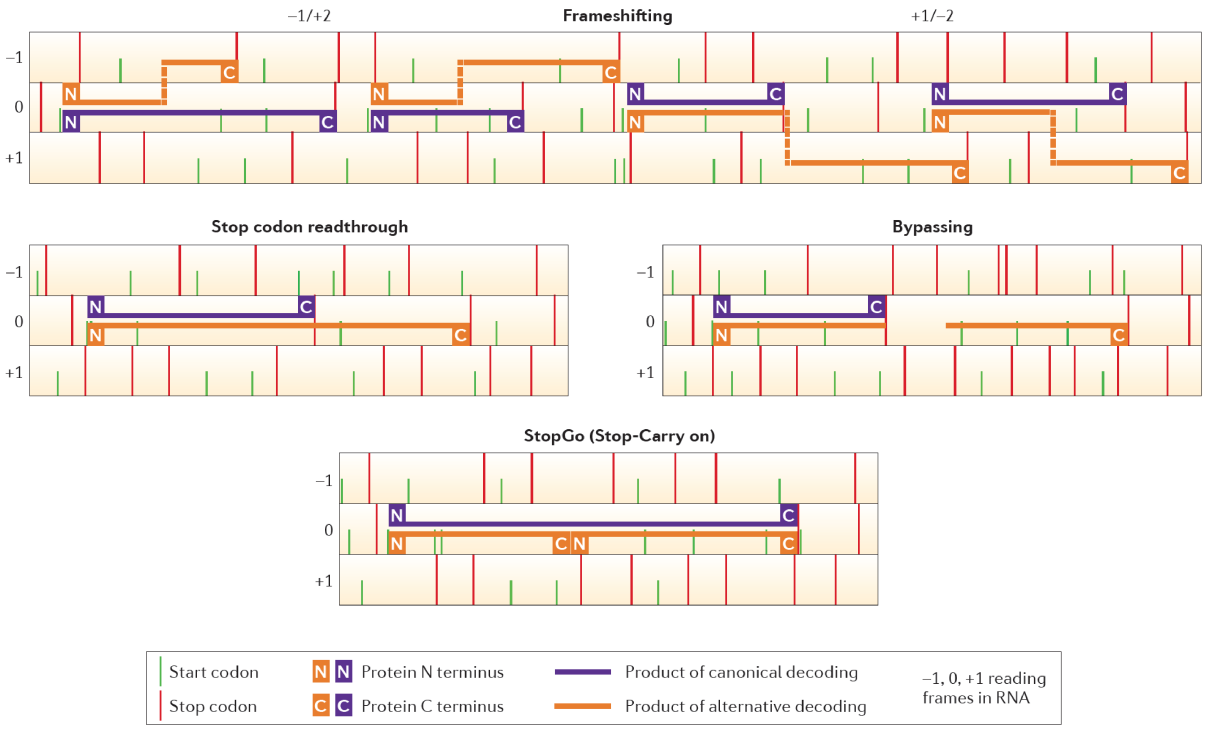


Figure 1.7 Example mechanisms of recoding.

Nucleotide sequences are displayed as horizontal lines signifying three open reading frames. Start codons and stop codons are depicted as green and red vertical lines, respectively. Canonically translated protein sequences are displayed as thick purple lines, while non-canonically translated protein sequences are shown as thick orange lines. The N- and C-termini are labeled. Transitions between reading frames are indicated with a dashed line. Adapted with permission from (Baranov et al., 2015).

parameters necessary for -1 PRFs. Each element involves a slippery heptanucleotide motif X_XXY_YYZ, where X, Y, and Z are three distinct nucleotides, and a stimulator located six to eight nucleotides downstream that is typically a pseudoknot or hairpin (Brierley et al., 1992; Kim et al., 1999; Yu et al., 2011).

Programmed ribosomal frameshifts are not restricted to shifts in the 'minus' or 5' end direction. Plus one (+1) frameshifts have be described in the regulation of cellular mRNAs and yeast transposable elements. Much like in retroviruses, the yeast Ty1 and Ty3 retrotransposons generate a Gag-Pol fusion protein that results from a programmed +1 frameshift. Ty1 frameshifting occurs at a CUU_AGG_C sequence that consists of overlapping Leu codons in the 0- and +1-frame. Slippage of the tRNA at the two leucine codons forces the ribosome to shift into the next frame and continuation of translation occurs after re-pairing of the tRNA with the +1-frame Leu codon (Belcourt and Farabaugh, 1990). In the case of Ty3, the PRF occurs at the GCG_AGU_U minimal element and does not involve tRNA slippage. Instead, the presence of a rare 0-frame serine codon (AGU) causes the ribosome to stall and the P site peptidyl-tRNA obstructs the neighboring 3' base, forcing decoding of the +1-frame codon (Farabaugh et al., 1993).

Frameshifting represents a powerful mechanism for the regulation of gene expression and to increase the coding capacity of an RNA message. As bioinformatics algorithms are further refined, frameshifting elements are found to be increasingly widespread throughout viruses and cellular mRNAs. For example, a -2 PRF element was recently discovered in arteriviruses that permits the synthesis of an alternative viral nsp 2 protein required for viral replication (Fang et al., 2012). Frameshifting, however, is not limited to RNA sequence motifs or structures. Recent reports have demonstrated that viral and cellular proteins can *trans*-activate PRFs in viral

genomes (Li et al., 2014b; Napthine et al., 2017; Napthine et al., 2016). During EMCV infection it was found that there is a temporal increase in frameshifting as infection progresses. Specifically, the viral protein 2A binds to a stem-loop structure 3' of the slippage site to induce a -1 frameshift with efficiency up to 70% (Napthine et al., 2017). Determining the features of RNA elements and proteins that mechanistically facilitate frameshifting will undoubtedly aid in our understanding of viral infection and development of new antiviral targets in addition to answering fundamental questions in translation.

1.4.2 Translational bypass

Perhaps one of the strangest and most unique mechanisms of recoding, translational bypassing could be thought of as a stochastic +50 frameshift. Unlike ribosomal shunting that is a form of translation initiation, translational bypass occurs with actively elongating ribosomes. Discovered in 1988, the most comprehensively described instance of this recoding mechanism occurs in bacteriophage T4 gene *60* (Huang et al., 1988). Here, ribosomes translate the first 45 codons of gene *60* up to a glycine GGA codon. Fifty percent of the translating ribosomes then halt at the adjacent stop codon; however, the other half then skips 50 nucleotides (nt) downstream and resumes translation at a downstream Gly codon (Maldonado and Herr, 1998). In an incredible case of molecular acrobatics, the peptidyl-$tRNA^{Gly}$ disengages from the mRNA at the stop codon ('take-off'), the ribosome jumps the 50 nt gap landing at the downstream GGA codon (the 'landing site') where the peptidyl-$tRNA^{Gly}$ re-pairs to the mRNA (Herr et al., 1999; Herr et al., 2000). Translation then continues to create a single polypeptide chain (Figure 1.7).

Extensive analysis both *in vitro* and *in vivo* has identified essential requirements and elements for bypassing to occur in gene *60*: (1) the $tRNA^{Gly}$ and the matching GGA codons at

the take-off and landing sites flanking the 50 nt 'gap', (2) an upstream nascent peptide signal that potentially interacts with the ribosomal exit tunnel, (3) a 3' stem-loop structure (bypassing tetraloop) that contains the take-off GGA codon and adjacent stop codon, and (4) a Shine-Dalgarno-like sequence 6 nt upstream of the landing site to ensure proper re-initiation (Herr et al., 2000). Interestingly, all the determinants for bypassing are specified in the mRNA itself. Until recently there was no mechanistic understanding of how the ribosome traverses the gap and resumes translation downstream; however, using single molecule fluorescence resonance energy transfer methodology, Chen *et al.* (2015a) discerned movements of the ribosome in relation to the mRNA that drive bypassing. In this model, as the ribosome approaches the GGA take-off codon, the $tRNA^{Gly}$ is delivered while the 5' stem-loop of the *gene 60* mRNA is refolded. The ribosome remains in a rotated state after peptide bond formation and the nascent peptide interacts with the exit tunnel 'pulling' on the peptidyl-tRNA. This, combined with the propensity of the bypassing tetraloop to re-fold, drives the ribosome to slip forward and leads to un-coupled translocation. This allows the bypassing tetraloop to completely refold and since the 5' stem-loop blocks backward movement the ribosome is launched forward. The ribosome then scans forward looking to resume translation (Chen et al., 2015a).

The complexity and number of sequence elements required for efficient bypassing detracts from the likelihood of its existence throughout nature. However, recent studies have identified bypassing elements in bacteriophages of *Streptomyces* species and in the mitochondria of the yeast *Magnusiomyces capitatus* termed 'byps' (Lang et al., 2014; Smith et al., 2013). Despite similarities to gene *60* between features at the primary and secondary structure level, when byps are expressed in *Escherichia coli* no bypassing occurs. This may reflect specialized features of the *M. capitatus* translational machinery that are necessary for bypassing to occur.

Whether bypassing elements are functional in other organisms or not remains unambiguous. However, given these examples it is apparent that recoding is a dynamic process and investigation into these mechanisms will undoubtedly challenge the ways in which we interpret genomic sequence and protein-coding information.

1.5 Mechanisms of non-enveloped RNA virus release

Following gene expression and successful host takeover, assembly and release of the viral particles are the next steps in fulfilling the replicative cycle of viruses. The release of infectious progeny virions is critical to the transmission to naïve cells as well as between host organisms. Indeed, viruses utilize a plethora of mechanisms to egress; however, each one traditionally falls into either a lytic or non-lytic release pathway. As this thesis is focused on non-enveloped RNA viruses, this section will provide a brief overview on the predominant pathways that these viruses manipulate in the host cell to liberate virions.

1.5.1 Lytic release

Cell lysis is the primary pathway by which non-enveloped viruses exit the host cell. It consists of a disruption of cellular membranes, leading to cell death and the release of cytoplasmic compounds into the extracellular milieu. Lysis is actively induced by many viruses as cells seldom trigger it on their own. In general, the mechanisms of how RNA viruses disrupt cell membranes and enact lysis is not well understood; however, there exist a few cases for which a mechanism has been deciphered.

Some eukaryotic lytic viruses utilize specialized proteins, termed viroporins, to induce membrane permeability and disrupt cellular membrane structures (Carrasco, 1995). Viroporins are small 60-120 amino acid proteins that contain a highly hydrophobic domain capable of

forming an amphipathic helix. Insertion into the host membrane leads to oligomerization forming a hydrophilic pore that functions as channel for ions or small molecules (reviewed in (Nieva et al., 2012)). As well, some viroporins contain additional hydrophobic regions that interact with membranes (or a membrane) and may disturb the organization of lipids in the bilayer. The picornavirus 2B protein serves as a model to study viroporins in non-enveloped viruses. Encoded by poliovirus and coxsackievirus, 2B oligomerizes as dimers and tetramers in the membrane and enhances membrane permealibilization (Agirre et al., 2002; van Kuppeveld et al., 2002). In addition to permeabilizing membranes, 2B functions in intracellular membrane remodeling that leads to disruption of vesicle and glycoprotein trafficking (Cho et al., 1994; Jackson, 2014).

Alternative methods of lytic release include the use of polyhedral occlusion bodies or potentially lytic phospholipids. Cypoviruses are dsRNA viruses belonging to the family *reoviridae* that infect insects. At late stages during infection, these viruses induce the formation of crystalline polyhedral structures that embed several thousand viral particles (Coulibaly et al., 2007). The exact mechanism of how cypoviruses induce lysis upon polyhedral formation is unclear. The concept of lytic phospholipids has been observed in dsDNA phycodnaviruses that infect phytoplankton. Phycodnaviruses encode a cluster of biosynthetic genes that are involved in the synthesis of viral glycosphingolipids which subsequently induce programmed cell death (Vardi et al., 2009). Sphingolipids, particularly ceramides, serve as potent inducers of apoptosis in animals and plants (Hannun and Obeid, 1995; Liang et al., 2003). Therefore, since many non-enveloped RNA viruses utilize and interact with host cell lipids/membranes it is tempting to think that they too manipulate this pathway to trigger viral release (Konan and Sanchez-Felipe, 2014).

1.5.2 Non-lytic release

Viruses are categorized into enveloped or non-enveloped classes depending on the presence or absence of a lipid membrane surrounding the capsid. Enveloped viruses escape the host cell via budding from distinct host cell membranes such as the ER, Golgi, or plasma membrane (e.g. hepatitis C virus, hantavirus, and influenza virus, respectively). On the other hand, as discussed in section 1.5.1, non-enveloped viruses have predominantly been thought to trigger lysis for viral release. However, evidence demonstrating the non-lytic spread of classical 'naked' viruses continues to accumulate, challenging this paradigm.

Early observations that non-enveloped viruses could spread through a non-lytic pathway came from studies examining cell lines persistently infected with poliovirus (Colbere-Garapin et al., 1989; Lloyd and Bovee, 1993). Consistent with these observations, poliovirus was found to be released exclusively from the apical surface of polarized intestinal epithelial cells with no apparent cell lysis. This prompted the authors to hypothesize that poliovirus was released through a 'vectorial secretory mechanism', which was supported by electron micrographs of the virus associated with vesicles on the apical surface (Tucker et al., 1993). The idea of non-lytic release of picornaviruses serves as an attractive hypothesis for multiple viruses in this family. Theiler's virus spreads from infected neurons to surrounding glial cells in mice harboring neurons highly resistant to axonal degradation (Roussarie et al., 2007). Coxsackievirus B3 (CVB3) can be transmitted between cells in the presence of neutralizing antibodies (Paloheimo et al., 2011). Lastly, HAV does not lyse infected cells in tissue culture or in humans like most picornaviruses (Lemon, 1985). In fact, HAV particles isolated from tissue culture cells exhibited three distinct buoyant densities after centrifugation suggest that different populations of the virus could exist (Lemon et al., 1985). Indeed, from the initial observations, HAV, poliovirus, and

CVB3 have all been found to exit host cells non-lytically by varying mechanisms (Bird et al., 2014; Chen et al., 2015b; Feng et al., 2013; Jackson et al., 2005).

For some time now, poliovirus has been found to modulate the host autophagy pathways during infection. Originally it was presumed that this pathway was involved in the re-structuring of cellular membranes for viral replication. However, knockdowns of critical autophagosomal components, such as LC3 and ATG12, only modestly decreased intracellular viral loads. In the same study however, depletion of LC3 or ATG12 led to drastic decreases in extracellular virus (Jackson et al., 2005). This was followed by the detection of extracellular poliovirus at early time points during infection (Taylor et al., 2009). It was not until live-cell imaging was combined with single cell analysis that poliovirus was confirmed to spread to cells through a non-lytic mechanism dependent on autophagic vesicles (Bird et al., 2014). It was proposed that poliovirus exits the cell in one of two mechanisms after being engulfed by a double-membraned autophagic vesicle: (1) the inner membrane degrades and the outer membrane fuses with the plasma membrane, releasing the virus, or (2) the double-membraned vesicle fuses with the plasma membrane discharging virions encased in a microvesicle (Bird et al., 2014). Supporting the second model, poliovirus and CVB3 have been observed to be released from cells in microvesicles (Chen et al., 2015b; Robinson et al., 2014). Both viruses were found to be released *en bloc* in phosphatidylserine-rich vesicles and are infectious in naïve cells contingent on the viral receptor being present. Multiple copies of viral particles or RNA within a given vesicle is thought to enhance the infectivity through increasing genetic diversity (Chen et al., 2015b).

Different from its related picornavirus family members, HAV exploits the host exosome pathway to acquire an envelope and mediate non-lytic release (Figure 1.8)(Feng et al., 2013). Exosomes are small vesicles (30-100 nm) that are derived from endosomes. They are released by

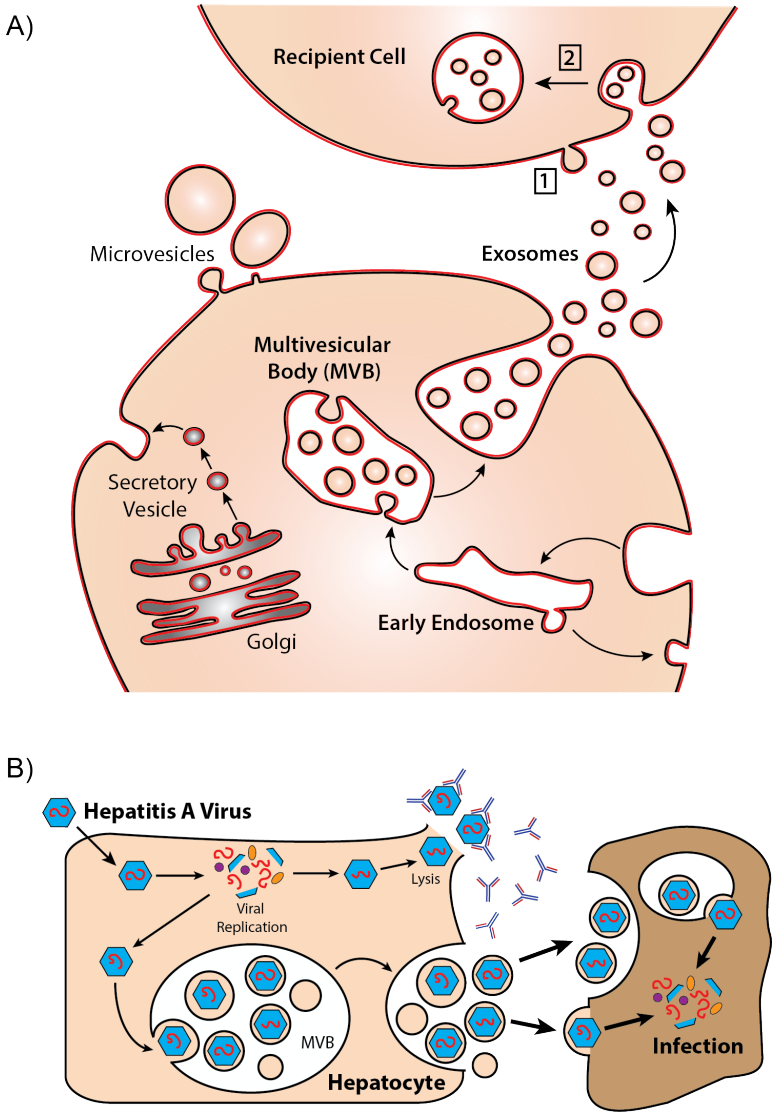


Figure 1.8 Exosome biogenesis and HAV hijacking.

(A) Exosomes derive from invagination of an endosomal membrane to form multivesicular bodies (MVBs). MVBs contain intraluminal vesicles (i.e. exosomes) that hold cargo of ubquitylated cytoplasmic and membrane proteins, mRNAs, micro RNAs, and metabolites. MVBs by Rab proteins where it fuses with plasma membrane, releasing the exosomes. After which, exosomes are delivered to recipient cells via direct membrane fusion (1) or receptor-mediated endocytosis (2). **(B)** hepatitis A virus (HAV) typically exits the cell in a lytic manner that deems the virus non-enveloped. However, a subset of virus can commandeer the exosome pathway to acquire an envelope and infect adjacent cells in the body. This exosome-derived envelope protects the virion from neutralizing antibodies.

all mammalian cell types and carry active enzymes, mRNAs, and regulatory micro RNAs for transport to adjacent cells (Figure 1.8). Over the past years, the role of exosomes has become increasingly evident in viral pathogenesis (reviewed in (Meckes and Raab-Traub, 2011)). Exosomes are generated after invagination of membrane into the lumen of a late endosome. This produces a multi-vesicular body (MVB) that is trafficked, with the help of Rab proteins, to the plasma membrane where it fuses with it, releasing the vesicles into the extracellular milieu. The most well understood mechanism for biogenesis of the vesicles is through the endosomal sorting complex required for transport (ESCRT)(reviewed in (Colombo et al., 2014)). Over thirty proteins that compose four main complexes (ESCRT-0, -I, -II, and -III) work together to facilitate membrane deformation, cargo recruitment, and vesicle scission (Figure 1.9). ESCRT-0 works to recruit ubiquitylated proteins to the endosomal membrane, while the component HRS recruits the ESCRT-I complex via TSG101. ESCRT-I and –II act to deform the endosomal membrane. Finally, ESCRT-III is recruited through ESCRT-II or ALIX and facilitates vesicle scission. The entire complex is then disassembled and recycled by the AAA-ATPase VPS4 (Figure 1.9)(Colombo et al., 2014). The ESCRT pathway has shown to be important for the life cycle of numerous viruses. For example, HIV-1 usurps this pathway to aid in viral budding from the host cell plasma membrane. HIV-1 encodes 'late' domains in the viral protein p6. Late domains are amino acid sequence motifs characterized by P(S/T)XP, $YPX_{1-3}L$, or PPXY that enable proteins to bind TSG101 or ALIX (Freed, 2002). The HIV-1 late domain in p6 allows the virus to bind TSG101 and recruit the ESCRT machinery for budding (Fujii et al., 2007). Interestingly, HAV encodes two tandem 'late' domains in the viral capsid protein VP1. Disruption of these domains perturbs hijacking of the exosomal pathway and development of

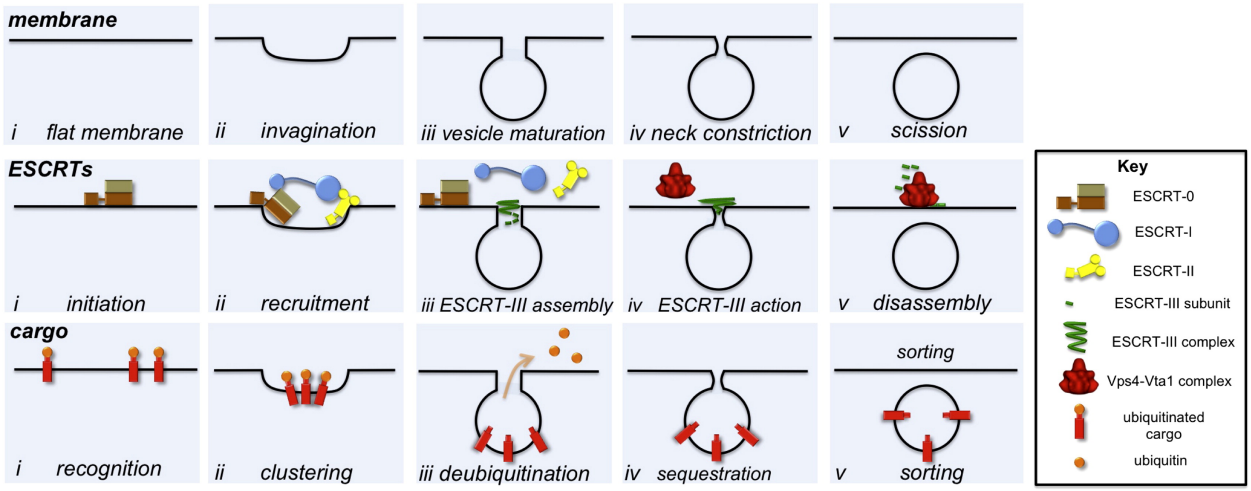


Figure 1.9 ESCRT-mediated vesicle biogenesis.

Each row of panels represents the perspective of the distinct components of vesicle formation: the membrane, ESCRT machinery, and cargo. ESCRT-0 initiates the pathway by engaging ubquitylated cargo (i). Clustering of the cargo following binding of ESCRT-I and –II results in a ESCRT-cargo-enriched zone (ii). ESCRT-II nucleates the assembly of ESCRT-III which triggers recruitment of de-ubquitylation machinery (iii). The vesicle neck is constricted and cargo is sequestered through the action of ESCRT-III (iv). Final budding and scission occurs after ESCRT-III complex disassembly by the AAA-ATPase, VPS4 (v). Reproduced with permission from (Henne et al., 2011).

enveloped HAV. Thus, it is presumed that HAV utilizes these domains to bind ALIX and facilitate sorting into exosomes (Feng et al., 2013).

Why non-enveloped viruses invoke non-lytic release pathways during infection is not well understood. In the case of HAV, enveloped virions are found to circulate in the blood of infected patients while naked virions reside in the stool (Feng et al., 2013). It may be that adopting a biphasic viral lifecycle is evolutionary advantageous. Intrinsic advantages exist to both enveloped and non-enveloped virions. Envelopes provide a means to readily manipulate the contents of the lipid shell allowing for flexibility in receptors and evasion of the immune response. Capsids, on the other hand, offer a robust protein cage that protects against environmental hazards such as desiccation and detergents. Whatever the reason for evolving the ability to acquire an envelope, understanding the mechanisms behind viral egress for both enveloped and non-enveloped viruses is critical for answering fundamental questions and developing novel antiviral therapies.

1.6 The *Dicistrovirdae* family of viruses

Originally classified as Picorna-like viruses, the *Dicistroviridae* are a rapidly expanding family of (+)ssRNA viruses within the order *Picornavirales* that infect arthropods. At the time of writing, *Dicistroviridae* contains 15 viruses classified into three genera: *Cripavirus*, *Aparavirus*, and *Triatovirus* (Adams et al., 2017). Classification within each genus is determined through phylogenetic distances and distinct differences in the intergenic region IRES (discussed more in Section 1.4.3.2). Notable members of the *Dicistroviridae* include Cricket paralysis virus (CrPV) and Drosophila C virus (DCV) both of which infect the genetically amenable *Drosophila* and have provided excellent models of virus host interactions in insect cells. The type species, CrPV,

was first isolated in 1970 from Australian field crickets (Bonning and Miller, 2010; Reinganum, 1970). Consequently, many studies have used CrPV as a model to delineate translational control mechanisms and antiviral immune mechanisms in insects, such as the RNAi and Imd (Immunodeficiency) pathways (Sabin et al., 2010). Furthermore, the unique dicistronic organization of its genome has led to insights into an unprecedented mechanism of translation initiation. The IGR IRES contains domains that functionally mimic a tRNA to recruit ribosomes and initiate translation in a factor-independent manner (Costantino et al., 2007; Jan et al., 2003; Pestova and Hellen, 2003; Pfingsten et al., 2006; Wilson et al., 2000a).

Other members of this family have been linked to the health of beneficial organisms. For example, Israeli acute paralysis virus (IAPV) and Kashmir bee virus (KBV) have been linked to the health of honeybees (Cox-Foster et al., 2007b; de Miranda et al., 2010). This is especially vital for crops reliant on honey bee pollination, which have an estimated worth of US$215 billion worldwide (Gallai et al., 2009). In addition, as the causative agent of Taura syndrome in panaeid shrimp, Taura syndrome virus (TSV) has devastated the shrimp farming industry throughout the Americas (Bonami et al., 1997; Lightner and Redman, 1998). Dicistroviruses can also be potentially advantageous. Eight of the currently known dicistroviruses are known to be pathogenic to pests of agricultural or medical importance (i.e. CrPV, ALPV, RhPV, PSIV, HiPV, HoCV-1, SINV-1, TrV). RhPV, ALPV, and PSIV are all pathogens of phytophagous insects that are major issues in crop production due to their feeding habits and transmission of plant viruses. Moreover, the host of TrV is the vector of the protozoan parasite *Trypanosoma cruzi*, which is the etiological agent of Chagas disease, afflicting an estimated 8 to 11 million people in Latin America (Muscio et al., 1987). Due to growing amount of invasive species being introduced into the United States, dicistroviruses have been considered for use as bio-pesticides. For example,

the red fire ant, *Solenopsis invicta,* causes an estimated $5 billion USD in annual costs due to crop and livestock losses, control measures, and medical treatment of victims. Solenopsis invicta virus 1, which was identified in a sequence library of *S. invicta* and with infection rates as high as 88% in all developmental stages, has been proposed as a potential method to control the spread of *S. invicta* (Hashimoto and Valles, 2007; Valles et al., 2004; Valles et al., 2008).

Lastly, dicistroviruses and discistro-like viruses have been identified in metagenomic surveys in from marine environments, invertebrates, and even the human gut (Culley et al., 2003; Shi et al., 2016; Victoria et al., 2009). Therefore, dicistroviruses are likely to be ubiquitous and exist in vast quantities in the environment. Taken together with their impact on economically and medically important arthropods, understanding their biology is vital.

1.6.1 Characteristics and Genome organization

Genomes in the *dicistroviridae* are monopartite, linear (+)ssRNA and range from 8 to 10 kb in size. Each genomic RNA copy contains a 5' viral protein cap (VPg), and a 3' poly(A) tail (Figure 1.10)(Bonning and Miller, 2010). The origin of the name for this viral family derives from its unique dicistronic genome arrangement consisting of two non-overlapping open reading frames. In addition to having two ORFs, a key distinguishing feature from the *picornaviridae* family is the location of structural proteins encoded at the 3' end of the dicistrovirus genome. Nonetheless, comparison between the two families reveals significant similarities at the amino acid level; e.g., in the chymotrypsin-like viral proteases (3C) and the super group type I RNA-dependent RNA polymerases.

Distinct IRESs drive translation of each open reading frame. The 5'UTR IRES directs translation of ORF1, which encodes viral non-structural proteins such as the suppressor of RNA

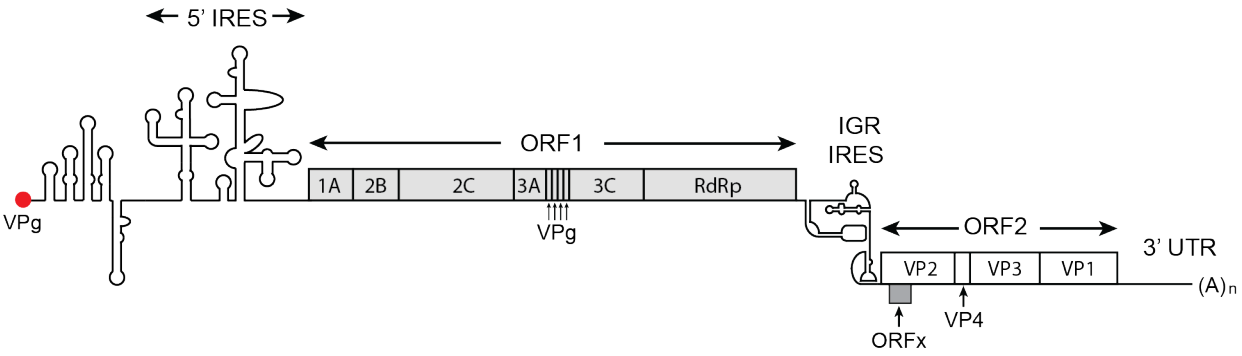


Figure 1.10 Genome organization of dicistroviruses.

Displayed is the genome organization of dicistroviruses using CrPV as a model. The genome is capped at the 5' end by VPg along with a 3' poly(A) tail. Translation of the two non-overlapping ORFs is differentially regulated by two distinct internal ribosome entry sites. The 5'UTR IRES facilitates translation of ORF1, which encodes the viral non-structural proteins including the suppressor of RNAi (1A), helicase (2C), protease (3C), and RNA-dependent RNA polymerase (RdRp). Translation of the 3' proximal ORF is driven by the intergenic region IRES and encodes the viral structural proteins, VP1 to VP4. In a subset of dicistroviruses, namely the honey bee and fire ant viruses, there exists a hidden overlapping ORF termed, ORFx.

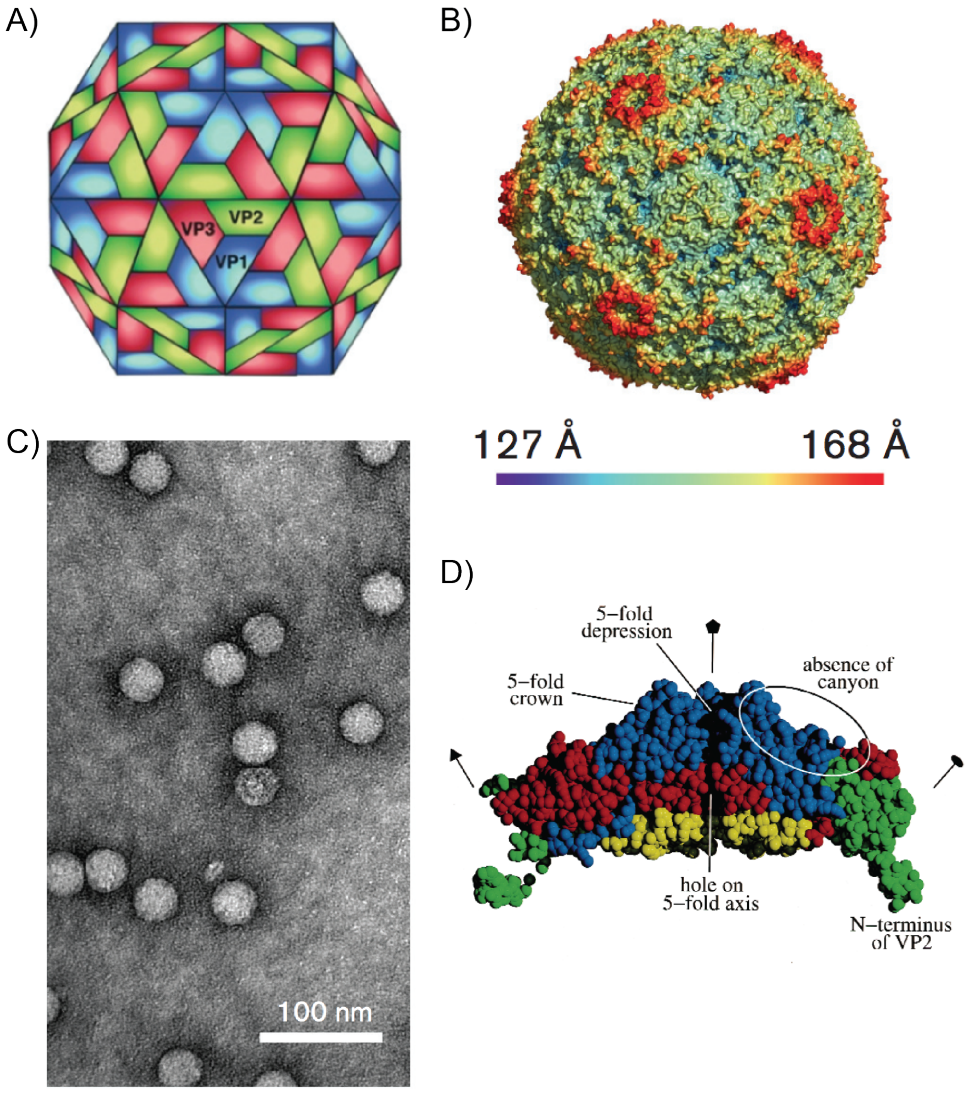


Figure 1.11 Morphology and structure of the dicistrovirus virions.

(A) Illustration demonstrating the packing surface of the viral structural proteins VP1, 2, and 3. VP4 is located on the inner surface of the capsid. **(B)** Rendered X-ray crystal structure of the full CrPV virion. Surface colour depicts the distance from the center of the virion. **(C)** Negative stain electron micrograph of the Triatoma virus. **(D)** Cross-section of a CrPV pentamer. VP1, 2, 3 and 4 are shown in blue, green, red, and yellow, respectively. The major interactions are formed by VP1 along the 5-fold axis. The N-terminus of VP3 coils around the axis beneath VP1 to form a rigid β-cylinder. VP4 forms a disk-like structure beneath VP3 at the interface of the protein shell and encased viral RNA. All figures are reproduced with permission from (Adams et al., 2017) and (Tate et al., 1999).

interference (RNAi), RNA helicase, 3C protease, and the RdRp. Interestingly, the arrangement of this ORF mimics the organization of non-structural proteins in picornaviruses; from N- to C-termini the order in both families is the superfamily 3 helicase (2C), VPg for priming viral replication, the 3C protease, and lastly the RdRp. The second open reading frame (ORF2) encodes viral structural proteins and is translated via the IGR IRES (Bonning and Miller, 2010).

The X-ray crystal structure of CrPV shows that VP1, VP2, and VP3 assemble to form the characteristic icosahedral capsid while VP4 resides on the inside (Figure 1.11). VP1-3 each have a jellyroll core composed of eight beta sheets that assemble to form the capsid protomer, similar to picornaviruses. Sixty copies of each structural protein assemble to form the full viral capsid, roughly 30 nm in size. VP4 is packaged as a fusion protein with the N-terminus of VP3 and cleavage of the fusion protein is thought to mature the virion (Tate et al., 1999).

Individual differences in the capsid proteins likely contribute to the wide host and tissue tropism for dicistroviruses. To date, each dicistrovirus has been found to infect a single order of host (e.g. *Diptera* or *Hymenoptera*), apart from CrPV which can infect over 20 species in 5 different orders (Bonning and Miller, 2010). Recent work examining the tissue tropisms of DCV and CrPV in a *Drosophilia melanogaster* host has revealed insights into the pathogenesis of these viruses. While both viruses infect the midgut, it was discovered that DCV, but not CrPV, infects the smooth muscles surrounding the crop. This leads to intestinal obstruction, accumulation of food in the crop of fruit flies, and eventual death of the host (Chtarbanova et al., 2014). Conversely, CrPV infection of adult flies paralyzes them, eventually leading to death through starvation or dehydration (Reinganum, 1970).

1.6.2 *Dicistrovirus* virus host interactions

Currently there is very limited knowledge on how dicistroviruses usurp the host cell processes to facilitate infection. In general, the replication of these viruses is thought to occur much like picornaviruses where the genomic RNA serves as both a template for the synthesis of viral proteins and RNA replication. The life cycle of dicistroviruses is on the order of hours; for instance, the CrPV life cycle is 8 to 10 hours in tissue culture cells. How these viruses enter the host is unclear, as the receptor(s) remains to be determined. However, it was shown that DCV enters the cell through a vesicular pathway that is dependent on clathrin-mediated endocytosis (Cherry and Perrimon, 2004). Following entry, the genome uncoats and the RNA is translated through two distinct IRES elements (discussed in more detail below). Non-structural proteins are translated first via the 5'UTR IRES and remain constitutively expressed throughout infection (Khong et al., 2016). These proteins act to facilitate viral replication and disrupt cellular process such as the antiviral response. Specifically, the CrPV protein 1A has a multifunctional role during infection. First, it binds to Argonaute 2 and prevents its ability to slice target RNA, thus inhibiting the antiviral RNAi response in host cells (Nayak et al., 2010b). Secondly, it acts to inhibit the formation of stress granules (SGs) during replication (Khong and Jan, 2011; Khong et al., 2017). Stress granules are RNA-protein aggregations that form under times of cellular stress and exhibit antiviral activity. Rapid shutoff of host translation during CrPV infection through the dissociation of eIF4G and 4E in addition to eIF2α phosphorylation creates a cellular environment amenable to SG formation. SGs formed early during infection may sequester nascent viral RNA or protein that slows replication. Indeed, when SG formation is inhibited prior to CrPV infection there is an increase in intracellular viral loads (Garrey et al., 2010; Khong et al., 2017). Akin to picornaviruses, dicistrovirus non-structural proteins facilitate the formation of

a membranous web structure during infection. The exact mechanism of how this occurs is unknown although DCV was shown to remodel the Golgi apparatus into viral replication vesicles in a process that is dependent on the coat protein complex I and fatty acid biosynthesis pathways of the host cell (Cherry et al., 2006). At roughly three hours post infection, synthesis of the structural proteins from the IGR IRES begins. Non-structural proteins translated earlier during infection act to promote the production of structural proteins in supra-molar excess (Khong et al., 2016). Subsequently, progeny virions are assembled and viral release is presumed to occur through disruption of the host cell plasma membrane and lysis.

1.6.2.1 The 5' untranslated region internal ribosome entry site

The dicistrovirus 5'UTR IRESs vary considerably in sequence and size, and to date there have been very few studies examining the mechanism(s) they use to initiate translation. *In vitro* reconstitution experiments with the 5'UTR IRES from RhPV have demonstrated that it has a complete reliance on eIF2, eIF3, and eIF1 for translational activity. This activity is enhanced with the addition of eIF1A, eIF4A, and eIF4G. Nevertheless, the exact mechanism remains undefined as 5'- and 3'-deletions of the UTR still support IRES activity. It has been proposed that it functions through non-specific binding of initiation factors near the AUG initiation codon (Terenin et al., 2005). Recently, a report suggested that the 5'UTR IRES from the dicistro-like Halastavi árva virus (HaIV) may employ a similar mechanism as RhPV. The HaIV 5'UTR IRES does not rely on structural integrity of the RNA like a typical IRES element. Instead the ribosome is recruited to an unstructured region flanking the AUG start codon of the IRES where it undergoes retrograde scanning before initiating translation (Abaeva et al., 2016). In this case, eIF2 and eIF3 are essential for IRES activity while it is enhanced in the presence of eIF1 and

eIF1A. Interestingly, both RhPV and HaIV utilize unstructured sequences proximal to the initiation codon to recruit the ribosome and demonstrate cross-kingdom activity (Abaeva et al., 2016; Royall et al., 2004; Woolaway et al., 2001). This suggests that this mode of translation initiation may be conserved throughout discistroviruses.

Challenging this hypothesis, the CrPV 5'UTR IRES was revealed to be highly structured, have a defined element, and resembles a Group III IRES (Figure 1.6)(Gross et al., 2017). Using SHAPE (selective 2' hydroxyl analysis by primer extension), the CrPV 5'UTR was found to be comprised of three distinct domains. Domain I precedes the IRES element and is dispensable for translational activity: it is proposed that the stem loop structures in this region may be involved in replication of the viral RNA or in assembly of the viral capsid. Conversely, Domains II and III are essential for IRES activity and are arranged into several hairpins with three- and four-way junctions, in addition to a pseudoknot structure. These elements recruit both the ribosome and eIF3 to direct translation initiation at the cognate viral AUG codon without any scanning akin to the HCV IRES (Gross et al., 2017). Additionally, both the HCV IRES and CrPV 5'UTR IRES require the auxiliary ribosomal protein RACK1 (Receptor of Activated Protein C Kinase 1) for function; further supporting the notion that the CrPV 5'UTR IRES falls into the Group III class (Majzoub et al., 2014).

1.6.2.2 The intergenic region internal ribosome entry site

As one of the most well studied and arguably simplest IRES, the IGR IRES has provided a wealth of insights into not only fundamental viral translation strategies but also into general translation mechanisms such as ribosome-tRNA functions and ribosome reading frame selection (Butcher and Jan, 2016). The IGR IRES is ~190 nucleotides in length, is highly structured with

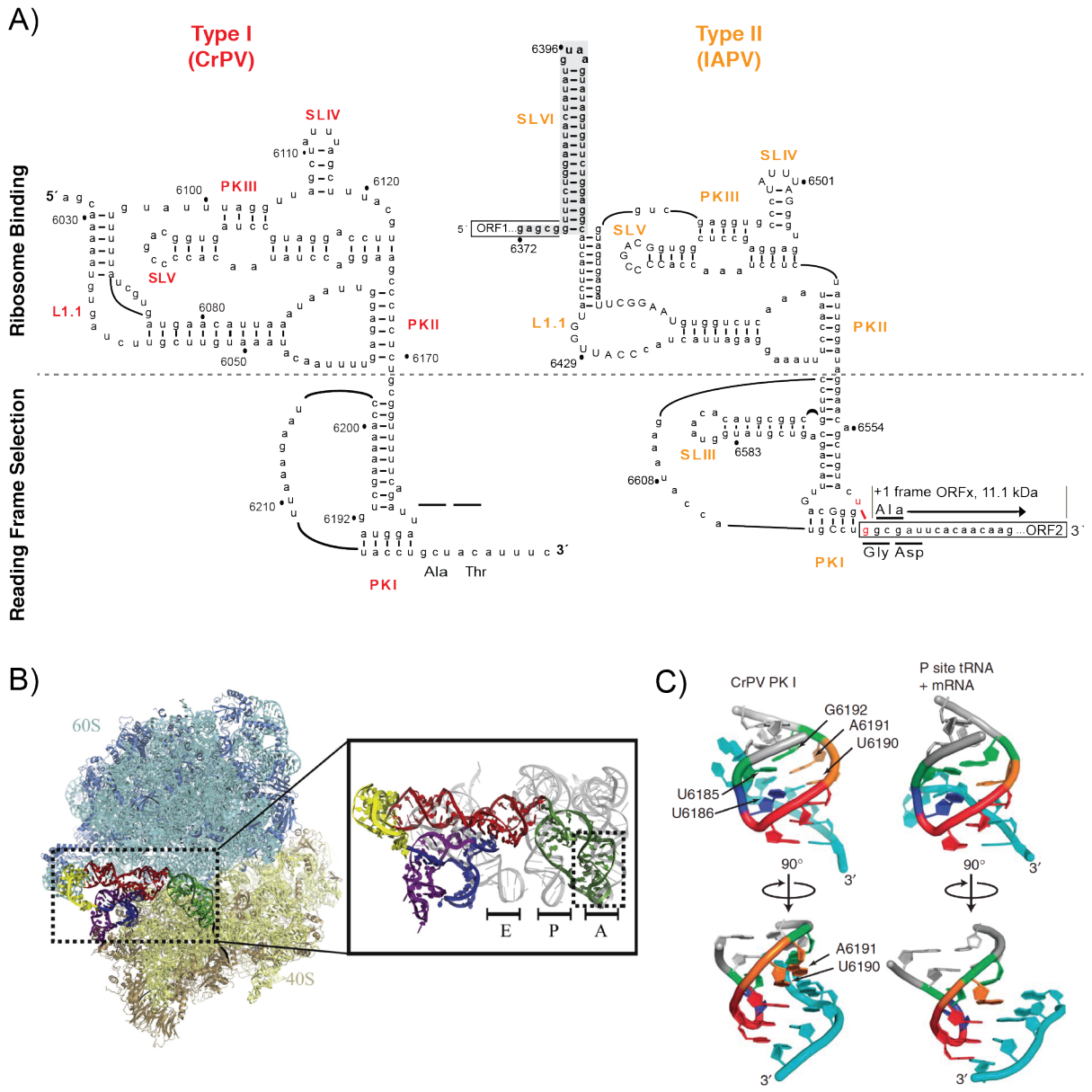


Figure 1.12 The *Dicistroviridae* intergenic region internal ribosome entry site.

(A) The secondary structure of Type I and Type II IGR IRESs from CrPV (left) and IAPV (right) as determined by empirical data. Type II IRESs harbor two additional stem-loops (SLIII and SLVI), a larger L1.1 loop, and lack a 'shoulder' between SLIV and PKII. Each IRES is modular with PKII and III comprising the ribosome binding domain, while PKI is involved in ribosome position and reading frame selection. **(B)** Cryo-EM structure of the CrPV IGR IRES bound to the yeast *Kluyveromyces lactis* 80S ribosome solved at 3.7 Å. The IRES spans the entire ribosome across the A, P, and E sites with PKI (green) occupying the A site, SLIV and V (blue and purple, respectively) binding the 40S subunit, and L1.1 (yellow) and PKII (red) interacting with the 60S subunit. Inset:

superposition of the IRES with A, P, and E site tRNAs displaying the location of PKI in the decoding center. Reproduced with permission from (Fernandez et al., 2014). **(C)** Comparison of the anticodon:codon-like interactions of the CrPV PKI to authentic P site tRNA-mRNA interactions. Analogous bases are highlighted in the same colour. Adapted with permission from (Costantino et al., 2008).

---

multiple RNA:RNA contacts, and is in general conserved throughout the *Dicistroviridae*. The IRES are classified into two subgroups, Type I and Type II, which are best represented by the CrPV and IAPV IRESs, respectively (Figure 1.12). The main distinguishing features between the subgroups include a larger L1.1 loop, a lack of 'shoulder' in the ribosome-binding domain, and an extra stem-loop (SLIII) in tRNA-like domain of Type II IRESs (Figure 1.12). Nevertheless, IRESs in each subgroup adopt a similar fold, thus are modular allowing for chimeric IRESs to be generated (Au and Jan, 2012; Hertz and Thompson, 2011b; Jang and Jan, 2010).

Several biochemical and structural studies employing crystallography and cryo-electron microscopy have revealed that the IGR IRES is comprised of three pseudoknots (termed PKI, II, and III) (Fernandez et al., 2014; Pfingsten et al., 2006; Schuler et al., 2006); PKII and PKIII together form the core ribosome-binding domain that recruits both the 40S and 60S ribosomal subunits, while PKI acts to position the 80S ribosome at the correct initiation codon (Figures 1.12 and 1.13). The most remarkable aspect of the IGR IRES is that it requires no translation initiation factors or initiator Met-tRNA to recruit the ribosome and initiate translation from a non-AUG codon (Figure 1.12) (Jan and Sarnow, 2002; Pestova and Hellen, 2003; Wilson et al., 2000b). The IGR IRES binds with high affinity to the conserved core of the ribosome and spans all three tRNA binding sites (A/P/E sites; Figure 1.12). IRES binding to the ribosome promotes a rotation of the 40S subunit head domain and the formation of latch interactions between helices 18 and 34 of the 18S rRNA near the mRNA channel entrance (Fernandez et al., 2014; Schuler et

al., 2006; Spahn et al., 2004). In an extraordinary case of molecular mimicry, the PKI domain adopts a tRNA-like fold to mimic the anticodon stem of a tRNA, thereby enabling access to the ribosome decoding center (A site; Figure 1.12) (Au et al., 2015; Costantino et al., 2008; Fernandez et al., 2014; Koh et al., 2014). However, PKI strays from the canonical codon:anticodon interaction of a tRNA and instead harbours five base-pairing interactions. After positioning, the IRES undergoes a pseudotranslocation step that involves the movement of the PKI domain to the P site thereby allowing the delivery of the first aminoacyl-tRNA to the non-AUG codon in the A site, essentially setting the ribosome into an elongation competent state (Figure 1.13) (Fernandez et al., 2014; Muhs et al., 2015). By mimicking a fundamental step in the host translation cycle, the dicistrovirus IGR IRESs can entirely bypass the heavily regulated initiation step of translation. Indeed, the IGR IRES can direct translation in eukaryotes, from yeast to human, and in prokaryotes, suggesting a universal translation mechanism across kingdoms (Colussi et al., 2015; Thompson et al., 2001).

How does the IGR IRES manipulate the ribosome? The IGR IRES makes extensive contacts with the 80S ribosome to recruit both the 40S and 60S subunits. Recruitment of the ribosomes to the IRES occurs largely through a sequential pathway where the 40S subunit is assembled prior to 60S subunit joining (Figure 1.13). Although at a significantly lower frequency, the fully assembled 80S ribosome can be recruited to the IRES *in vitro* (Petrov et al., 2016). In canonical eukaryotic translation initiation, it is proposed that the dynamics of the 40S head play an import role in correctly positioning the initiator Met-tRNA (Aylett et al., 2015; Llacer et al., 2015). For the CrPV IRES, conformational flexibility of the 40S head is restricted through insertion of SLIV and SLV in the cleft between the head and body of the 40S subunit where they make contacts with uS7, uS11, and eS25. This conformation allows PKI to bind

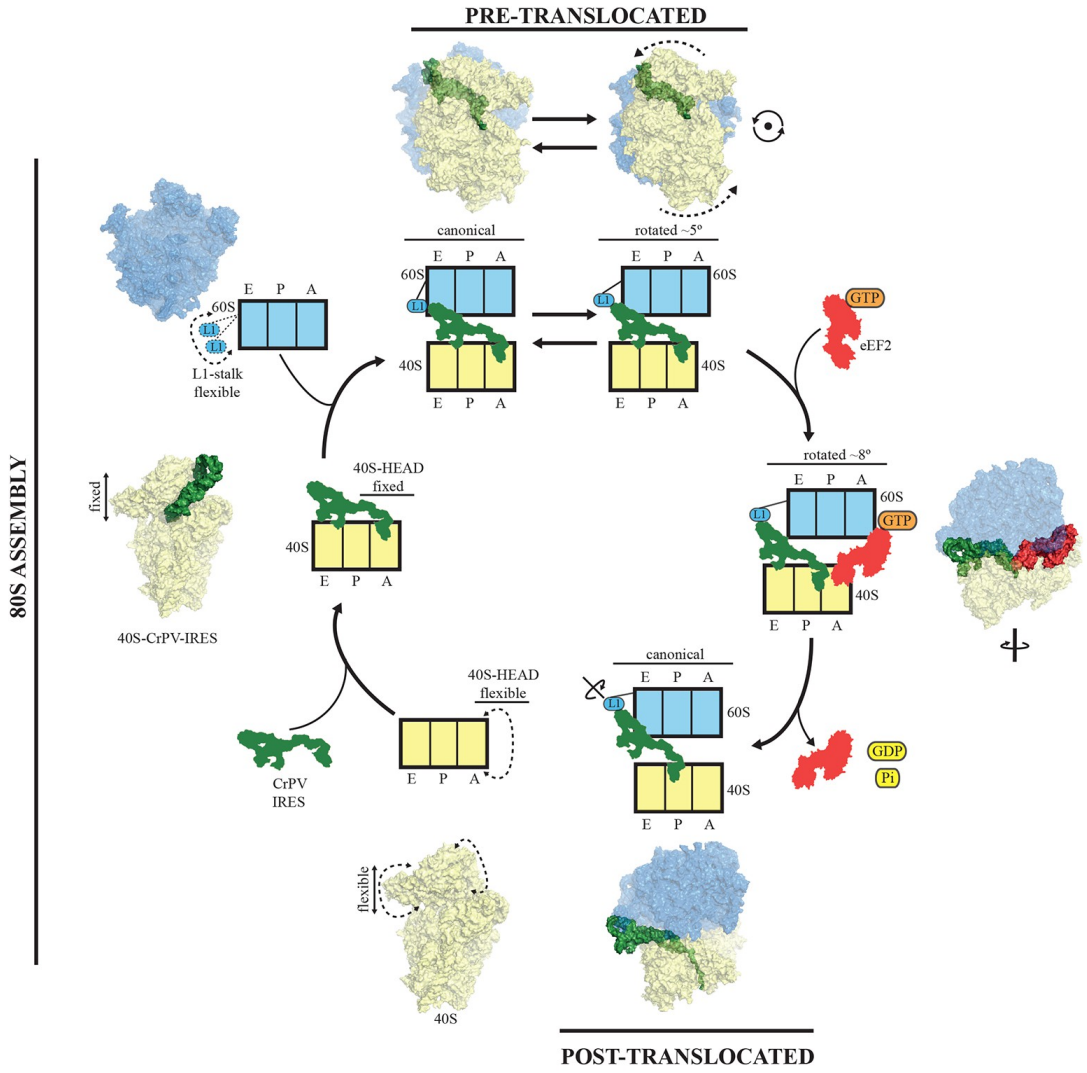


Figure 1.13 Translation initiation pathway of the IGR IRES.

Beginning at the bottom left, the 40S subunit is bound to the IGR IRES through SLIV and V; restricting the conformational flexibility of the 40S head. This allows for decoding of PKI and recruitment of the 60S subunit by binding the L1.1-PKII region. Following this, the binary IRES-80S complexes alternates between canonical and rotated states with PKI residing in the A site (top; 'pre-translocated states'). eEF2 binding triggers an additional ~ 3° rotation of the 40S, forcing the L1 stalk of the 60S subunit into a wider conformation. Domain IV of eEF2 interacts with PKI forming a translocation intermediate, reminiscent of an A/P site hybrid state. GTP hydrolysis induces conformational changes in eEF2 and the ribosome. PKI translocates to the P site while the original contacts made by SLIV and V with 40S are disrupted (bottom; 'post-translocated state'). Reproduced with permission from (Murray et al., 2016).

tightly to the decoding center and mimic interactions that occur during canonical decoding. For instance, the conserved bases A1755 and A1756 in the 18S rRNA interact with the minor groove of the PKI codon:anticodon minihelix (Fernandez et al., 2014; Koh et al., 2014; Muhs et al., 2015; Murray et al., 2016). Despite the striking resemblance to tRNAs, the IRES still maintains specific contacts in the decoding center. In standard decoding, C1273 of the 18S rRNA interacts with the wobble base pair in the tRNA anticodon loop. With the CrPV IRES, a rotation about the glycosidic bond of C1273 allows this nucleotide to engage in tight stacking with the first nucleotide in the initiation codon of the IRES (Abeyrathne et al., 2016; Murray et al., 2016). Intriguingly, except for PSIV, all known IGR IRESs have a guanine residue as the first nucleotide of the initiation codon, suggesting that this base pair may play an important role in interacting with the 18S rRNA.

Upon recruitment of the 60S subunit, pseudotranslocation ensues (Figure 1.13). The binary IRES:80S complex oscillates between a canonical and rotated state that is reminiscent of hybrid state tRNAs (Koh et al., 2014; Muhs et al., 2015; Murray et al., 2016). The rotated state involves a counter-clockwise rotation of the 40S subunit relative to the 60S along with 40S head domain swiveling and movement of the L1 stalk. Since the L1.1 loop of the IRES interacts with the L1 stalk, it couples the IRES to the movements of the ribosome and forces it to adopt multiple dynamic conformations that are directly dependent on the rotational state of the ribosome. In the rotated state, eEF2-GTP binds and induces a further ~3-degree rotation of the 40S subunit. This extra rotation is also a feature observed with canonical tRNAs; however, in the case of the IRES the L1 stalk is pushed to a wider conformation. At this stage, Domain IV of eEF2 contacts PKI of the IRES and stabilizes it in an intermediate A/P site hybrid state. Delivery of the first aminoacyl-tRNA only occurs following pseudo-translocation of PKI to the P site. In

the absence of eEF1A and cognate aminoacyl-tRNA the IRES is vulnerable to a spontaneous back-translocation event unless it is stabilized by an A site ligand. Following GTP hydrolysis, there is a conformational change in eEF2 and the ribosome, PKI translocates to the P site, and the original contacts made by SLIV and SLV with the 40S are broken while interactions with the L1 stalk are maintained. Once in the P site, interactions of SLIV and SLV with uS7 and eS25 are restored (where 'u' and 'e' denote universally and eukaryotic conserved ribosomal proteins, respectively)(Abeyrathne et al., 2016; Ban et al., 2014; Murray et al., 2016). This cycle of conformational states during rotation of the ribosome continues until the IRES has fully translocated through the ribosome and elongation ensues.

Generally, the IRES is thought to mechanistically exhibit an 'inchworm-like' movement through this process where it varies from extended (initiation state), to compact (pre-translocation), and back to extended conformations (post-translocation) as it translocates (Figure 1.13)(Abeyrathne et al., 2016). Recent kinetic studies have shed light into the details of this mechanism (Zhang et al., 2016). They illustrate clearly that the first two cycles of elongation are retarded compared to subsequent steps due to the slow, rate-determining, pseudo-translocation and translocation steps. After displacement of the IRES from the ribosome and translocation of the tripeptidyl-tRNA the pace of peptide synthesis increased substantially (Zhang et al., 2016). Interestingly, there is slow rate observed on the second elongation cycle after the IRES has presumably exited the E site of the ribosome. While the consequences of stalling or slowing are not understood, it may be that this occurrence contributes to positioning of PKI for proper reading frame selection.

1.6.2.3 IGR IRES-mediated alternative reading frame selection

As viral genomes are extremely compact with limited sequence space, viruses have evolved a variety of strategies to expand the coding capacity. Recently, bioinformatic analyses identified an +1-frame overlapping open reading frame within a subset of dicistrovirus genomes (Figure 1.10)(Firth et al., 2009; Sabath et al., 2009). A subset of IGR IRESs have been shown to facilitate the translation of this +1-frame overlapping ORF, termed ORFx, in addition to translation in the 0 frame (Au et al., 2015; Ren et al., 2012). Translation of +1-frame ORFx occurs at ~20% of that of ORF2 in IAPV and initiates through a U:G wobble base-pair directly adjacent to the IRES translational start site. Specific mutations within the tRNA-like PKI domain uncouple 0 and +1 frame translation, that is correlated with distinct, subtle conformations, which may drive reading frame selection (Au et al., 2015; Ren et al., 2014). Unlike most viral RNA signals that shift reading frames of translating ribosomes, such as -1 frameshift signals in retroviruses and coronaviruses, IGR IRES-mediated 0 and +1 frame translation represents a novel translation initiation strategy to increase the coding capacity of a viral genome. Normally, translation is highly accurate and the ribosomal reading frame is maintained. As the IGR IRES adopts a tRNA-like domain that recruits the ribosome directly and sets it into a specific reading frame, it provides a unique yet powerful model for not only elucidating RNA-driven recoding strategies in virus infections but also into fundamental mechanisms of ribosome reading frame maintenance.

1.6.2.4 Significance of the IGR IRES

While the IGR IRES serves as a powerful model for other viral IRESs, it has also proven to be an important tool for studying other aspects of translation. Since the IGR IRES bypasses

the canonical translation initiation process, which would otherwise require at least 9 core translation factors, the IGR IRES can be exploited in a simple purified system for studying translation elongation and termination and its initiation factor-independent nature permits investigations into whether a process is initiation factor-dependent. For example, the use of the IGR IRES has provided significant insights into miRNA-dependent translational control and nonsense-mediated mRNA decay (Isken et al., 2008; Petersen et al., 2006). It is anticipated that IGR IRES translation can be utilized to probe many aspects of translational regulation.

Studies of the IGR IRES mechanism have also provided new insights into ribosome heterogeneity. As mentioned above, ribosomal protein eS25 interacts with a specific IGR IRES domain to mediate function (Landry et al., 2009; Muhs et al., 2011). However, it has also been shown that when other ribosomal proteins are depleted (e.g. eS6) there is a specific block in IGR IRES-dependent translation but not overall cap-dependent translation, thus hinting that the ribosome may have specificity for specific mRNAs (i.e. ribosome filter hypothesis) (Cherry et al., 2005). Indeed, this process extends past IGR IRES translation. Depletion of ribosomal proteins also affects other viral and cellular IRESs such as the more complex hepatitis C virus IRES (Landry et al., 2009). In addition to ribosomal proteins, the IGR IRES has also expanded our understanding of the role of pseuoduridylation (pseudoU), an RNA modification within the rRNA (Jack et al., 2011). PseudoU of rRNA is highly specific and conserved, however, the exact biological impact was elusive. Jack *et al* demonstrated that lack of pseudoU resulted in reduced ribosome binding to the IGR IRES and a decrease in IGR IRES-mediated but not cap-dependent translation. These experiments suggest that ribosomes may be heterogeneous in nature and may be regulated via RNA modification and/or post-translational modification to modulate translation of specific mRNA. It is likely that the regulatory effects of ribosomal proteins and modifications

extend beyond IRES-mediated translation, however this has given an initial hint into their biological roles.

Overall, many different RNA structures have defined functions to facilitate viral infection. Remarkably, the ~190 nucleotide IGR IRES is an unprecedented minimal RNA structure that can hijack a 3 MDa molecular machine to facilitate translation and does so without the need of any additional factors, all while directing translation from two reading frames. The IGR IRES continues to prove invaluable for understanding mechanisms of viral IRES-dependent translation initiation as well as serving as a useful research tool. Finally, this IRES truly highlights the ingenuity and diverse means that viruses have evolved to exploit host processes for expression of their genomes.

1.7 Thesis investigation

Although dicistroviruses have served as models for studying cellular immune pathways and the deep understanding that we have of the IGR IRES mechanism, very little is known about the life cycle of these viruses. Much of our current understanding of various viral elements in dicistroviruses has been conducted *in vitro* with very few studies examining the contribution to *in vivo* pathogenesis. Thus, in my thesis I look to further elucidate how dicistroviruses usurp cellular pathways to facilitate the viral life cycle. In Chapter 2, I hypothesized that the 5'UTR of Cricket paralysis virus is indispensable for viral replication. To examine this, I utilized molecular biological methods to construct an infectious clone of Cricket paralysis virus, the first of the *dicistroviridae* family, and explored the role of a duplicated nucleotide sequence in the 5'UTR in viral infection. In Chapter 3, I hypothesized that functional and structural regions of the IGR IRES are critical for successful viral infection. By transferring established *in vitro* mutants of the

IGR IRES to the infectious clone developed in Chapter 2, I defined the importance of specific features in the IGR IRES for CrPV replication. Chapter 4 revolves around the hypothesis that the CrPV IGR IRES is capable of +1-frame translation and the encoded +1-frame protein is necessary for viral infection. To this end, I utilized a combination of biochemical techniques to explore the role of the CrPV protein, ORFx, in viral infection and the mechanism by which it is translated from the IGR IRES. Finally, in Chapter 5, I hypothesized that extracellular vesicles promote CrPV infection in *Drosophila*. Using quantitative proteomics and biochemical methods, this chapter investigates the release mechanism of CrPV from insect cells. In summary, my thesis strives to illuminate the different aspects of the CrPV life cycle and its host interactions and strengthens the understanding of the *Dicistroviridae* family of viruses.

Chapter 2: The 5'UTR of a *Cricket paralysis virus* infectious clone modulates viral infection

2.1 INTRODUCTION

The *Dicistroviridae* are a family of non-enveloped, single-stranded RNA viruses that infect arthropods (Bonning and Miller, 2010). Dicistrovirus genomes range from 8 to 10 kilobases (kb) in size, contain 5' a viral protein cap (VPg), a 3' poly(A) tail, and are characterized by a unique dicistronic genome arrangement. Distinct internal ribosome entry sites (IRES) drive translation of each open reading frame (ORF). The 5'UTR IRES directs translation of ORF1, which encodes viral non-structural proteins such as the suppressor of RNA interference (RNAi), RNA helicase, 3C protease, and the RNA-dependent RNA polymerase (RdRp). The intergenic region IRES (IGR IRES) facilitates expression of the viral structural proteins (ORF2) (Bonning and Miller, 2010).

These viruses are of economic and medical importance. Studies have linked a subset of bee dicistroviruses to the decline in North American honeybee populations and panaeid shrimp worldwide (Bonami et al., 1997; Cox-Foster et al., 2007b; de Miranda et al., 2010; Lightner and Redman, 1998). The type species, Cricket paralysis virus (CrPV), was first isolated in 1970 from Australian field crickets and has a wide host range, including *Drosophila melanogaster* (Bonning and Miller, 2010; Reinganum, 1970). Consequently, many studies have used CrPV and other dicistroviruses as models to delineate translational control mechanisms and antiviral immune mechanisms in insects, such as the RNAi and Imd (Immunodeficiency) pathways (Sabin et al., 2010). Furthermore, the unique dicistronic organization of its genome has led to insights into an unprecedented mechanism of translation initiation. The IGR IRES contains domains that functionally mimic a tRNA to recruit ribosomes and initiate translation in a factor-independent

manner (Costantino et al., 2007; Jan et al., 2003; Pestova and Hellen, 2003; Pfingsten et al., 2006; Wilson et al., 2000a). In general, the 5'UTR IRES of dicistroviruses is not well understood, however recently the mechanism of the CrPV 5'UTR IRES was shown to be similar to IRESs in the Type III family, such as the HCV IRES, and that translation is RACK1-dependent (Gross et al., 2017; Majzoub et al., 2014).

Despite the importance of dicistroviruses, the molecular mechanism and replication life cycle during infection are not completely understood. This is primarily due to a lack of an established infectious clone for any member of the *Dicistroviridae* family. *In vitro* transcripts derived from PCR amplicons of the Black queen cell virus (BQCV) RNA were reported to be infectious, however cloned infectious cDNA was not obtained (Benjeddou et al., 2002). The genomic sequence of RhPV was cloned, but it harbored a mutation in its first open reading frame that prohibited infection (Boyapalle et al., 2008; Pal et al., 2014). Here, we report the generation of an infectious full-length clone of the CrPV genome, which compared to the published sequence (Wilson et al., 2000b), contains unique nucleotide changes including a 196-nt duplication in the 5'UTR. *In vitro* transcribed RNA of the CrPV clone transfected into *Drosophila* S2 cells recapitulates CrPV infection and produces virions that infect and replicate in *Drosophila* S2 cells. Furthermore, we demonstrate that the duplication in the 5'UTR is not an artifact of cloning, but occurs within viral stocks and can stimulate translational activity of the 5'UTR IRES. Using reverse genetics, an infectious clone containing the duplicated element in the 5'UTR is less fit than one containing a single copy of the element in S2 cells. Finally we show evidence that infection of virus generated by the CrPV infectious clone causes mortality in adult flies. To our knowledge, this is the first established infectious clone of the *Dicistroviridae* family.

2.2 MATERIALS & METHODS

2.2.1 Cell culture and virus

Drosophila Schneider line 2 (S2) cells were maintained and passaged in Shield's and Sang medium (Sigma) supplemented with 10% fetal bovine serum. Kc167 Cells were maintained in passaged in Sf-900 III medium (Invitrogen) and supplemented with 10% fetal bovine serum.

Propagation of CrPV in *Drosophila* S2 cells has been previous described (Garrey et al., 2010). CrPV was isolated from *Drosophila* S2 cells using an adapted protocol (Krishna et al., 2003) with cesium chloride (CsCl) gradients. Briefly, S2 cells were infected with CrPV and incubated for 24 h. Cells were dislodged into the media, treated with 0.5% Igepal CA-630 (Nonidet P-40) and 0.1% 2-mercaptoethanol, and incubated on ice for 10 min. Cell debris was cleared by centrifugation at 13,800 RCF for 15 min at 4°C. The supernatant was then treated with RNase A (100 μg/mL) for 30 min at 27°C. Viral particles were then pelleted through a 30% (w/w) sucrose cushion in VP buffer (50 mM HEPES, 0.1% 2-mercaptoethanol, and 5 mM $CaCl_2$) by ultracentrifugation at 141,000 RCF for 2.5 h at 11°C. The pellet was resuspended in VP buffer and added to a 50% CsCl gradient in VP buffer. The gradient was ultracentrifuged at 274,000 RCF for 18 h at 4°C. The viral band was extracted and dialyzed with phosphate-buffered saline (PBS) overnight. Viral titres and yield were performed as previously described (Garrey et al., 2010).

Virus was inactivated by UV light for 30 minutes on ice with a Stratagene 1800 UV crosslinker. Lack of cytopathic effects upon infection verified inactivation of the virus.

2.2.2 CrPV cDNA construction

RNA was extracted from CsCl purified CrPV using TRIzol (Invitrogen). The construction of the full-length infectious clones was generated as follows. First strand CrPV

cDNA was synthesized through reverse transcription (RT)-PCR using Superscript III Reverse Transcriptase (Invitrogen) and an Oligo dT primer at 50°C for 2 h (Figure 3A). The CrPV cDNA was amplified using primers P1 (5' – GAGGGTACCATGTCTTTTCAACAAAC AAACAAC – 3' [KpnI site is underlined]) and P3 (5' – GGAGCTCAGA AAACAACTATTAATCAAAAACCAAATTG – 3' [SacI site is underlined]), which introduces KpnI and SacI sites. The resulting PCR product, which contains the CrPV genome from the beginning of ORF1 to the end of the 3'UTR, was cloned into pAC5.1/V5 His B (Invitrogen) using KpnI and SacI sites to create pCrPV-Δ5'UTR (Figure 3A). Amplification of the CrPV genomic cDNA in its entirety was performed using primers P3 and P2 (5' - GAGGGTACCTTTAATAAGT GTTGTGCAGATTAATCTG – 3' [KpnI site is underlined]). To create the full-length clone of CrPV, a KpnI/HindIII fragment from the P2/P3 PCR product was cloned into the pCrPV-Δ5'UTR using the same restriction sites. The resulting plasmid was denoted as pCrPV-1 (Figure 3A). A T7 RNA polymerase promoter flanking the 5' end of the CrPV genome was incorporated into pCrPV-1 using primer incomplete polymerase extension (PIPEs; (Klock and Lesley, 2009)) with primers P4 (5' - *GAAGATTAATACGACTCACTATAGGGAGA*TTTAATAAGT GTTGTGCAGATTAATCTG – 3' [T7 promoter is underlined]) and P5 (5' – *TCTCCCTATAGTGAGTCGTATTAATCTTC*TCTCCCTATAGTGAGTCGTATTAATCTTC – 3' [T7 promoter is underlined]) to create pCrPV-2 (Figure 3A). Primers P2 and P6 (5 – TGTAAGCTTTGTTCTTAAGTTCTCTG – '3 [HindIII site is underlined]) were used to amplify the 5'UTR(ΔDup) and part of ORF1 from the 2006 CrPV stocks (Figure 1). This was inserted in between KpnI and HindII sites of pCrPV-2 to generate pCrPV-3 (Figure 3A). Finally, a T7 promoter was added to pCrPV-3 using primers P4 and P5.

Stop codon mutations in either ORF1 or ORF2 of pCrPV-2 (pCrPV-2-ORF1-STOP and pCrPV-2-ORF2-STOP, respectively) were introduced via PCR-based site-directed mutagenesis. UAA stop codons were introduced by mutating nucleotides C775T(C971T) and G778T(G974T) of ORF1 and nucleotide A6232T(A6428T) of ORF2. The numbering of the nucleotides is based on CrPV-3, with CrPV-2 in parentheses.

2.2.3 Bicistronic and minigenome constructs

For the bicistronic reporter construct, a PCR-derived fragment containing the CrPV or CrPV-2 5'UTR plus 54 nt of ORF1 was ligated into the EcoRI and NcoI sites of pMC53 upstream of the firefly luciferase open reading frame. Subsequently, the 5'UTR-firefly luciferase fragment was inserted into EcoRI and XbaI sites of pEJ551 (Wang et al., 2013). For the minigenome construct, a PCR-derived fragment containing the IGR IRES and a short region of VP2 (nucleotides 6413-6571) was inserted between EcoRI and NcoI restriction sites in pEJ551. Subsequently, overlapping PCR was used to insert the CrPV 5'UTR (without the duplication) in front of Renilla luciferase open reading frame. Lastly, a synthesized CrPV 3'UTR (Integrated DNA Technologies) was inserted between XbaI and BamHI sites downstream of the firefly luciferase open reading frame. The resulting plasmid is called pMG-ΔDup. The minigenome containing the 5'UTR duplicated element was constructed using the Gibson Assembly approach (New Enlgand Biolabs). The CrPV 5'UTR containing the duplication was amplified using pCrPV-2 as a template with primers P10 (5' – ACGACTCACTATAGGGTTTAATAAGTGTTGTGCAGATTAATC – 3') and P11 (5' – TAAACTTTCGAAGTGATCACATTGTAAGAATCGG – 3'). The minigenome containing the Renilla Luc-IGR IRES-Firefly Luc-3'UTR was amplified with primers P12 (5' – CATGACTTCGAAAGTTTATGATC – 3') and P13 (5' –

CCCTATAGTGAGTCGTATTAAAG – 3'). The CrPV 5'UTR amplicon was mixed with the minigenome amplicon and incubated for 15 min at 50°C in the Gibson Assembly mix (NEB). The reaction was transformed into competent *E. coli* DH5α cells and sequenced to confirm that the 5'UTR duplication was present.

2.2.4 *In vitro* transcription and RNA transfection

Purified plasmid was linearized with either Ecl136II (pCrPV-2 and derivatives) or BamHI (minigenome constructs). RNA was transcribed in a T7 RNA polymerase reaction and subsequently purified with a RNeasy kit (Qiagen). The integrity and purity of the RNA was confirmed on a 1.2% denaturing formaldehyde agarose gel.

Transfection of *in vitro* synthesized RNA into S2 cells was performed using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions. 3 µg of RNA derived from either the full-length CrPV clone or its cognate mutants and 1 µg of reporter RNA were used for transfection using 2.5 x 10^6 cells.

2.2.5 RT-PCR and Northern blots

Total RNA was isolated from cells using TRIzol reagent. RT-PCR was performed on 5 µg of RNA using Superscript Reverse Transcriptase III (Invitrogen) at 55°C followed by DNase I treatment for 30 min at 37°C. For reverse transcription of the negative sense CrPV viral RNA, tagged primer P7 (5' – CTATGGATCCATGGGAGAAGATCAGCAAAT – 3' [tag is underlined]) was used. Primers P8 (5' – CTATGGATCCATGGGAGAAG – 3') and P9 (5' – GTGGCTGAAATACTATCTCTGG – 3'), were used for PCR amplification of the negative sense strand of the CrPV genome. The 5'UTR of CrPV was RT-PCR amplified using primers P2 and P14 (5' – GAGCCATGGCTCAAGGGAGTTGATGTTGTT – 3' [NcoI site is underlined]). Diagnostic RT-PCRs of Drosophila C virus (DCV), Drosophila X virus (DXV) or Flock house

virus (FHV) cDNA was detected using the following primers: DCV genome, P15 (5' – TCTGCAGGAGTTCCCGATG – 3') and P16 (5' – CAATGCGCTTCCGGAGAC -3'); Segment B of DXV, P17 (5' – TGGACATCGAAACAGGGTACAC – 3') and P18 (5' – CCGCGGATAGAGTTTGGTAACC – 3'); FHV RNA1, P19 (5' – CGCCAATGATAAGAAACCAAGCG – 3') and P20 (5' – GGTCCACAAATACACGAGACAGG – 3'). Actin was amplified using primers P21 (5' – CGACAACGGCTCTGGCATGTGCA – 3') and P22 (5' – ACGCAGCTCATTGTAGAAGGTG – 3').

Northern blots were performed by loading 5 µg of RNA on a denaturing agarose gel and subsequently transferred to Zeta-probe blotting membrane (Bio-Rad). DNA probes were radiolabelled with a DecaLabel DNA labeling kit (Fermentas) and hybridized overnight. Radioactive bands were detected via phosphoimager analysis (Typhoon; GE Healthcare).

2.2.6 Western blots

Equal amounts of S2 protein lysates (20 µg) were resolved on a 12% SDS-PAGE gel and then transferred to a polyvinylidene difluoride Immobilon-FL membrane (Millipore). Membranes were blocked for 30 min at room temperature with 5% skim milk in TBST. Blots were incubated for 1 h at room temperature with the following antibodies: CrPV ORF1 (raised against CrPV RdRp) rabbit polyclonal (1:10,000) or CrPV ORF2 (raised against CrPV VP2) rabbit polyclonal (1:10,000)(Garrey et al., 2010). Membranes were washed 3 times with TBST and incubated with either IRDye 800CW goat anti-rabbit IgG (1:20,000; LI-COR Biosciences) or donkey anti-rabbit IgG-horseradish peroxidase (1:20,000; GE Healthcare) for 1 h at room temperature. An Odyssey imager (LI-COR Biosciences) and enhanced chemiluminescence (Thermo Scientific) was used for detection.

2.2.7 Electron microscopy

Negative stained specimens were prepared by adsorbing samples to glow discharged carbon-coated copper grids and subsequently staining with uranyl formate. A Tecnai Spirit transmission electron microscope (FEI) operated at an accelerating voltage of 120 kV was used to examine these specimens. Images were acquired using a 4K x 4K Eagle charge-coupled device (CCD) camera (FEI) at a nominal magnification of 49,000x.

2.2.8 Transwell assay

S2 cells were transfected with *in vitro* transcribed viral genomic RNA for 24 h at 25°C. Untreated cells were seeded onto a coverslip pre-treated with concanavalin A (0.5 mg/mL). Transfected cells were washed 3 times with PBS and then seeded onto a transwell insert within the 6-well plate. The coverslips were washed with PBS and treated with 3% paraformaldehyde for 15 min followed by methanol treatment for 10 min. The cells were washed with PBS and incubated with anti-ORF2 antibody (1:500 in 5% bovine serum albumin [BSA] in PBS) for 1 h at room temperature. Cells were washed 3 times with PBS and incubated with goat anti-rabbit Texas Red IgG (1:200 in 5% BSA in PBS; Invitrogen) for 1 h at room temperature, followed by 0.5 µg/mL Hoechst to stain the nuclei. Slides were imaged using a confocal microscope (Olympus FV1000 using Olympus Fluoview, version 2.0a) with a 60X oil immersion lens. Images shown represent a single Z-section and were processed in Photoshop CS6.

2.2.9 *In vitro* and *in vivo* translation assays

S2 cell translation extracts were obtained using an adapted protocol (Brasey et al., 2003; Roy et al., 2004). Briefly, 2.0 x 10^9 cells were resuspended in PBS and either mock infected or infected with CrPV (MOI = 10). Cells were then diluted to 1.0 x 10^7 cells/mL and incubated at 25°C for 6 h. Cells were pelleted (1000 *g*, 8 min), washed once with PBS, and resuspended in

hypotonic buffer (10 mM HEPES-KOH [pH 7.4], 10 mM KOAc, 0.5 mM MgOAc, 1 mM DTT) and subsequently incubated on ice for 5 min. Cells were lysed by extrusion 25 times through a 23-gauge needle. Resulting extracts were adjusted with 50 mM KOAc. Cell debris was cleared by centrifugation (16,000 RCF 5 min at 4°C) and the supernatant was stored at -80°C.

Bicistronic or minigenome RNA (90 nM) were incubated in either mock-infected or CrPV-infected S2 cell extracts for 30 min at 30°C. Luciferase activity was measured using a dual Luciferase assay (Promega). *In vitro* translation of the full-length viral RNA genome (uncapped) was performed in *Spodoptera frugiperda* 21 (Sf-21) cell extract (Promega) in the presence of [^{35}S]-methionine/cysteine. Reactions were loaded on a SDS-PAGE. Gels were dried and radioactive bands were monitored by phosphoimager analysis.

For *in vivo* translation assays, mock- or CrPV-infected S2 cells were incubated with [^{35}S]-methionine/cysteine for the last 30 minutes of the infection. Equal amounts of lysate (10 µg) were then loaded on a SDS-PAGE. Gels were dried and radioactive bands were monitored by phosphoimager analysis. For luciferase assays, S2 cells (1.5 x 10^6 cells) were transfected with *vitro* transcribed RNA for 2 h, followed by mock- or CrPV infection (MOI 10) for 6 h. Cells were harvested and lysed and luciferase activity was measured (Promega) using a microplate luminometer (Berthold Technologies, Centro LB 960).

2.2.10 Fly stocks and viral injections

Isogenic w^{1118} flies were maintained on standard cornmeal food at 25°C and 70% humidity with a 12 h light-dark cycle. Freshly eclosed virgin males and females were separated and collected in five groups of 10 each. Flies (10 males and 10 females) were injected with 200 nL of PBS, CrPV-2, or CrPV-3 (500 or 5000 FFU) using a Neurophore BH-2 Injection System (Harvard Apparatus) and transferred to standard food and flipped to fresh food every 3 days.

2.2.11 Nucleotide sequence accession numbers

The sequences of pCrPV-2 and pCrPV-3 have been deposited in GenBank under accession numbers KP974706 and KP974707, respectively.

2.3 RESULTS

2.3.1 Selection of a 196-nucleotide element within the CrPV 5'UTR that augments translation

In general, the mechanism underlying CrPV 5'UTR IRES translation is poorly understood. To address this, we initially RT-PCR cloned the 5'UTR from CrPV stocks that have been propagated in our S2 tissue culture cells. Using a CrPV stock from 2006 (CrPV-2006), an RT-PCR product corresponding to the published predicted size of the CrPV 5'UTR plus 54-nt of ORF1 (763 nucleotides) was detected (Figure 1A). Interestingly, using a CrPV stock from 2013 (CrPV-2013), RT-PCR analysis detected a band corresponding to the 5'UTR that was larger in size by roughly 200 bp (Figure 1A). Moreover, a CrPV stock from 2009 (CrPV-2009) produced a product that is consistent with the size of the 5'UTR of CrPV-2006 as well as a very weak band consistent with that of the 5'UTR of CrPV-2013 (Figure 1A). Sequencing analysis revealed that CrPV-2013 acquired a tandem duplication of nucleotides 76-271 (196 nt) in the 5'UTR that was not previously described in the original CrPV sequence (Figure 1B; NC_003924). As RNA viruses largely exist as quasispecies (Lauring and Andino, 2010), we surmised that the duplication in the 5'UTR likely arose due to selection during infection which presumably increased viral fitness.

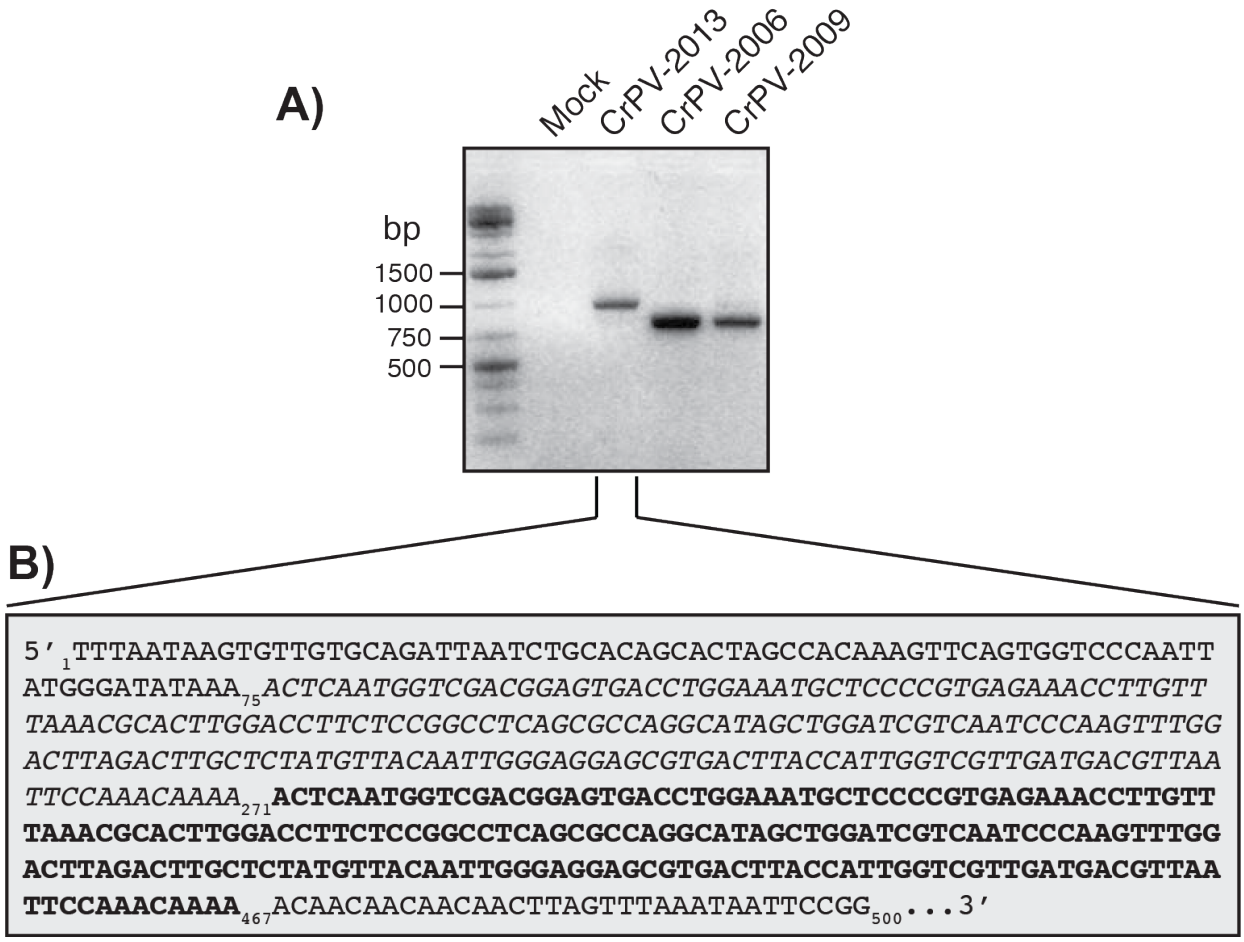


Figure 2.1 RT-PCR and sequence of the CrPV 5'UTR from different viral stocks.

(A) RT-PCR analysis of the CrPV 5'UTRs from cells that were infected with different CrPV stocks from circa (ca.) 2006, 2009, or 2013 for 24 hours. (B) Sequence of the first 500 nucleotides of the 5'UTR from CrPV-2013, which contains the tandem 196-nt duplication (italics and bolded).

We next investigated the mechanism by which the 196-nucleotide duplication affects CrPV infection. As the duplication occurred within a known IRES sequence, we hypothesized that the duplication may influence the translational activity of the 5'UTR IRES. To address this, we generated a minigenome reporter construct that encodes Renilla and firefly luciferase genes in lieu of CrPV ORF1 and ORF2, respectively (Figure 2.2A). Consequently, expression of Renilla luciferase (RLuc) is driven by the 5'UTR IRES while firefly luciferase (FLuc) expression is directed by the IGR IRES (Figure 2.2A). *In vitro* transcribed minigenome RNA with (5'UTR^{+Dup}) or without (5'UTR$^{\Delta Dup}$) the 5'UTR duplication was incubated in translation extracts derived from mock- or CrPV-infected S2 cells and luciferase activities were measured. In both mock and CrPV-infected cell extracts, the 5'UTR^{+Dup} containing reporter RNA resulted in greater than 4-fold more Renilla luciferase activity than the 5'UTR$^{\Delta Dup}$ RNA (Figure 2.2B). In all cases, IGR IRES-mediated translation of the firefly luciferase ORF was similar in mock and CrPV-infected lysates. To determine whether this effect is similar *in vivo*, the same minigenome RNAs were transfected into S2 cells followed by mock or CrPV infection. Cells were harvested at 6 h post infection and luciferase activities were measured. We previously determined that the luciferase activities increased linearly from 1-6 h after transfection of the reporter RNAs, suggesting that the RNA is intact and engaged in translation during this time (Wang et al., 2013). Similar to the *in vitro* results, the 5'UTR containing the duplication resulted in higher Renilla luciferase activity as compared to that without the duplication (Figure 2.2C), indicating that the duplication within the 5'UTR stimulates translation in the context of the minigenome reporter RNA.

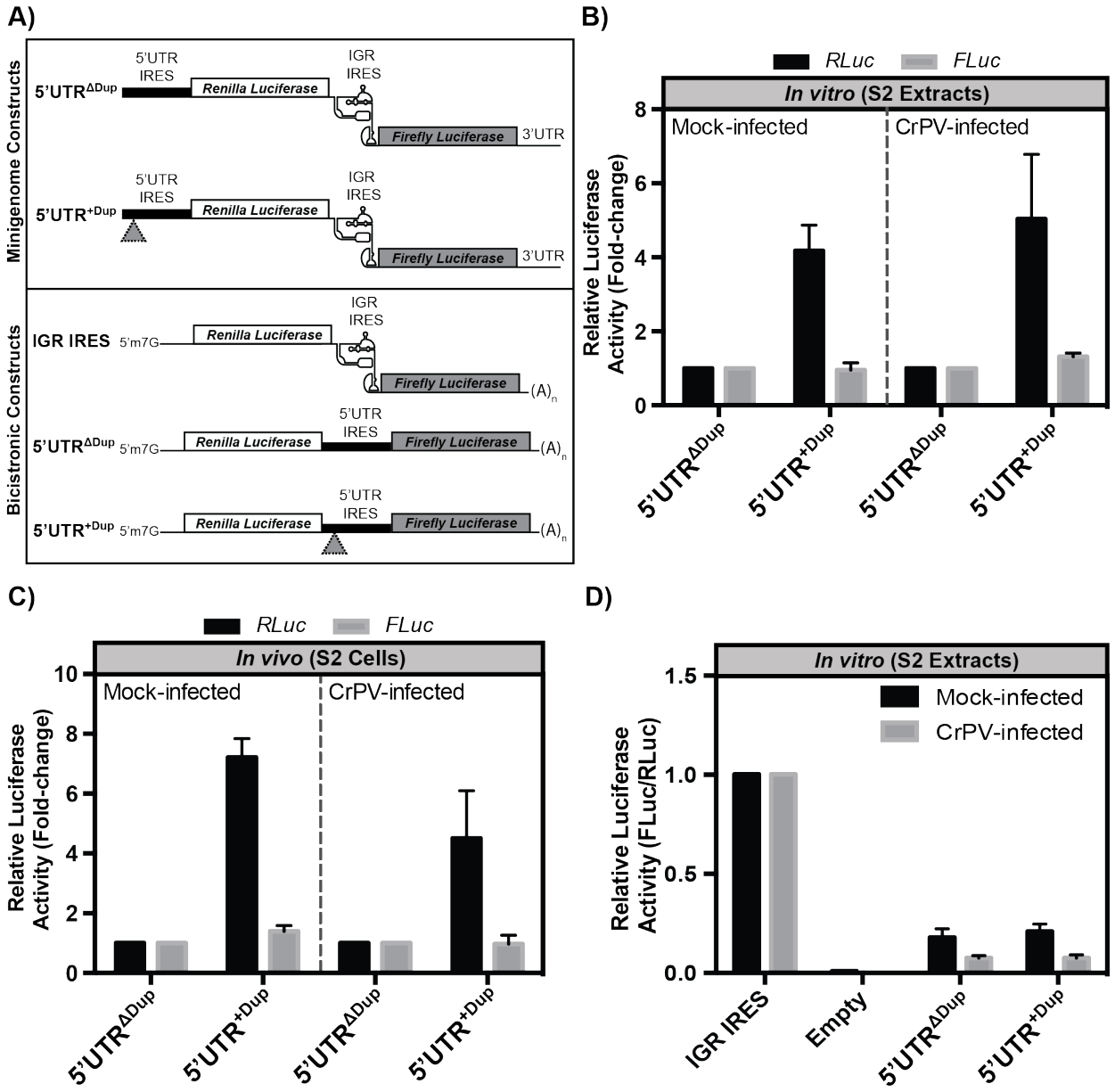

A)
Minigenome Constructs
5'UTR IRES
IGR IRES
5'UTR^ΔDup
Renilla Luciferase
Firefly Luciferase
3'UTR
5'UTR^+Dup
Bicistronic Constructs
IGR IRES 5'm7G
(A)n
B)
RLuc
FLuc
In vitro (S2 Extracts)
Mock-infected
CrPV-infected
Relative Luciferase Activity (Fold-change)
C)
In vivo (S2 Cells)
D)
In vitro (S2 Extracts)
Mock-infected
CrPV-infected
Relative Luciferase Activity (FLuc/RLuc)
IGR IRES
Empty

Figure 2.2 Translational control mediated by the duplication in the CrPV-2 5'UTR.

(A) Schematics of reporter constructs. Minigenome reporter RNAs (top) contain Renilla and firefly luciferase genes in lieu of the CrPV ORF1 and ORF2, respectively, but maintain the untranslated regions of the CrPV genome. The bicistronic reporter RNA (bottom) contain a 5'-m7G cap structure and a 3'-poly(A) tail. Renilla luciferase is translated via 5'end scanning-dependent translation while firefly luciferase is IRES-dependent. Gray triangles represent the duplicated sequence element. (B) In vitro translation. in vitro transcribed minigenome RNA containing the 5'UTR with (5'UTR+Dup) or without (5'UTR-Dup) the 196 bp duplication were incubated in mock- or CrPV-infected S2 cell extracts for 30 min at 30°C. The Renilla and firefly luciferase activities monitor 5'UTR- and IGR IRES-dependent translation, respectively. Luciferase activity is normalized to the minigenome reporter construct containing the 5'UTR-Dup. (C) Translational activity in S2 cells. S2 cells, transfected with the indicated minigenome RNAs for 2 h at 25°C, were either mock- or CrPV-infected (MOI 10) for 6 h. Luciferase activity is normalized to the minigenome reporter construct containing the 5'UTR-Dup. (D) IRES activity. Bicistronic RNAs containing either the CrPV 5'UTR+Dup or 5'UTR-Dup 5'UTR or an empty intergenic region were incubated in mock- or CrPV-infected S2 cell extracts for 30 min at 30°C. Renilla luciferase measures scanning-dependent translation while firefly luciferase measures IRES-mediated translation. The ratios of firefly to Renilla luciferase activities were normalized to the IGR IRES activity. Shown are averages from at least three independent experiments (±S.D.)

The 5'UTR of dicistroviruses has been reported to contain IRES activity (Masoumi et al., 2003; Shibuya and Nakashima, 2006; Woolaway et al., 2001). To determine whether the duplication has an effect on IRES translation, we inserted the 5'UTR with (5'UTR^{+Dup}) or without (5'UTR$^{\Delta Dup}$) the duplication into the intergenic region of a bicistronic reporter construct (Figure 2.2A). Here, the 5'UTR directs translation of firefly luciferase while Renilla luciferase translation occurs by a 5'end-dependent scanning mechanism (Wang et al., 2013). *In vitro* transcribed bicistronic RNAs were incubated in S2 cell extracts and the ratio of FLuc/RLuc activities were measured. As shown previously, the IGR IRES directed translation of the

luciferase ORF as compared to a no IRES control (Figure 2.2D). Translation driven by both the 5'UTR^{+Dup} and 5'UTR$^{\Delta Dup}$ was approximately 20% and 10% of IGR IRES translation in mock- and CrPV-infected lysates, respectively (Figure 2.2D), indicating that the duplication within the 5'UTR does not affect IRES activity. This result is in contrast to that observed in the minigenome reporter RNA where the 5'UTR^{+Dup} stimulated translation compared to the 5'UTR without the duplication (Figure 2.2B, 2.2C). These results show an important distinction of the location of the 5'UTR within the reporter RNA and the mechanism of translation via the 5'UTR.

2.3.2 Construction and characterization of a CrPV infectious clone

To determine the role of the 5'UTR duplication during CrPV infection, we constructed an infectious clone of CrPV with (pCrPV-2) and without (pCrPV-3) the 5'UTR duplication (Figure 2.3A). Viral RNA isolated from pure CrPV virions was used to RT-PCR amplify a partial CrPV cDNA lacking the 5'UTR using primers P1 and P2. The resulting PCR product was cloned into the pAC5.1/V5 His B plasmid creating pCrPV-Δ5'UTR with KpnI and SacI restriction sites flanking the genomic sequence (Figure 2.3A). Attempts to clone the full-length CrPV cDNA using a primer that hybridizes to the 5' end of the CrPV genome were unsuccessful. Therefore, we took advantage of a unique endogenous HindIII site in the CrPV genome in order to construct a full-length clone. An RT-PCR product of CrPV cDNA that contains the 5'UTR was ligated into the KpnI-HindIII sites of pCrPV-Δ5'UTR, thus generating pCrPV-1 (Figure 2.3A). A T7 RNA polymerase promoter was inserted between the KpnI site and the 5' end of the CrPV genome to create the construct pCrPV-2 (Figure 2.3A). To construct pCrPV-3, a PCR fragment from CrPV-2006 was inserted into the HindIII and KpnI sites of pCrPV-2 (Figure 2.3A).

Propagation of pCrPV-2 and pCrPV-3 plasmids was performed in *E. coli* DH5α cells at 30°C. The entire sequence of each CrPV genomic clone was confirmed by sequencing. When grown at

37°C, all plasmids containing the CrPV genome contained deletions or insertions within the viral cDNA. Previous reports have shown that clones of viral genomic sequences can be highly unstable (Khromykh and Westaway, 1994; Lai et al., 1991; Ryan et al., 1989; Sumiyoshi et al., 1992). This is attributed to the fact that propagation of viral cDNA clones in *E. coli* can result in toxic products that prevent cloning efforts or accumulation of mutations in the genomic sequence causing clones to become defective (Boyer and Haenni, 1994; Li et al., 2011). Accordingly, the CrPV genome lacking the 5'UTR was stable (pCrPV-Δ5'UTR; Figure 2.3A), however during propagation of the full-length clone (pCrPV-2 and pCrPV-3; Figure 2.3A) in *E. coli* DH5α at 37°C it acquired large nucleotide insertions (up to ~1kb) within the genome (data not shown). To overcome this issue, we found that growth of *E. coli* cells harboring pCrPV-2, pCrPV-3 or other clones at 30°C permitted propagation with no apparent aberrations.

Sequence analysis of CrPV-2 and CrPV-3 revealed 34 nucleotide substitutions that differ from the previously published CrPV sequence (NC_003924; (Wilson et al., 2000b))(Table 2.1). Of these, 2 nucleotide changes are located in the 5' and 3' UTRs and 32 are within ORF1 and ORF2. No changes were found within the IGR IRES, reflecting the importance of this element for viral IRES activity. Inspection of ORF1 and ORF2 nucleotide changes revealed that 14 are synonymous while 19 are non-synonymous (Table 2.1; Figure 2.3B). In addition to the nucleotide substitutions, insertion and deletion events occurred at 4 positions: nucleotides 3671 (insertion of a G), 3678 (deletion of an A), 5245 (insertion of an A), and 5270 (deletion of a C) which resulted in an alternative amino acid sequence, but maintained the reading frame (Table 2.1; Figure 2.3B). The numbering refers to the nucleotides within the published CrPV sequence (NC_003924; (Wilson et al., 2000b)).

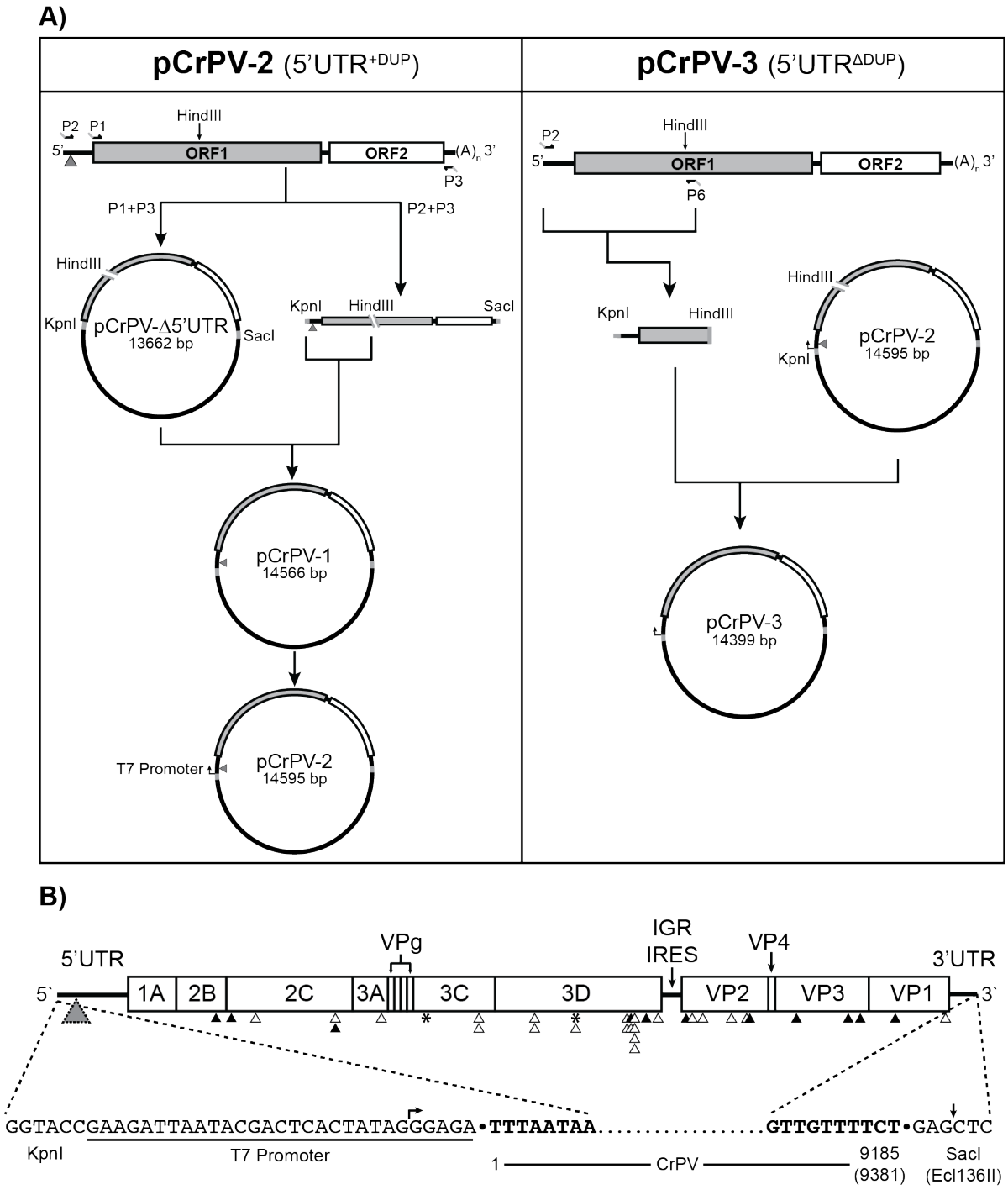

A)
pCrPV-2 (5'UTR+DUP)
pCrPV-3 (5'UTRΔDUP)
HindIII
P2
P1
5'
ORF1
ORF2
(A)n 3'
P3
P1+P3
P2+P3
KpnI
SacI
pCrPV-Δ5'UTR
13662 bp
pCrPV-1
14566 bp
T7 Promoter
pCrPV-2
14595 bp
P6
pCrPV-3
14399 bp
B)
5'UTR
VPg
IGR
IRES
VP4
3'UTR
1A
2B
2C
3A
3C
3D
VP2
VP3
VP1
GGTACCGAAGATTAATACGACTCACTATAGGGAGA•TTTAATAA.....GTTGTTTTCT•GAGCTC
KpnI
T7 Promoter
1
CrPV
9185
(9381)
SacI
(Ecl136II)

Figure 2.3 Construction and map of the full-length cDNA clones CrPV-2 and CrPV-3.

(A) RNA from CsCl purified CrPV was reverse-transcribed using an oligo-dT primer. Primers P1 and P3 were used to clone the CrPV genome without the 5'UTR into pAC5.1/V5 His B while primers P2 and P3 containing KpnI and SacI restriction sites, respectively, were used to amplify the CrPV genome. CrPV contains an endogenous HindIII site in ORF1. Therefore, to clone the 5'UTR, both the P1+P2 amplicon and pCrPV-Δ5'UTR were digested with HindII and KpnI. The resulting fragments were ligated together to generate pCrPV-1. A T7 polymerase promoter was added to pCrPV-1 via primer incomplete polymerase extension (PIPEs) to create pCrPV-2. To create pCrPV-3, a PCR fragment derived from the 2006 stock of CrPV was inserted into the KpnI and HindIII sites of pCrPV-2 followed by a T7 promoter. (B) Top: Organization of the CrPV-2 and CrPV-3 cDNA genomic sequences showing all annotated proteins. Arrowheads denote either synonymous (black) or non-synonymous (white) changes that differ from the published CrPV sequence (NC_003924). Insertion-deletion events are indicated by an asterisk. A gray triangle depicts the location of the 196-nt duplication in CrPV-2. See Table 1 for specific nucleotide changes. Bottom: Sequence of the 5' and 3' ends of the cloned viral genomes. Nucleotides corresponding to the viral genome are bolded. Arrows indicate the beginning and end of the *in vitro* transcript obtained from using T7 polymerase with Ecl136II-linearized plasmid.

Altogether, the changes within CrPV-2 and CrPV-3 resulted in 31 alterations of the amino acid sequence, of which 26 occurred in ORF1 and 5 occurred in ORF2 with no differences occurring in the predicted conserved motifs of the non-structural proteins (Table 2.1; Figure 2.3B). As mentioned previously, a tandem duplication of nucleotides A76-A271 (196 nt) is present in the 5'UTR of CrPV-2 that is not in CrPV-3 or the previously published CrPV genome (Table 2.1; Figure 2.3B).

Table 2.1 Nucleotide and amino acid sequence differences between CrPV-2 and CrPV Sequences[a]

	ORF1			ORF2			3'UTR
5'UTR nucleotide change	Nucleotide change	Amino acid change	VP[b]	Nucleotide change	Amino acid change	VP[b]	nucleotide change
A76-A271, A272-A467[c]	C1545T		2B	C6279A		VP2	G9005A
G312A	A1719G		2C	A6316G	I34V	VP2	
	T1979C	V424A	2C	G6426A	G70R	VP2	
	C2127T		2C	C6707G	T164R	VP2	
	T2128C	F474L	2C	C6860T	T215I	VP2	
	T3190A	F828I	3A	C6882T		VP2	
	G3671[d]	A988G	3C				
	A3678[e]	T989N, A990S	3C	C7359A		VP3	
			3C	T7896C		VP3	
	A4189G	N1161D	3C	A8019T		VP3	
	A4255G	P1183A	3C	C8355T		VP1	
	C4745T	T1346I	3D	A8870G	E885G	VP1	
	T4814C	F1369S	3D				
	A5245[d]	F1513I;	3D				
	G5262T	F1514L;	3D				
	C5270[e]	G1515W, Q1516A, S1617I, C1518F, G1519W, K1520E, S1521V	3D				
	T5693A	F1662Y	3D				
	C5696G	P1663R	3D				
	C5706A		3D				
	C5715G	H1669Q	3D				
	A5724G		3D				
	A5726G	K1673R	3D				
	A5728G	R1674G	3D				
	C5745G	H1679Q	3D				
	T5784A		3D				
	A5971G	K1755E	3D				

[a]Nucleotide numbers correspond to those of the published CrPV Sequence. The first amino acid letter corresponds to CrPV (GenBank accession number NC_003924) while the second is the amino acid in the same position in CrPV-2.
[b]VP, Viral Protein.
[c]Duplicated sequence.
[d]Insertion;
[e]Deletion

2.3.3 Transfection of *in vitro* transcribed CrPV-2 and CrPV-3 RNA resembles CrPV infection

To validate the infectious clones of CrPV, we first determined whether viral proteins are expressed using *in vitro* translation in Sf-21 extracts. Incubation of *in vitro* transcribed RNA from pCrPV-2 or pCrPV-3 in the presence of [^{35}S]-methionine/cysteine resulted in the expression of proteins ranging from <25 kDa to >170 kDa (Figure 2.4, lane 4). If unprocessed, ORF1 and ORF2 polyproteins are 1771 and 895 amino acids in length with predicted molecular weights of 203.8 and 100.4 kDa, respectively. To determine whether the proteins are translated from ORF1 or ORF2, we introduced stop codons into the CrPV-2 cDNA. Stop codons at nucleotides C971T and G974T within ORF1 of CrPV-2 (CrPV-2-ORF1-STOP) should prevent the synthesis of viral non-structural proteins, including the viral protease, the RNA helicase, and RNA-dependent RNA polymerase (RdRp). We inserted two stop codons to prevent translational read-through. Conversely, a stop codon at nucleotide C6428T within ORF2 (CrPV-2-ORF2-STOP) should halt synthesis of viral structural proteins. *In vitro* translation using Sf-21 extracts demonstrated that the stop codons introduced into ORF1 or ORF2 of CrPV-2 prevented synthesis of the viral non-structural or structural proteins (Figure 2.4, lanes 5 and 6). Specifically, compared to reactions containing CrPV-2 or CrPV-3, the CrPV-2-ORF1-STOP resulted in loss of expression of all proteins except for a protein that migrated at ~110 kDa (Figure 2.4, lanes 4 and 5). Because CrPV-2-ORF1-STOP prevents expression of all ORF1 proteins and in particular the 3C-like protease, the ~110 kDa protein is presumably the unprocessed ORF2 polyprotein.

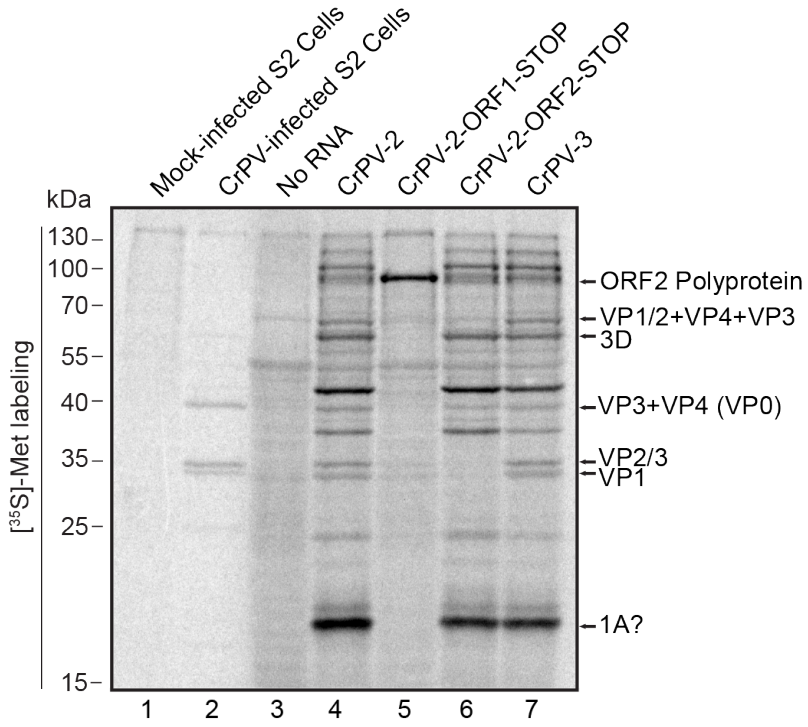


Figure 2.4 CrPV-2 and CrPV-3 RNA produces viral proteins *in vitro*.

In vitro synthesized RNA from CrPV-2, CrPV-3, CrPV-2-ORF1-STOP, or CrPV-2-ORF2-STOP were incubated in Sf-21 translation extracts for 2 h at 30°C in the presence of [^{35}S]-methionine/cysteine. Reactions were resolved by SDS-PAGE and radioactive proteins were visualized by phosphoimager analysis. As a control, lysates from [^{35}S]-methionine/cysteine-labeled mock- and CrPV-infected S2 cells were resolved by SDS-PAGE in parallel (lanes 1 and 2). Shown is a representative gel from at least three independent experiments.

Furthermore, the proteins expressed in reactions containing CrPV-2-ORF2-STOP likely represents a combination of processed and unprocessed ORF1 proteins (Figure 2.4, lane 6). Moreover, the lack of expression of proteins at 70 and 35 kDa in the CrPV-2-ORF2-STOP reactions confirms that these proteins represent ORF2 structural proteins. Importantly, we annotated the viral proteins expressed *in vitro* based on [^{35}S]-labeled lysates from CrPV-infected cells and previous reports showing the migration of CrPV structural proteins (Figure 2.4, lane 2) (Garrey et al., 2010; Moore et al., 1980; Wilson et al., 2000b). Thus, we conclude that CrPV-2 and CrPV-3 can express all known CrPV proteins *in vitro*.

We next asked whether CrPV-2 and CrPV-3 are infectious. Transfection of S2 cells with *in vitro* transcribed CrPV-2 and CrPV-3 RNA resulted in cytopathic effects (CPE) that include membrane blebbing, detachment of cells from the substratum, cell clumping and cell lysis at 48 h post transfection (h.p.t), which are phenotypes also observed in cells infected with CrPV (Figure 2.5A). Northern blot analysis showed CrPV-2 and CrPV-3 viral RNA accumulated over time, strongly suggesting that the infectious clones can replicate in S2 cells. Interestingly, transfection of CrPV-3 RNA reproducibly accumulated at an earlier time point than that of CrPV-2 (Figure 2.5B). Furthermore, RT-PCR analysis using tagged primers specific to the negative strand followed by nested PCR using a primer specific to the primer tag demonstrated that cells transfected with CrPV-2 or CrPV-3 RNA or infected with CrPV result in synthesis of negative sense viral RNA at 24 h.p.t (Figure 2.5C). The use of the tagged primer RT-PCR approach was necessary as the infectious CrPV-2 and CrPV-3 RNA led to false-priming in the RT reaction in the absence of a primer, thus producing an RT-PCR product (data not shown), which is similar to that observed for dengue virus, hepatitis C virus, and alphavirus RNA (Gunji, 1993; Peyrefitte et al., 2003; Plaskon et al., 2009). Finally, expression of both nonstructural (3CD) and structural

(VP2) viral proteins was detected by immunoblot analysis at 48 h.p.t (Figure 2.5D). In summary, these results collectively demonstrate that CrPV-2 and CrPV-3 can replicate and produce viral proteins in S2 cells.

To determine whether replication by the infectious clones is dependent on the synthesis of viral proteins, we transfected the CrPV-2 stop codon mutant RNAs and assessed the formation of CPE, viral RNA replication, and viral protein synthesis. As expected, S2 cells transfected with CrPV-2-ORF1-STOP RNA failed to display CPE as this mutation prevents expression of nonstructural proteins (Figure 2.4, 2.5A). Interestingly, transfection of CrPV-2-ORF2-STOP RNA resulted in CPE, suggesting that expression of ORF1 proteins can still induce CPE (Figure 2.5A). To confirm these results, Northern blot analysis showed no accumulation of viral RNA over time for CrPV-2-ORF1-STOP. However, transfection of the CrPV-2-ORF2-STOP led to a constant level of viral RNA, presumably due to the ongoing translation of the RdRp within ORF1 (Figure 2.5B). Accordingly, RT-PCR analysis detected negative sense viral RNAs in CrPV-2-ORF2-STOP transfected cells (Figure 2.5C). By contrast, CrPV-2-ORF1-STOP RNA-transfected cells produced very low levels of negative strand viral RNA (Figure 2.5C). The faint negative sense band observed in CrPV-2-ORF1-STOP reactions is likely attributed to the fact that negative sense RNA is also detected by RT-PCR using *in vitro* transcribed RNA only as a template (data not shown). Thus, these results suggest that replication is not occurring in CrPV-2-ORF1-STOP transfected cells. Viral non-structural (3CD antibody) and structural proteins (VP2 antibody) were not detected by Western blot analysis in both CrPV-2-ORF1-STOP and CrPV-2-ORF2-STOP transfected cells (Figure 2.5D). The observation that the non-structural protein RdRp (3CD) is absent in CrPV-ORF-2-STOP RNA-transfected cells despite displaying

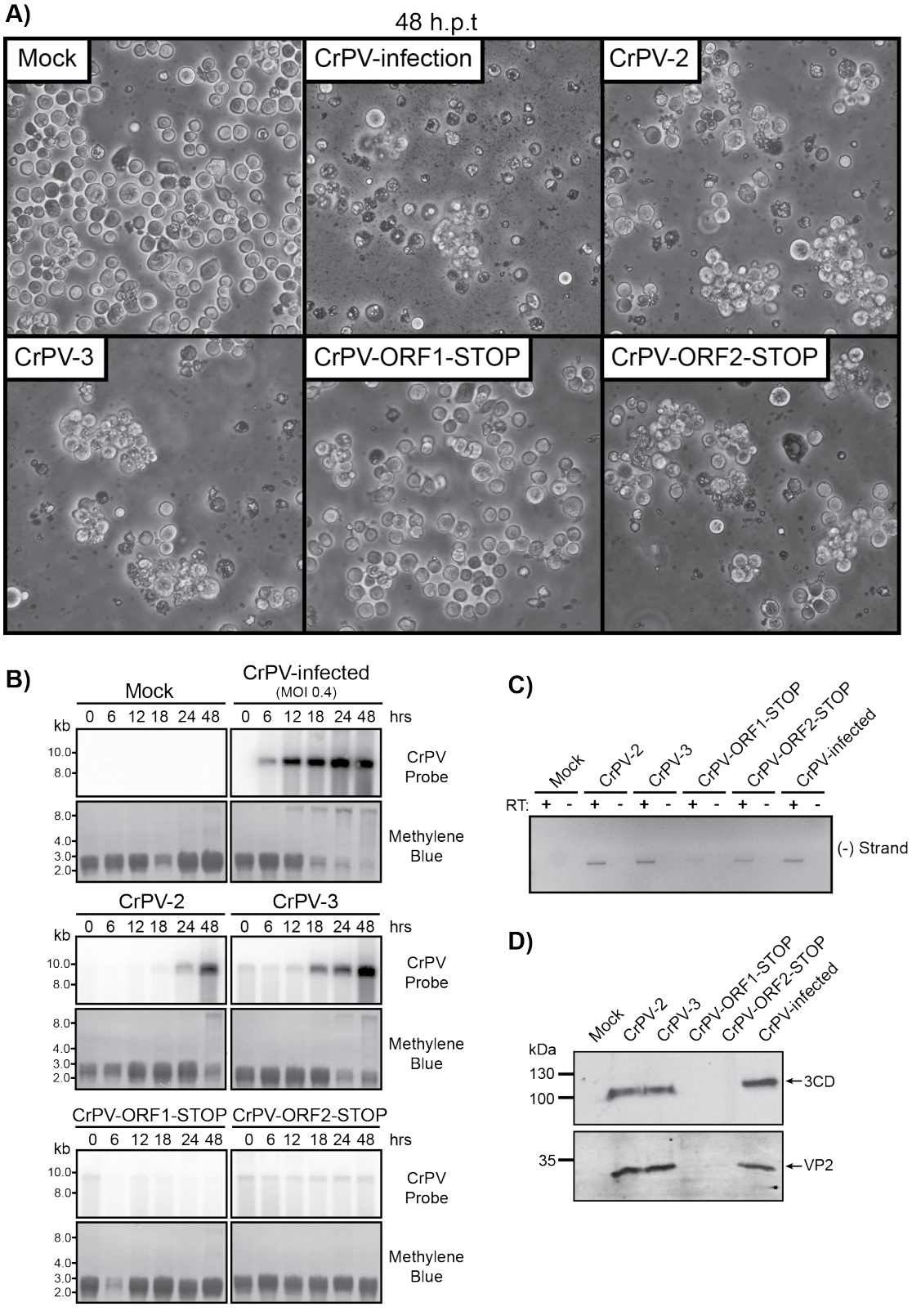

A)
48 h.p.t
Mock
CrPV-infection
CrPV-2
CrPV-3
CrPV-ORF1-STOP
CrPV-ORF2-STOP
B)
Mock
CrPV-infected
(MOI 0.4)
0 6 12 18 24 48 hrs
kb
10.0
8.0
CrPV Probe
8.0
4.0
3.0
2.0
Methylene Blue
CrPV-2
CrPV-3
CrPV-ORF1-STOP
CrPV-ORF2-STOP
C)
Mock
CrPV-2
CrPV-3
CrPV-ORF1-STOP
CrPV-ORF2-STOP
CrPV-infected
RT: + - + - + - + - + - + -
(-) Strand
D)
kDa
130
100
35
3CD
VP2

Figure 2.5 Transfection of *in vitro* transcribed CrPV-2 and CrPV-3 RNA in S2 cells.

(A) Phase contrast images of S2 cells mock-transfected, CrPV-infected (MOI 0.4) or transfected with *in vitro* transcribed CrPV-2, CrPV-3, CrPV-2-ORF1-STOP, or CrPV-2-ORF2-STOP RNA at 48 h post transfection (h.p.t) or post infection (h.p.i). (B) Northern blots of CrPV RNA genome from RNA isolated from CrPV-infected S2 cells or cells transfected with the indicated CrPV genomic RNA. Methylene blue staining of blots are shown below. (C) RT-PCR analysis of CrPV negative (-) strand synthesis from RNA of cells transfected with the indicated CrPV genomic RNA or CrPV-infected at 24 h.p.t. Reactions in the presence (+RT) or absence (-RT) of reverse transcriptase are shown. (D) Western blots of viral 3CD (ORF1) and VP2 (ORF2) from lysates of cells transfected with the indicated CrPV genomic RNA or CrPV-infected at 48 h.p.t.

---

negative strand synthesis and CPEs reflects that multiple rounds of infection after transfection are needed to detect viral proteins whereas RT-PCR analysis is sensitive enough for detection. In fact, in cells transfected with CrPV-2 or CrPV-3, viral proteins are not detected until 48 h.p.t. (Figure 2.5D). In summary, transfection of the CrPV-2 and CrPV-3 clones in S2 cells can lead to CPE, viral negative strand synthesis, and viral protein expression.

2.3.4 S2 cells transfected with CrPV-2 or CrPV-3 RNA produce infectious virions

To determine whether virions are produced, particles were isolated from CrPV-2- or CrPV-3-transfected S2 cells at 24 h post transfection. Transmission electron microscopy revealed the presence of virus-like particles in both cases with a diameter of approximately 26 nm and an icosahedral shape, which is similar to particles isolated from CrPV-infected S2 cells (Figure 2.6A).

To address whether the CrPV-2 and CrPV-3 viral particles produced by cells are infectious, we employed a two-pronged approach. First, a transwell assay coupled with immunofluorescence was used to determine if cells transfected with CrPV-2 or CrPV-3 RNA can infect naïve cells. Transfected or infected cells were seeded on the top of the transwell insert, which overlays naïve

S2 cells at the bottom of the well. Particles produced by the transfected cells will traverse the transwell membrane and infect the naïve cells. We monitored infection of naïve cells by expression of the structural VP2 protein using immunofluorescence analysis. As expected, cells transfected with CrPV-2 RNA, CrPV-3 RNA, or infected with CrPV displayed VP2 expression whereas mock-transfected cells or cells transfected with CrPV-2-ORF1-STOP or CrPV-2-ORF2-STOP RNA did not (Figure 2.6B). Second, we measured viral titre over time of cells transfected with CrPV-2 or CrPV-3 RNA (Figure 2.7). As expected, clones containing stop codons in either ORF1 or ORF2 produced no detectable titre up to 48 h.p.t. By contrast, infectious virions could be detected as early as 24 h.p.t for cells transfected with CrPV-2 or CrPV-3 RNA. Interestingly, a significantly higher amount of CrPV-3 was detected than CrPV-2 at both 24 and 48 h.p.t (Figure 2.7). This may reflect the earlier accumulation of CrPV-3 viral RNA observed previously by Northern blot analysis (Figure 2.5B). Altogether, these results strongly demonstrate that the CrPV-2 and CrPV-3 RNA are infectious in S2 cells.

2.3.5 Infection by CrPV-2 and CrPV-3 clones is not dependent on helper viruses

Insect RNA viruses are notoriously difficult to clone as illustrated in a recent review (Carillo-Tripp et al., 2015). A main challenge is that insect cells lines are commonly persistently infected with RNA viruses, such as Flock house virus (FHV) and Drosophila X virus (DXV), which can lead to artifacts in the development of infectious viral clones (Carillo-Tripp et al., 2015). To determine if CrPV-2 and CrPV-3 can recapitulate infection in the absence of other viruses, we first tested for the presence of three viruses commonly found in S2 cells lines, DXV, FHV, and Drosophila C virus (DCV; another dicistrovirus) (Flynt et al., 2009; Huszar and Imler, 2008; Nayak et al., 2010a; Wu et al., 2010).

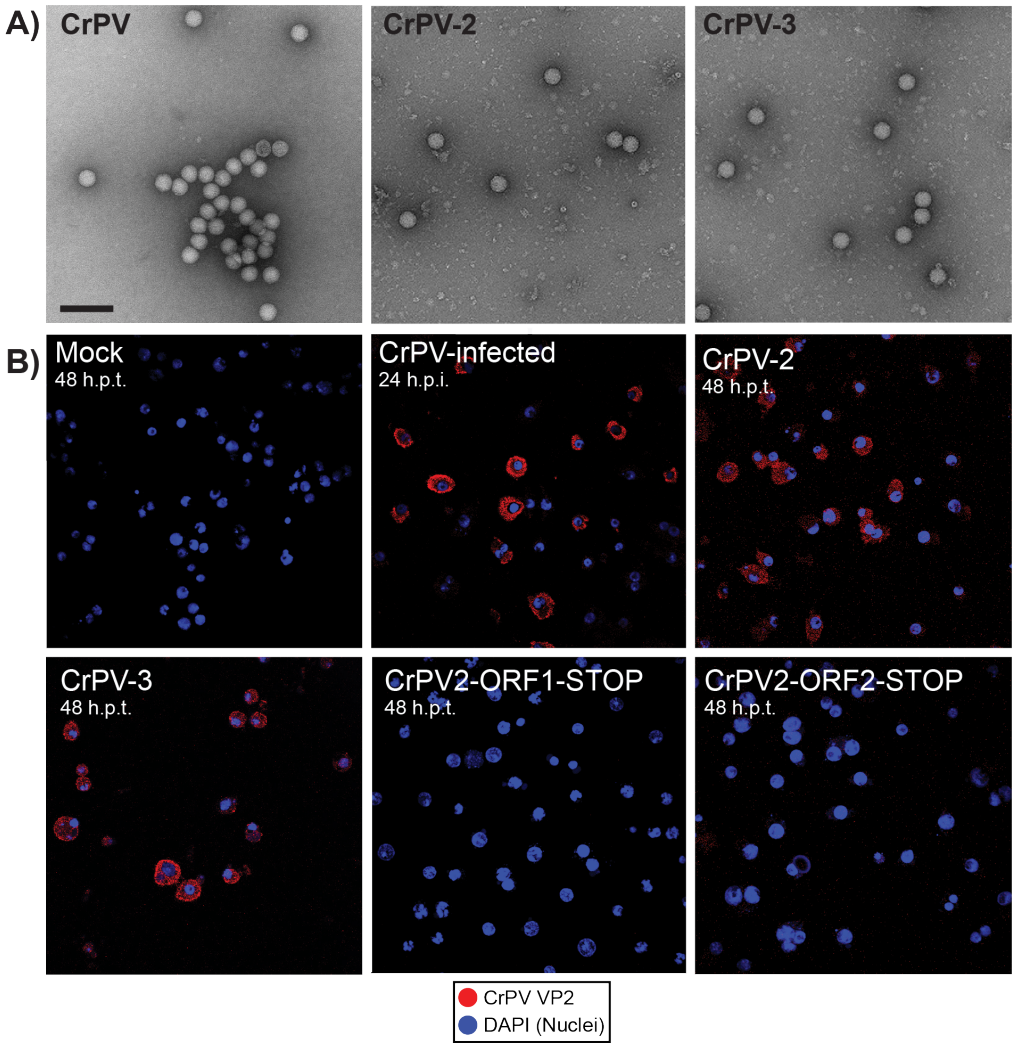


Figure 2.6 Transfected CrPV RNA clones produce infectious viral particles.

(A) Negative stain electron micrographs of viral particles purified from CrPV-infected S2 cells (left panel) or from S2 cells transfected with CrPV-2 or CrPV-3 RNA at 48 h post transfection (right panel). Scale bar represents 100 nm. (B) Transwell assay. S2 cells transfected with the indicated *in vitro* transcribed CrPV genomic RNA or infected with CrPV at an MOI of 0.4 for 24 h were seeded on a 0.4 µm transwell insert, which overlay naive S2 cells at the bottom of the well. Cells were then incubated for 24 or 48 h. Cells were analyzed by indirect immunofluorescence using an anti-CrPV VP2 (red) and the nuclei stained with DAPI (blue). Shown are representative images from at least three independent experiments.

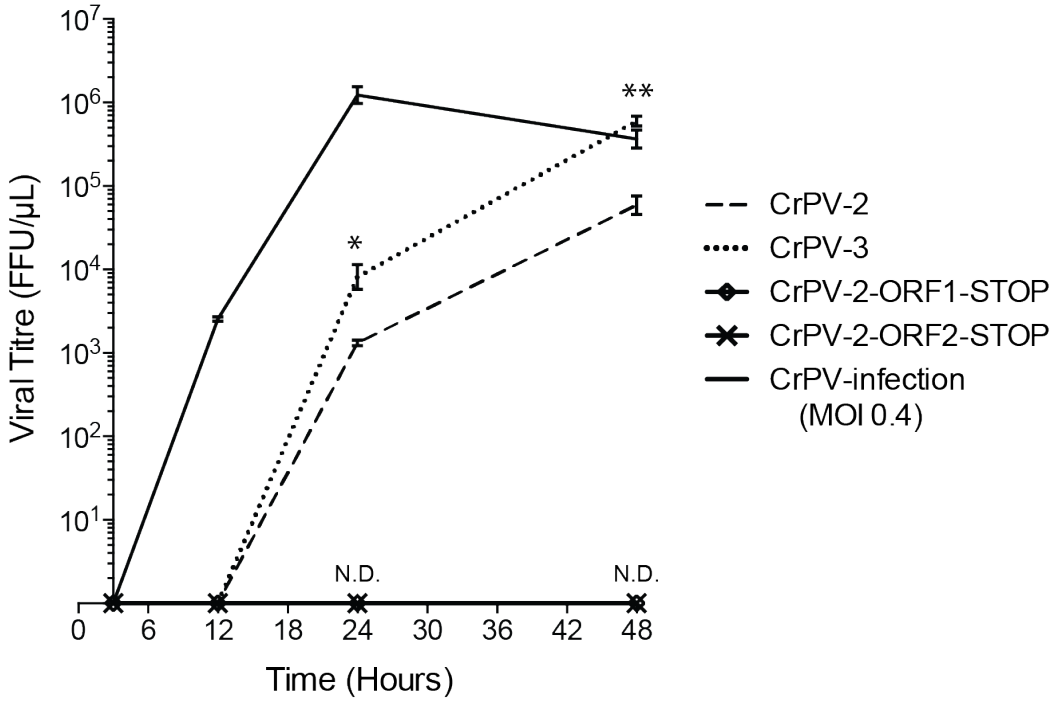


Figure 2.7 S2 Cells transfected with *in vitro* transcribed CrPV-2 and CrPV-3 RNA accumulate infectious virions over time.

2.5 x 10^6 S2 cells were transfected with 3 µg of the indicated *in vitro* transcribed genomic RNA or infected with CrPV (MOI 0.4). Titres were measured as described in the Materials and Methods at 3, 12, 24, and 48 h.p.t. Shown are averages from at least 3 independent experiments (±S.D.). *$P < 0.05$, **$P < 0.005$. N.D. = Not detected.

RT-PCR analysis of S2 cells within our lab (UBC S2 cells) against the DCV genome, DXV genome segment B, or the FHV genome segment RNA1 tested positive for the presence of DXV and FHV (Figure 2.8A). By contrast, another *Drosophila* cell line, Kc167, tested positive for only DXV while S2 cells from Invitrogen were negative for all three viruses, as previously reported (Figure 2.8A) (Nayak et al., 2010a). We next tested the capacity for CrPV-2 and CrPV-3 to replicate in Kc167 and Invitrogen S2 cells. In both cases, cells transfected with either CrPV-2 or CrPV-3 RNA accumulated viral RNA over time as analyzed by Northern blot, albeit the accumulation was delayed compared to our stock of S2 cells (UBC S2 cells) (Compare Figure 2.5B to 2.8B). Conversely, cells transfected with CrPV-ORF1-STOP RNA did not accumulate viral RNA over time for either Kc167 or Invitrogen S2 cells (Figure 2.8B). Viral titres of Kc167 and Invitrogen S2 cells transfected with CrPV-2 or CrPV-3 RNA produced infectious virions (Table 2.2). In agreement with UBC S2 cells, CrPV-3 had a reproducibly higher titre than CrPV-2. Taken together, the replication of CrPV-2 and CrPV-3 in S2 cells is not dependent on these viruses.

2.3.6 CrPV-2 and CrPV-3 virions are infectious in *Drosophila melanogaster* flies

CrPV can infect *Drosophila*, which has served as a model system to study virus host interactions (Costa et al., 2009; Kemp et al., 2013; Nayak et al., 2010a; van Rij et al., 2006; Wang, 2006). To determine whether CrPV-2 and CrPV-3 is infectious in adult *Drosophila* other than in tissue culture cells, we first propagated CrPV-2 and -3 virus after transfection in S2 cells. We also used CrPV that has been propagated in S2 cells. 5000 fluorescent focus forming units (FFU) of CrPV, CrPV-2 or CrPV-3, UV-inactivated viruses or PBS were injected intrathroacically into *D. melanogaster* (*iso-w*^{1118}). Mortality was observed in adult flies injected with CrPV, CrPV-2, and CrPV-3 at two days post-injection (Figure 2.9).

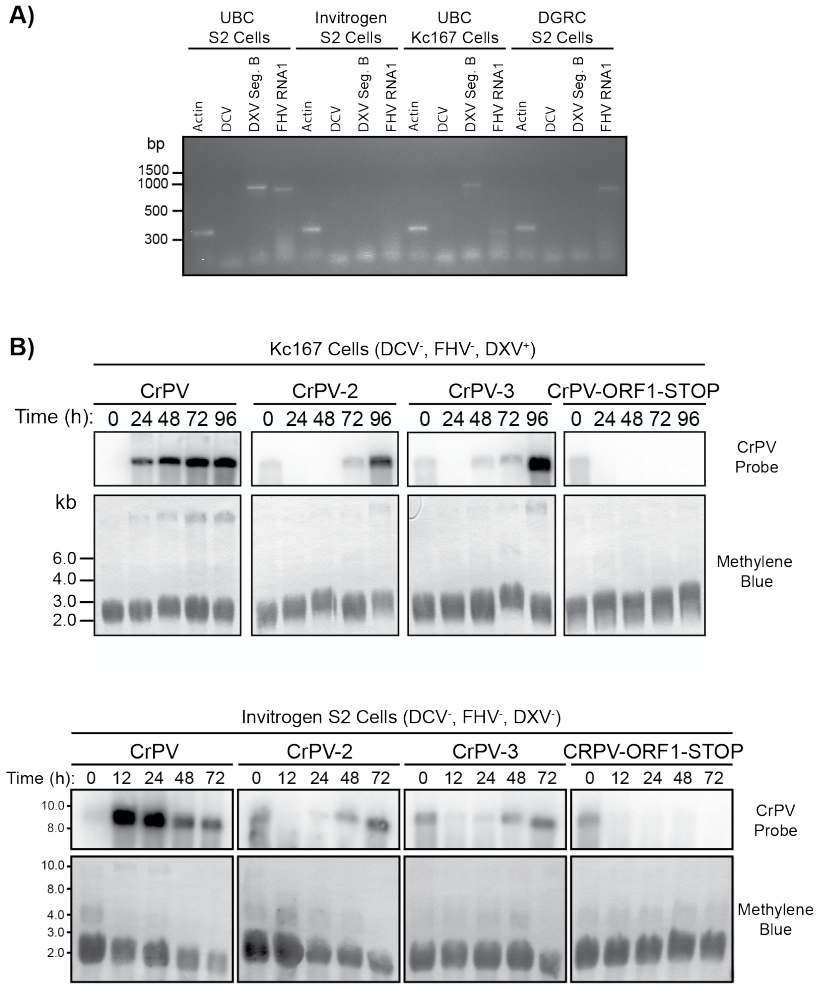


Figure 2.8 CrPV infectious clones can infect Drosophila S2 cells devoid of Drosophila C virus, Flock house virus and Drosophila X virus.

(A) Detection of Drosophila C virus (DCV), Flock house virus (FHV) and Drosophila X virus (DXV) by RT-PCR analysis in UBC S2 cells, Kc167 cells, Invitrogen S2 cells, and S2 cells from DGRC. (B) Northern blots of CrPV RNA genome from RNA isolated from Kc167 cells and S2 cells from Invitrogen that were transfected with the indicated CrPV genomic RNA. Methylene blue staining of blots are shown below.

Table 2.2 Titre of CrPV-2, CrPV-3, and CrPV-ORF1-STOP in Invitrogen S2 cells and Kc167 cells.

Cell Type	Transfected RNA	Titre* (FFU/µL)
Invitrogen S2 Cells (48 h.p.t.)	CrPV-2	2.34×10^4
	CrPV-3	5.47×10^4
	CrPV-ORF1-STOP	N.D.
Kc167 Cells (96 h.p.t.)	CrPV-2	1.03×10^5
	CrPV-3	2.09×10^6
	CrPV-ORF1-STOP	N.D.

*Shown is the average titre determined from duplicate experiments.
N.D. = Not detected.

By day 4, all flies injected with either CrPV, CrPV-2, or CrPV-3 succumbed to death, however those injected with PBS, or UV-inactivated virus survived and remained healthy (Figure 2.9). These results strongly suggest that like CrPV, CrPV-2 and -3 virions are infectious in adult *D. melanogaster*.

2.3.7 The duplication in the 5'UTR of CrPV-2 is stable and reduces viral fitness

The observation that the 5'UTR containing the duplication stimulated translation prompted us to test the hypothesis that CrPV may have acquired this duplication to influence viral fitness. Strikingly, cells transfected with CrPV-3 RNA accumulated viral RNA much earlier and had significantly higher viral titres than those cells transfected with CrPV-2 RNA (Table 2.2; Figures 2.5B and 2.7). Altogether, the duplication within the 5'UTR of CrPV-2 retarded viral RNA synthesis and overall production of infectious virions as compared to CrPV-3 despite having a stimulatory effect on translation.

It is possible that the fitness loss seen in CrPV-2 is due to either a defect in viral RNA replication, translation, egress, or entry. To determine if the fitness loss is due to a disruption in the first replication cycle of CrPV, we infected S2 cells with CrPV-2 or CrPV-3 at an MOI of 10 and monitored viral protein and RNA synthesis (Figure 2.10A). Viral protein synthesis was similar for both CrPV-2 and CrPV-3 as detected by [^{35}S]-met/cys pulse-labeling (Figure 2.10A). Despite the enhanced translation from the 5'UTR^{+dup} observed with reporter constructs, there was no visible enhancement in the production of ORF1 proteins during CrPV-2 infection by immunoblot analysis (Figure 2.10A). Furthermore, there was no detectable difference in RNA accumulation between CrPV-2 and CrPV-3 as visualized by Northern blot analysis (Figure 2.10A). Altogether this suggests that there is no defect in viral RNA replication or translation between CrPV-2 and CrPV-3 infections in S2 cells.

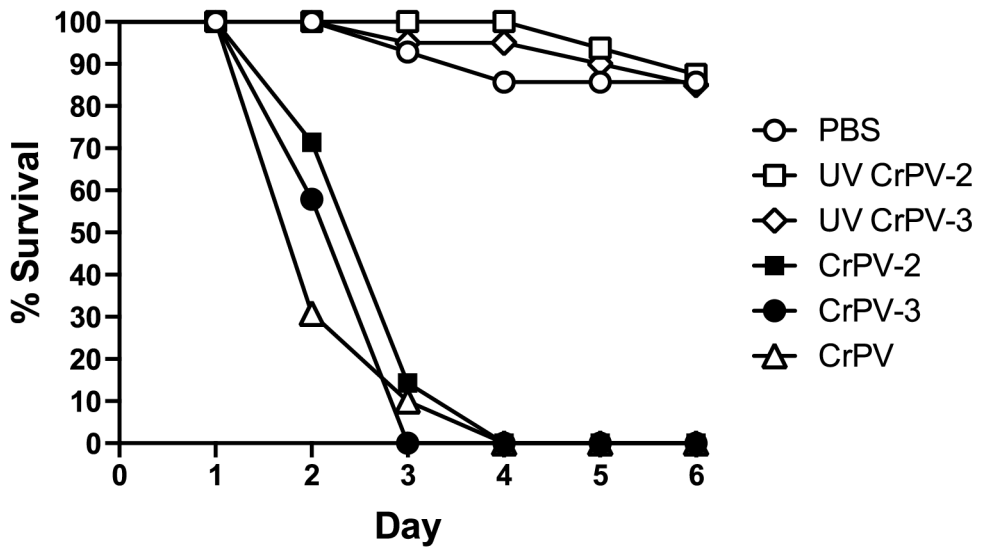


Figure 2.9 Injection of CrPV-2 and CrPV-3 virions into adult *Drosophila melanogaster* flies.

Virgin *Iso* w^{1118} flies (10 males and 10 females) were injected intrathoracically with 5000 FFU of CrPV, CrPV-2, CrPV-3 or UV inactivated CrPV-2 and CrPV-3 or PBS. Subsequently, flies were flipped onto standard media and survival was monitored daily. Shown is a representative graph from at least three independent experiments.

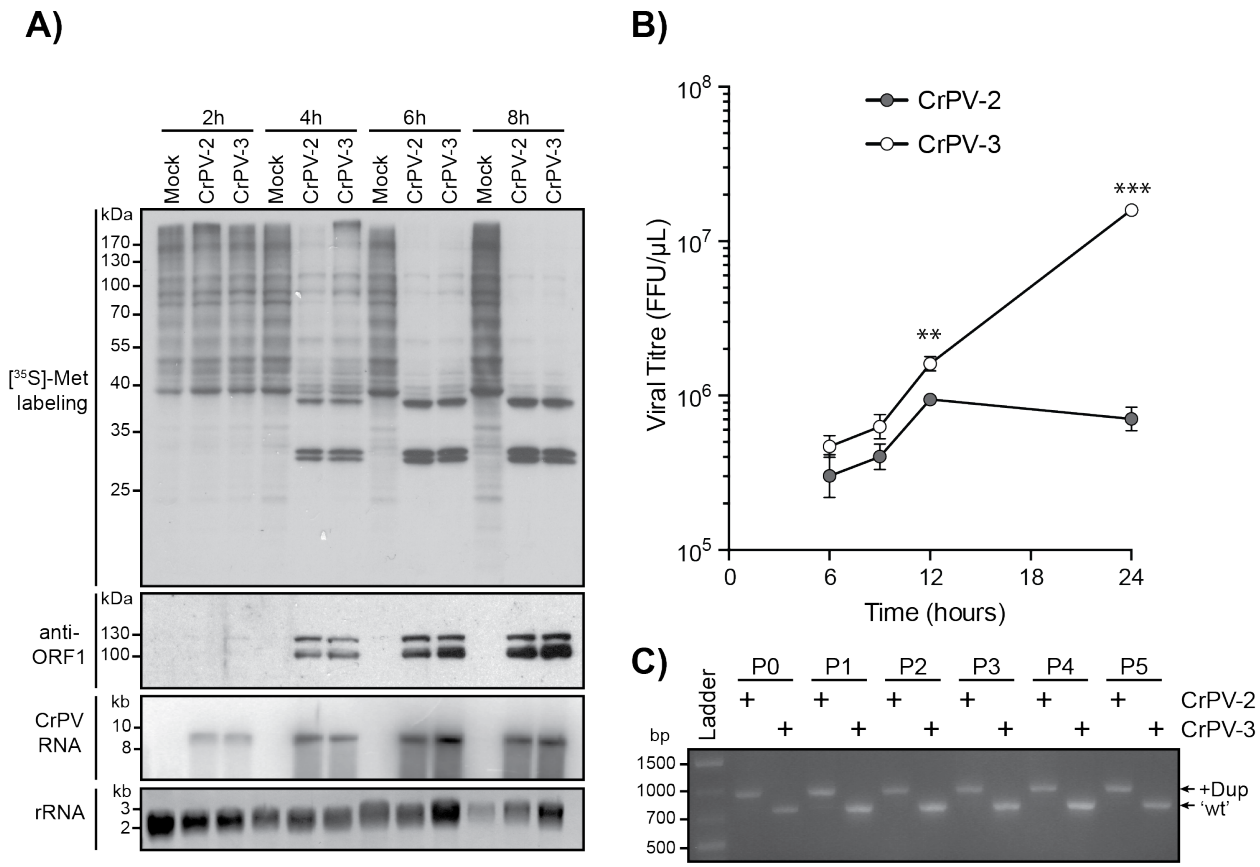


Figure 2.10 The CrPV-2 5'UTR reduces viral fitness.

(A) S2 cells were mock-infected or infected with CrPV-2 or CrPV-3 (MOI 10). Cells were metabolically labeled with [^{35}S]-methioine/cysteine for the last 30 minutes of infection. Lysates were subjected to SDS-PAGE, immunoblot or Northern blot analysis. Accumulation of viral RNA was monitored by probing for CrPV RNA genome. Expression of CrPV ORF1 was assessed by anti-RdRp antibody. (B) S2 cells were infected with CrPV-2 or CrPV-3 (MOI 1) and viral titres were measured as described in the Materials and Methods at 6, 9, 12, and 24 h.p.i. $^{**}P < 0.005$, $^{***}P < 0.0005$. Shown are averages from at least 3 independent experiments (±S.D.). (C) RT-PCR analysis of the CrPV-2 and CrPV-3 5' UTRs after serial passage in S2 cells. Cells were transfected with either CrPV-2 or CrPV-3 RNA and virus was harvested (passage 0 [P0]). Cells were then infected at an MOI of 10 and passaged in S2 cells 5 times (P1 to P5). At each passage, RNA was isolated and subjected to RT-PCR analysis.

To determine if the phenotype exhibited by CrPV-2 occurs during the subsequent rounds of infection, we infected cells at a MOI of 1 and measured viral titre over time (Figure 2.10B). At 6 and 9 h.p.i., titres for CrPV-2 and CrPV-3 were not significantly different, however by 12 h.p.i the titre was reproducibly higher for CrPV-3 than CrPV-2 and continued to generate a log-fold more virus than CrPV-2 by 24 h.p.i, consistent with previous observations of cells transfected with CrPV infectious clones (Compare Figures 2.7 and 2.10B). To ensure that the 5'UTR duplication in CrPV-2 is stable and is not simply lost during the course of infection, we passaged CrPV-2 and CrPV-3 in S2 cells for 5 passages and performed RT-PCR analysis of the 5'UTR. Our results showed that the 5'UTR duplication is stable in CrPV-2 and that CrPV-3 does not gain the duplication during these passages (Figure 2.10C). Overall, we conclude that the duplication in the 5'UTR of CrPV-2, despite having increased translational activity in the context of reporter constructs, reduces viral fitness compared to a virus lacking this extra element. Furthermore, this fitness loss likely does not occur at the replication step of the CrPV life cycle, but may occur at another step such as viral packaging or entry into uninfected cells.

2.4 DISCUSSION

Despite the identification of several members of the *Dicistroviridae* family (Leat et al., 2000; Mari et al., 2002; Moon et al., 1998; Reddy et al., 2013; Reddy et al., 2014; Sasaki et al., 1998; Wilson et al., 2000b), the establishment of an infectious clone has been elusive (Carillo-Tripp et al., 2015). In this study, we generated infectious molecular clones of CrPV, termed CrPV-2 and CrPV-3 (Figure 2.3A). Several lines of evidence support this conclusion: Transfection of the CrPV-2 or CrPV-3 RNA in *Drosophila* S2 cells results in (i) cytopathic effects, (ii) expression of viral nonstructural and structural proteins by Western blot analysis, (iii)

synthesis of negative strand viral RNA by tagged RT-PCR analysis, (iv) and accumulation of viral RNA and infectivity over time as detected by Northern blot analysis and viral titres. (v) Furthermore, particles produced from CrPV-2- and CrPV-3-transfected S2 cells are infectious in naive cells using a transwell assay and that the (vi) CrPV-2 and CrPV-3 infectious particles resemble the shape and size of natural CrPV virions by EM. (vii) Importantly, insertion of stop codons within the viral coding regions of CrPV-2 attenuated virus infectivity. (viii) Finally, both CrPV-2 and CrPV-3 virions produced from S2 cells are infectious resulting in mortality when injected into adult *Drosophila*. In summary, we have conclusively demonstrated the generation of a robust dicistrovirus molecular clone that can be propagated in *Drosophila* S2 cells. Given that many studies use dicistroviruses as a model system, the CrPV-2 and CrPV-3 clones should provide useful tools to study the life cycle of dicistroviruses and host-virus interactions in insects.

During infection, RNA viruses largely exist as quasispecies (Lauring and Andino, 2010), thus the accumulation of nucleotide changes that deviate from the published sequence is not unexpected. CrPV-2 and CrPV-3 contain several nucleotide changes that likely evolved during viral propagation in S2 cells (Table 1; Figure 2.3B). None of the changes observed in either CrPV-2 or CrPV-3 are found in known active sites, although the majority occurs within the viral RdRp (3D), which may reflect changes in the RdRp for optimal RNA replication in S2 cells.

The most striking difference between CrPV-2 and the published CrPV sequence is the presence of a 196-nucleotide duplication within the 5'UTR (Table 2.1; Figure 2.3B). How did this duplication come about within the CrPV 5'UTR? RNA recombination occurs frequently in some (+) ssRNA viruses during genome replication and can drive viral evolution (Simon-Loriere and Holmes, 2011). The 5'UTR duplication may have arisen from a template-switching event

during genome replication, and subsequently selected for as a fitness gain. To our surprise, viral infectivity of CrPV-2 is reduced, not enhanced, when compared to a clone of CrPV that lacks the 5'UTR duplication (CrPV-3) even though the 5'UTR^{+Dup} stimulates translation (Figures 2.2, 2.7, 2.8B and 2.10B). CrPV-2 produced 1-log fold fewer infectious particles than CrPV-3 upon RNA transfection and infection at a low MOI, supporting the notion that infection is reduced with the duplication in the 5'UTR (Figures 2.7 and 2.10B). In transfected S2 cells, CrPV-3 RNA accumulated earlier than CrPV-2 RNA suggesting that there is a reduction in viral replication when the duplication is present (Figure 2.5B). Strikingly, in cells that are infected with CrPV-2 at a low MOI, replication appears to halt at 12 h.p.i., which is in contrast to that observed in cells that are transfected with CrPV-2 RNA where replication is seen up to 48 h.p.t. (Figures 2.7 and 2.10B). The reason for this difference is not clear, however, it may be due to the number of rounds of infection that have occurred in each experiment. Interestingly, infection at a high MOI did not produce any observable differences in replication or viral protein synthesis in cells infected with CrPV-2 or CrPV-3 (Figure 2.10A). Altogether, these results suggest that the rate-limiting step for CrPV-2 replication may occur at viral packaging or entry. A recent report suggests that the spatial organization of packaging signals within the RNA sequence of ssRNA viruses is important to ensure efficient capsid assembly (Patel et al., 2015). It is possible that the duplication in the 5'UTR of CrPV-2 may disrupt the proper binding of capsomers to the viral RNA thus hindering efficient assembly. Alternatively, entry of some RNA viruses into host cells is dependent on conformational changes in their capsids and RNA. For example, upon exposure to acidic pH in the endosome, Rhinovirus capsids undergo a conversion to the porous subviral A-particle, which is then followed by a conformational change in the RNA and exit of the genome in the 3' to 5' direction into the host cytosol (Harutyunyan et al., 2013; Pickl-Herk et al., 2013).

The duplication in the 5'UTR of CrPV-2 may impede release by interfering with either RNA-protein or RNA-RNA contacts in the capsid. A recent study examined the structure of the CrPV 5'UTR revealing that it is composed of three domains containing multiple stem-loop and hairpins with three- and four-way junctions (Gross et al., 2017). While Domains II and III were required for IRES activity of the 5'UTR, Domain I was dispensable (Gross et al., 2017). Interestingly, the duplication we observed occurs directly within Domain I. It is possible that this region may be implicated in replication of the virus or may assist in directing ribosomes to the core region of the IRES before translation occurs. These hypotheses need to be examined further. Finally, the 5'UTR of CrPV-2 may simply be the consequence of the founder effect. The viral population is subjected to bottlenecks (i.e. inoculum size) allowing founder viruses with suboptimal fitness levels to succeed due to random sampling (Manrubia et al., 2005). This effect may explain how a low fitness strain of CrPV could emerge over a high fitness strain (e.g. CrPV-2 vs CrPV-3).

Viral infection in S2 cells was established by transfecting *in vitro* transcribed CrPV-2 and -3 RNA that lacked a 5'-linked VPg and a 3' poly (A) tail, both of which are naturally present on the CrPV RNA genome (Bonning and Miller, 2010). It is not surprising that CrPV-2 and CrPV-3 RNA can replicate in the absence of these elements as they are also dispensable for replication of other RNA viruses (Flanegan et al., 1977; Racaniello and Baltimore, 1981). However, the absence of a 5' linked VPg and a poly(A) tail may explain the delay in viral replication in CrPV transfected cells compared to infected cells (compare Figures 2.7 and 2.10). It remains to be investigated if inclusion of these elements to the CrPV infectious clones facilitates infection.

The 5'UTR of CrPV contains an IRES that facilitates the translation of viral non-structural proteins (Wilson et al., 2000b). Our results showed that the duplication element within

the CrPV-2 5'UTR stimulated translation only when positioned within the 5'UTR of reporter mRNAs (Figure 2.2B and 2.2C). This result may not be surprising as it reflects the natural location of this element within the viral genome. How does the duplication element stimulate 5'UTR translation? In general, the mechanism of translation for the majority of dicistrovirus 5'UTR IRESs is poorly understood. Moreover, the dicistrovirus 5'UTRs vary in length and the nucleotide sequences are not well conserved (Jan, 2006; Roberts and Groppelli, 2009). The lack in conservation may be due to the unique translation factor requirements of each dicistrovirus 5'UTR IRES to drive viral translation in their specific host cells. Indeed, not all 5'UTR IRESs function in all assays: the 5'UTR IRESs of the CrPV and the related dicistroviruses *Plautia stali intestine virus* and RhPV can each mediate translation in a subset of species-specific translation extracts (Groppelli et al., 2007; Shibuya and Nakashima, 2006; Wilson et al., 2000b; Woolaway et al., 2001). Through reconstitution experiments, the RhPV 5'UTR IRES absolutely requires the canonical translation factors, eIF1, eIF2 and eIF3, for efficient 48S complex formation and translation initiation and that this complex is stimulated by eIF1A, eIF4A and eIF4F (Terenin et al., 2005). It is possible that the enhanced translational activity observed with the CrPV 5'UTR duplication may be due to either an increase in initiation factor or ribosome recruitment or modulation in flexibility of its RNA structure in order to facilitate translation, the latter of which would be in agreement with previous reports that the RhPV 5'UTR is unstructured (Groppelli et al., 2007; Terenin et al., 2005). However, the enhanced translation observed from the 5'UTR in the context of the minigenome reporter appears at odds to that observed during infection of CrPV-2 and CrPV-3 (Figure 2.10A). It is likely that the minigenome reporter mRNAs are not regulated in the same manner as the viral genome, thus further investigations are needed to elucidate the role of the duplication in CrPV 5'UTR translation using the infectious clones.

In conclusion, we have constructed the first infectious clones of a dicistrovirus. The CrPV clones are infectious to both *Drosophila* S2 cells and adult flies, thus providing invaluable tools in dissecting the molecular mechanisms behind the *Dicistroviridae* life cycle, pathogenesis, and interactions with the host. Furthermore, the CrPV clones may provide a framework to develop other dicistrovirus infectious clones that have been difficult to obtain. Our *in vitro* system showed that CrPV-2 and CrPV-3 RNA are capable of producing CrPV nonstructural and structural proteins and that processing of the polyproteins is dependent on ORF1 protein expression, indicating that the viral 3C protease is active in Sf-21 cell extracts (Figure 2.4) (Garrey et al., 2010; Moore et al., 1980; Wilson et al., 2000b). Akin to poliovirus, it may be possible to exploit this *in vitro* system to investigate CrPV-2 or CrPV-3 replication and assembly (Barton et al., 1995; Molla et al., 1991). Finally, because CrPV infections have a wide host range, the infectious clones can be exploited for biological controls such as management of insect pests and induction of innate immune responses or expression of antiviral proteins in insects and arthropods.

Chapter 3: Analysis of CrPV IGR IRES-mediated translation during viral infection

3.1 INTRODUCTION

Canonical eukaryotic translation initiation is a highly orchestrated series of steps involving 40S recruitment to the 5'cap, scanning, 80S assembly and initiation at an AUG codon (Jackson et al., 2010). Twelve core translation factors are required to mediate cap-dependent translation (Jackson et al., 2010). All viruses utilize the host translation machinery for viral protein synthesis and some have evolved ingenious strategies to commandeer host ribosomes for their own benefit. Internal ribosome entry sites (IRESs) are one of the most well studied translation initiation mechanisms employed by some viruses to facilitate expression of their genomes. In general, IRESs are structured RNAs that directly recruit the ribosome using a subset of translation factors, thus providing an advantage during infection when cap-dependent translation is compromised (Au and Jan, 2014; Plank and Kieft, 2012). IRESs are classified based on nucleotide and structural conservation as well as mechanism of translation initiation. The picornavirus IRESs require most of the canonical translation initiation factors and IRES-trans-acting factors whereas the hepatitis C virus-like IRESs are streamlined requiring only eIF2 and eIF3 to recruit the ribosome and initiate translation (Pestova et al., 1998; Sweeney et al., 2014). Extensive biochemical studies and the identification of factors required for IRES translation have yielded detailed insights into IRES mechanisms. The challenge is to identify the key steps and factors in IRES translation that impact virus infection.

One of the most well-studied and perhaps simplest IRES to date is within the intergenic region of the *Dicistroviridae*. Dicistroviruses are single-stranded positive sense RNA viruses with genome sizes ranging from 8-10 kb that infect arthropods (Bonning and Miller, 2010).The

namesake of these viruses stems from their unique dicistronic genome arrangement where each open reading frame (ORF) is driven by a distinct IRES which allows differential and temporal regulation of each ORF during infection (Khong et al., 2016; Wilson et al., 2000b). The 5' untranslated region (5'UTR) IRES directs translation of ORF1, which encodes the viral non-structural proteins, such as the RNA-dependent RNA polymerase and 3C protease (Figure 1). Whereas the aforementioned intergenic region IRES (IGR IRES) facilitates translation of the viral structural proteins encoded in the second ORF (ORF2)(Bonning and Miller, 2010).

Through a unique mechanism, the IGR IRES binds directly to 40S and 80S ribosomes without the assistance of initiation factors or initiator Met-tRNA and initiates translation from a non-AUG codon (Sasaki and Nakashima, 2000; Wilson et al., 2000a). Several structural and *in vitro* biochemical studies have revealed that the CrPV IGR IRES adopts a structure comprised of three pseudoknots (termed PKI, II, and III) and two stem-loops (SLIV and SLV)(Costantino et al., 2008; Fernandez et al., 2014; Jan and Sarnow, 2002; Kanamori and Nakashima, 2001; Muhs et al., 2015; Muhs et al., 2011; Murray et al., 2016; Pfingsten et al., 2006; Schuler et al., 2006). Together PKII and PKIII form the ribosome-binding domain, which mediates 40S and 80S recruitment, and the PKI domain functionally mimics the anticodon stem of a tRNA, which permits the IGR IRES to first occupy the ribosomal A site (Costantino et al., 2008; Fernandez et al., 2014; Jan and Sarnow, 2002; Muhs et al., 2015; Pfingsten et al., 2006; Schuler et al., 2006; Spahn et al., 2004). Upon initial occupancy of the A site by PKI, eukaryotic elongation factor 2 (eEF2) facilitates a pseudo-translocation step that involves movement of the PKI domain to the P site, leaving the A site clear for delivery of the first aminoacyl-tRNA by eEF1A (Fernandez et al., 2014; Jan et al., 2003; Pestova and Hellen, 2003). The IRES undergoes a second eEF2-mediated translocation event without peptide bond formation; afterwards elongation then

proceeds (Abeyrathne et al., 2016; Jan et al., 2003; Pestova and Hellen, 2003; Petrov et al., 2016; Zhang et al., 2016). The tRNA-like anticodon domain of the PKI domain sets the reading frame for IRES translation as it translocates through the decoding center of the ribosome (Costantino et al., 2008; Kamoshita et al., 2009; Zhu et al., 2011). The IGR IRES makes specific contacts with both ribosomal subunits. Stem-loops SLIV and SLV interact with uS7 and eS25 of the 40S subunit. Biochemical and structural data suggest that that uS7 and eS25 interact with and bridge both SLV and SLIV, thus these ribosomal proteins may have redundant yet crucial roles in binding to and positioning of the IGR IRES relative to the 40S subunit (Fernandez et al., 2014; Landry et al., 2009; Muhs et al., 2011; Nishiyama et al., 2007). In support of these interactions, IGR IRES translation is abrogated in yeast lacking eS25 (Landry et al., 2009). The conserved L1.1 loop of the IGR IRES interacts with the L1 stalk of the 60S ribosomal subunit, which is reminiscent of interactions of the L1 stalk with an E-site deacylated tRNA (Jang et al., 2009; Pfingsten et al., 2006). The variable loop region (VLR) in the PKI domain, which is reported to facilitate ribosome positioning and eEF2-mediated pseudotranslocation, interacts with the β-hairpin loop of uS7 and helix 23 of 18S rRNA in the E site, suggesting that the VLR has a stabilizing role as the IRES translocates through the ribosome (Abeyrathne et al., 2016; Au and Jan, 2012; Ruehle et al., 2015). In addition to the ribosome, the IGR IRES may interact with eEF2 to facilitate its movement through the ribosome (Ruehle et al., 2015). IGR IRES recruitment of the ribosome is inhibited by depleting pseudouridylation of rRNA, suggesting that specific rRNA modifications can affect IRES function (Jack et al., 2011). Altogether, the contacts with the 80S ribosome allow the IGR IRES to operate as a highly tuned RNA element that can initiate translation in an unprecedented manner.

To date, studies on the *Dicistroviridae* IGR IRES have been limited to *in vitro* translation assays, reporter assays in tissue culture cells and orthologous systems (e.g. *S. cerevisiae,* cell extracts etc.). In this study, we use a recently developed infectious clone of CrPV, termed CrPV-3, which now allows the use of reverse genetics, to probe the IGR IRES mechanism in its native context(Kerr et al., 2015). Specifically, the physiological significance of well-established mutations in the IGR IRES is examined in the context of the entire CrPV genome and in CrPV-infected *Drosophila* cells.

3.2 MATERIALS & METHODS

3.2.1 Cell culture and virus.

Drosophila Schneider line 2 (S2) cells were maintained and passaged in Shield's and Sang medium (Sigma) supplemented with 10% fetal bovine serum. Determination of CrPV viral titres and yield were performed as previously described(Garrey et al., 2010). Briefly, a total of 1.5 x10^6 S2 cells we incubated with serial dilutions of virus for 30 min, then resuspended in media, plated into a 96-well plate coated with concavilin A (0.5 mg/mL; Calbiochem) and incubated at 25°C for 8 h. Cells were then washed with PBS before being fixed with 3% paraformaldehyde for 15 min followed by methanol for 10 min. The fixed cells were then incubated with an anti-ORF2 antibody (1:250 dilution in 5% bovine serum albumin in PBS) for 1 h at room temperature. Subsequently, cells were washed three times with PBS and incubated with a Texas Red IgG anti-rabbit (1:500 dilution in 5% bovine serum albumin in PBS; Invitrogen) for 1 h at room temperature. Finally, cells were washed with PBS stained with Hoechst dye (0.5 µg/mL). The amount of infected cells was quantified after plates were analyzed with a Cellomics Arrayscan HCS instrument. Through serial dilutions of CrPV, the FFU/mL can be calculated. Each titre is the result of at least three replicate experiments.

CrPV-3 and mutant viruses were generated from *Drosophila* S2 cells using an adapted protocol(Krishna et al., 2003). Briefly, 5.0 x 10^7 S2 cells were transfected with *in vitro* transcribed RNA derived from pCrPV-3 or mutant plasmids and incubated for 48h. Cells were dislodged into the media, treated with 0.5% Igepal CA-630 (Nonidet P-40) and 0.1% 2-mercaptoethanol, and incubated on ice for 10 min. Cell debris was cleared by centrifugation at 13,800 RCF for 15 min at 4°C. Viral particles were then concentrated by ultracentrifugation at 141,000 RCF for 2.5 h at 4°C. The pellet was resuspended in PBS and sterilized through a 0.2 μM filter. All viruses were sequence verified via RT-PCR with primers directed against the CrPV IGR IRES.

3.2.2 Construction of CrPV-3 mutants.

Deleterious mutations were introduced into the pCrPV-3 clone at nucleotides denoted in Table 1. A stop codon mutation (UAA) in ORF2 of pCrPV-2 was introduced by mutating nucleotide A6428T of ORF2(Kerr et al., 2015). The numbering of the nucleotides is based on CrPV-2 in this case. Note that the only difference between CrPV-2 and CrPV-3 is the presence of a 196-nt duplication in the 5'UTR that does not affect IGR IRES translation. All plasmids were fully verified by sequencing to ensure no other mutations were in place.

3.2.3 *In vitro* transcription and RNA transfection.

Purified pCrPV-3 and derivative plasmids were linearized with Ecl136II. RNA was transcribed in a T7 RNA polymerase reaction and subsequently purified with a RNeasy kit (Qiagen). The integrity and purity of the RNA was confirmed on a 1.2% denaturing formaldehyde agarose gel.

Transfection of *in vitro* synthesized RNA into S2 cells was performed using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions. 3 µg of RNA derived from either pCrPV-3 or its cognate mutants were used for transfection using 2.5 x 10^6 cells.

3.2.4 Northern blot analysis.

Total RNA was isolated from cells using TRIzol reagent. Northern blots were performed by loading 5 µg of RNA on a denaturing agarose gel and subsequently transferred to Zeta-probe blotting membrane (Bio-Rad). DNA probes were radiolabelled with a DecaLabel DNA labeling kit (Fermentas) and hybridized overnight. Radioactive bands were detected via phosphoimager analysis (Storm; GE Healthcare).

3.2.5 Western blots.

Equal amounts of S2 protein lysates (20 µg) were resolved on a 12% SDS-PAGE gel and then transferred to a polyvinylidene difluoride Immobilon-FL membrane (Millipore). Membranes were blocked for 30 min at room temperature with 5% skim milk in TBST. Blots were incubated for 1 h at room temperature with the following antibodies: CrPV ORF1 (raised against CrPV RdRp) rabbit polyclonal (1:10,000) or CrPV ORF2 (raised against CrPV VP2) rabbit polyclonal (1:10,000)(Garrey et al., 2010). Membranes were washed 3 times with TBST and incubated with either IRDye 800CW goat anti-rabbit IgG (1:20,000; LI-COR Biosciences) for 1 h at room temperature. An Odyssey imager (LI-COR Biosciences) was used for detection.

3.2.6 *In vitro* and *in vivo* translation assays.

In vitro translation of the full-length viral RNA genome (uncapped) was performed in *Spodoptera frugiperda* 21 (Sf-21) cell extract (Promega) in the presence of [^{35}S]-methionine/cysteine. Reactions were loaded on a SDS-PAGE. Gels were dried and radioactive bands were monitored by phosphoimager analysis.

For *in vivo* translation assays, mock- or CrPV-infected S2 cells (MOI 5) were incubated with [^{35}S]-methionine/cysteine for the last 30 minutes of the infection. Equal amounts of lysate (10 µg) were then loaded on a SDS-PAGE. Gels were dried and radioactive bands were monitored by phosphoimager analysis. Gels were quantified using ImageQuant software (GE Healthcare).

3.3 RESULTS

3.3.1 IGR IRES mutants of CrPV-3 can synthesize ORF2 proteins in Sf-21 extracts

To assess the relevance of IGR IRES translation in CrPV infection, we systematically introduced a panel of known mutations that affect specific properties of IGR IRES translation within the CrPV-3 infectious clone (Figure 3.1; Table 3.1). We generated a series of mutations as follows: (i) nucleotides $CC_{6214\text{-}15}$ to GG ($CC_{6214\text{-}15}GG$), which disrupts PKI base-pairing and effectively inhibits proper 80S positioning on the IRES (Jan and Sarnow, 2002). (ii) Nucleotides within the loop of SLIV. The SLIV loop (nucleotides $AUUU_{6111\text{-}14}$) makes contacts with ribosomal protein uS7 with the head of the 40S subunit. Mutating $AUUU_{6111\text{-}14}UAAA$ (mSLIV) markedly reduces translational activity and is reported to reduce 40S binding (Abeyrathne et al., 2016; Fernandez et al., 2014; Jan and Sarnow, 2002). (iii) Nucleotides within the loop or helical stem of SLV were altered. Mutating the loop of SLV (mSLV-L1; $CA_{6142\text{-}43}GU$) or disrupting the helical stem of SLV (mSLV-S; $CAC_{6148\text{-}50}GUG$]) reduce translational activity of the IGR IRES, presumably due to loss of interactions with ribosomal protein eS25 (and potentially uS7), however these mutations on their own largely do not affect 40S binding (Fernandez et al., 2014; Jan and Sarnow, 2002; Landry et al., 2009; Muhs et al., 2011; Nishiyama et al., 2007). We also made another mutation (mSLV-L2; $GC_{6144\text{-}45}CU$) within the SLV loop to determine effects on nucleotide identity in this region. (iv) A single point mutation in the L1.1a loop of the IGR IRES

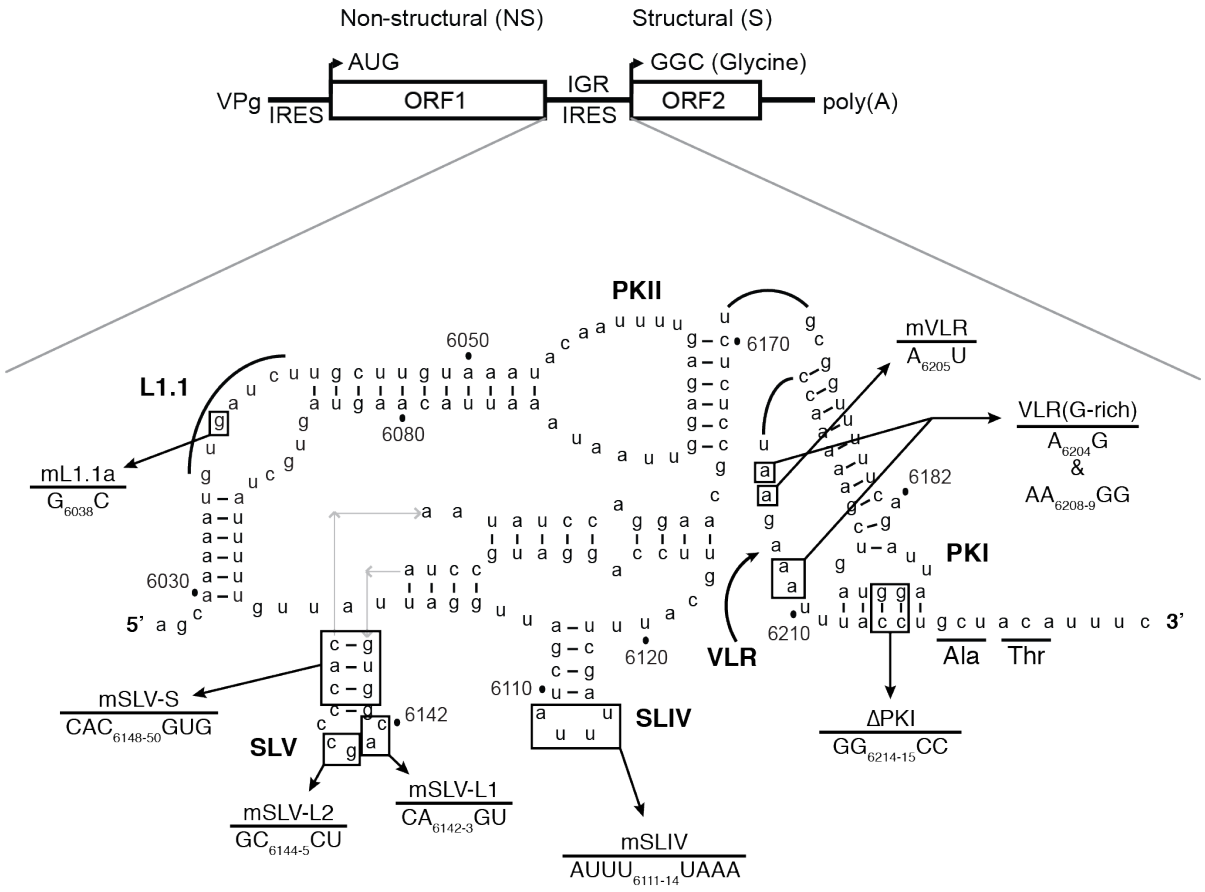
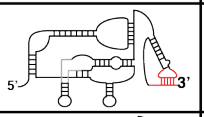
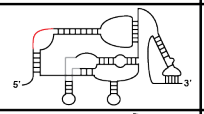
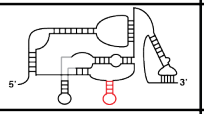
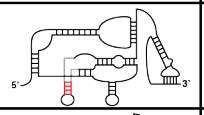
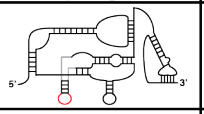
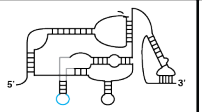
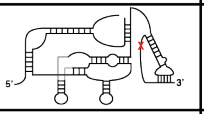
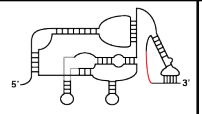
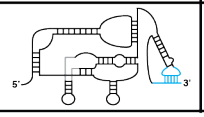


Figure 3.1 Secondary structure schematic of the CrPV Intergenic region IRES.

Pseudoknots (PK) I, II, and III as well as stem-loops (SL) IV and V are indicated. The CCU triplet initially occupies the A site of the 80S ribosome and moves to the P site; translation then initiates at the GCU codon. Numbering refers to the nucleotide position within the CrPV RNA genome. Black dashes represent helical regions within the IRES and underlined residues represent the first two amino acids of the viral capsid protein. Dashed grey lines indicate regions that interact with the ribosome. Mutations generated in this study are depicted.

Table 3.1 List of mutations generated in the CrPV IGR IRES

Mutation	Nucleotide(s)	Region	Effect on the IRES	Reference
ΔPKI	$CC_{6214\text{-}15}GG$		Disrupts 80S ribosome postioning at the GCU	Jan & Sarnow (2002)
mL1.1a	$G_{6038}C$		Interferes with 60S subunit joining	Jang *et al.* (2009)
mSLIV	$AUUU_{6111\text{-}14}UAAA$		Reduces IRES activity >90% (loss of eS7 interactions?)	Jan & Sarnow (2002)
mSLV-S	$CAC_{6148\text{-}50}GUG$		Reduces IRES activity >90% (disrupts 40S subunit interactions)	Jan & Sarnow (2002)
mSLV-L1	$CA_{6142\text{-}43}GU$		Reduces IRES activity >90% (loss of eS25 contacts?)	Jan & Sarnow (2002)
mSLV-L2	$GC_{6144\text{-}45}CU$		N/A	-
VLRm	$A_{6205}U$		Reduces IRES activity ~50%	Au & Jan (2014)
VLR (G-rich)	$A_{6204}G$ & $AA_{6208\text{-}9}GG$		Inhibits IRES activity; blocks translocation	Reuhle *et al.* (2015)
ePKI	$UA_{6190\text{-}91}GC$ & $UA_{6212\text{-}13}GC$		Enhances IRES activity 2.0-fold *in vitro*	-

Red = detrimental mutation; Blue = enhancing or no known affect mutations

($G_{6038}C$; mL1.1a). The $G_{6038}C$ mutation disrupts 60S but not 40S binding (Jang et al., 2009). The L1.1a domain interacts with the L1 stalk, resembling contacts of an E-site tRNA (Jang et al., 2009; Pfingsten et al., 2006). (v) Two sets of mutations within the Variable Loop Region (VLR), A6205 to U (mVLR) and A6205G & AA6208-9GG (VLR(G-rich)), both of which reduce IRES translation (Au et al., 2015; Ruehle et al., 2015). The mVLR inhibits ribosome positioning whereas the VLR(G-rich) mutation has been shown to inhibit IRES-mediated translocation (Au et al., 2015; Ruehle et al., 2015). (vi) Mutations that are predicted to strengthen specific helical regions of the IGR IRES ($UA_{6190-91}GC/UA_{6212-13}GC$ within PKI (ePKI)) and enhance IRES activity by ~2 fold using a bicistronic reporter assay in rabbit reticulocyte lysate (RRL) (unpublished, Chris Jang).

Most structure/function analysis on the IGR IRES was performed using bicistronic reporter constructs in orthologous systems (e.g. RRL, human and yeast ribosomes)(Fernandez et al., 2002; Jang and Jan, 2010; Thompson et al., 2001). To determine IGR IRES activity in its native context, we monitored viral protein synthesis of *in vitro* transcribed CrPV-3 RNA incubated in insect Sf-21 translation extracts. We chose to utilize Sf-21 extracts for two main reasons: 1) these extracts are from the insect order *Lepidoptera* which have be shown to be previously susceptible to CrPV infection (Bonning and Miller, 2010; Christian and Scotti, 1998) and 2) since the IGR IRES interacts with the conserved core of the ribosome, mechanisms observed between *Lepitdoteran* or *Dipteran* ribosomes are likely similar. Incubation of CrPV-3 RNA in the presence of [^{35}S]-cysteine/methionine resulted in the synthesis of viral proteins ranging from ≤15 kDa to ≥170 kDa, similar to what was observed previously (Kerr et al., 2015). The amount of translation was quantified by comparing the intensity of ORF2 protein synthesized to ORF1. Since ORF1 is translated independently from ORF2, the level of ORF1

translation can be used as a baseline for translation from the viral RNA. Placement of a stop codon downstream of the IGR IRES completely abolished structural protein synthesis (Figure 3.2)(Kerr et al., 2015). In agreement with *in vitro* studies, no structural protein synthesis was detected in mutant CrPV-3 ΔPKI, mL1.1a, mSLV-L1, mSLIV, VLR(G-rich), and a double mutant mSLIV/mSLV-L1 that contains mutations in both SLIV and SLV loops (Figure 3.2). mSLV-L2, mSLV-S and mVLR reduced protein synthesis by approximately 40-80% (Figure 3.2). In the case of mSLV-S, this is in agreement with previous *in vitro* data using compensatory mutations that the structure of SLV is integral to IRES function (Jan and Sarnow, 2002). For mVLR, this data suggests that the nucleotide identity of the VLR is important for IRES function as seen previously *in vitro* (Ruehle et al., 2015). Furthermore, specific mutations within the SLV loop (mSLV-L1 vs mSLV-L2) can affect IRES translation distinctly suggesting the nucleotide identity of the loop region may be important to IRES function. Mutant ePKI, which is reported to enhance IGR IRES activity in a bicistronic assay in RRL, resulted in an ~90% decrease in structural proteins synthesis compared to WT CrPV-3. This result demonstrates that the context of the IRES and/or the type of translation extract system can influence IRES activity. In general, mutations within the IGR IRES inhibited ORF2 structural protein synthesis in Sf-21 extracts in the CrPV-3 infectious clone.

3.3.2 IGR IRES perturbations in the CrPV infectious clone still produce viable virus

We next investigated whether these IGR IRES mutations affected viral infectivity. We transfected equal amounts of *in vitro* transcribed wild type or mutant CrPV-3 RNA into S2 cells and monitored cytopathic effects, viral protein expression and viral titres. Transfection of *in vitro* synthesized CrPV-3 RNA yielded cytopathic effects (CPE) including membrane blebbing, cell clumping, and lysis as seen previously (data not shown)(Kerr et al., 2015).

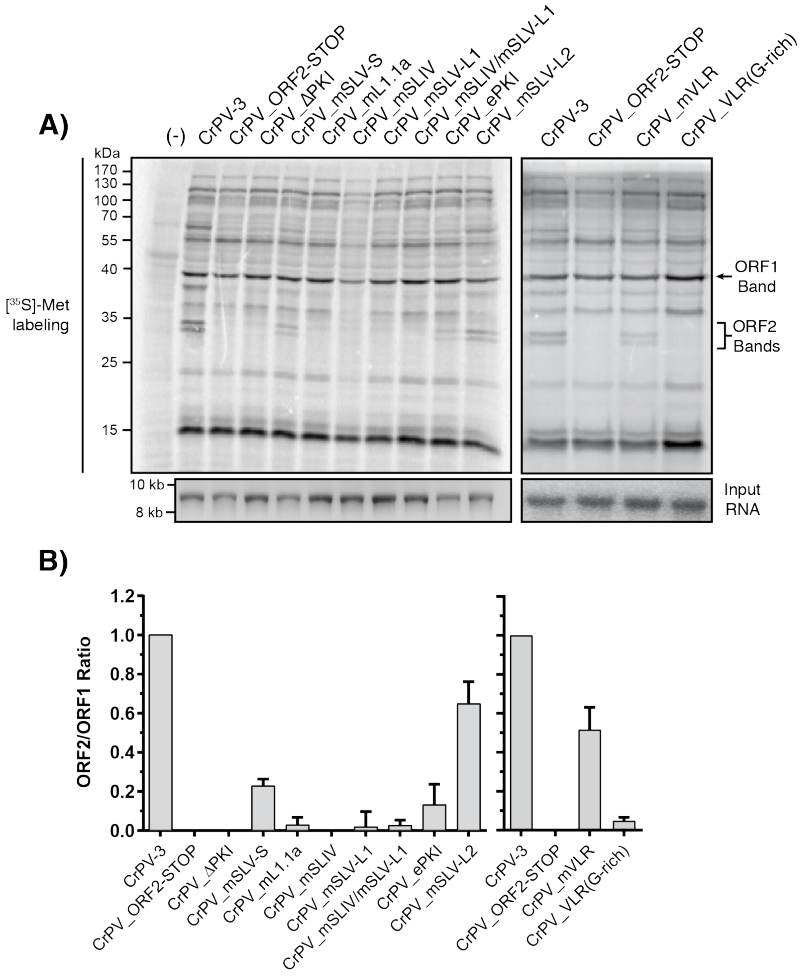


Figure 3.2 IGR IRES mutants can synthesize viral structural proteins in the context of the CrPV-3 genome. (A) *In vitro* synthesized RNA (2 µg) derived from CrPV-3 or the indicated mutant was incubated in Sf-21 extracts for 2h at 30°C in the presence of [^{35}S]-methionine/cysteine. Extracts were resolved via SDS-PAGE and radioactive proteins were analyzed by phosphoimager analysis. A mutant containing a STOP codon in ORF2 of CrPV-2 (lane 3) was used as a control as it does not express structural proteins (see Materials and Methods)(Kerr et al., 2015). Shown is a representative gel from 3 independent experiments. 3 different batches of *in vitro* transcribed RNA were used. *bottom:* Agarose gel of *in vitro* transcribed input RNA. A no RNA control is indicated by (-). Arrows indicate quantified bands. **(B)** Quantification of structural (ORF2) versus non-structural (ORF1) protein synthesis in Sf-21 extracts. Gels were quantified using ImageQuant. Bands used for quantitation are indicated with arrows in the gel above. Error bars represent the standard deviation.

Moreover, RdRP (ORF1) and VP2 (ORF2) proteins were readily detected by immunoblotting and virus production was observed 48 hours post transfection (Figure 3.3). As shown previously, introduction of a stop codon downstream of the IGR IRES did not lead to viral protein synthesis or virus production (ORF2-STOP)(Kerr et al., 2015). Similarly, no viral proteins or viral titre were detected in S2 cells transfected with mutant CrPV-3 ΔPKI, mL1.1a, VLR(G-rich), or double mutant mSLIV/mSLV-L (Figure 3.3). However, transfection of mutant CrPV-3 mSLV-S, mSLV-L, mSLIV, ePKI, eSLV and mVLR resulted in viral protein expression and yielded virus production to varying extents. Compared to wild-type CrPV-3, viral yield was lower for cells transfected with mutant CrPV-3 mSLV-S, mSLV-L, ePKI and mVLR (Figure 3.3). In the case of ePKI, it is possible that increasing the rigidity of PKI by strengthening the base-pairing is not favorable in this context as the IGR IRES may require innate flexibility in the tRNA-mimicry region to facilitate IRES translocation. For some mutations such as ΔPKI, mL1.1a, and VLR(G-rich), the loss of IRES translation activity correlated with the lack of virus production, thus demonstrating the importance of these domains in directing IRES translation during virus infection. Surprisingly, several mutations such as mSLIV, mSLV-S, and mSLV-L1 yielded robust virus production yet were IRES-translation compromised. For instance, while the $AUUU_{6111\text{-}14}$ loop sequence of SLIV is absolutely conserved throughout dicistroviruses, strongly suggesting nucleotide context importance, mutating AUUU to UAAA still yielded similar viral titres as the wild-type CrPV-3. By contrast, mutant mVLR, which decreased translation in vitro by only 50% (Figure 3.2), resulted in a 2-log decrease in viral yield (Figure 3.3). This corroborates the importance of this region of the IRES in viral structural protein synthesis during infection.

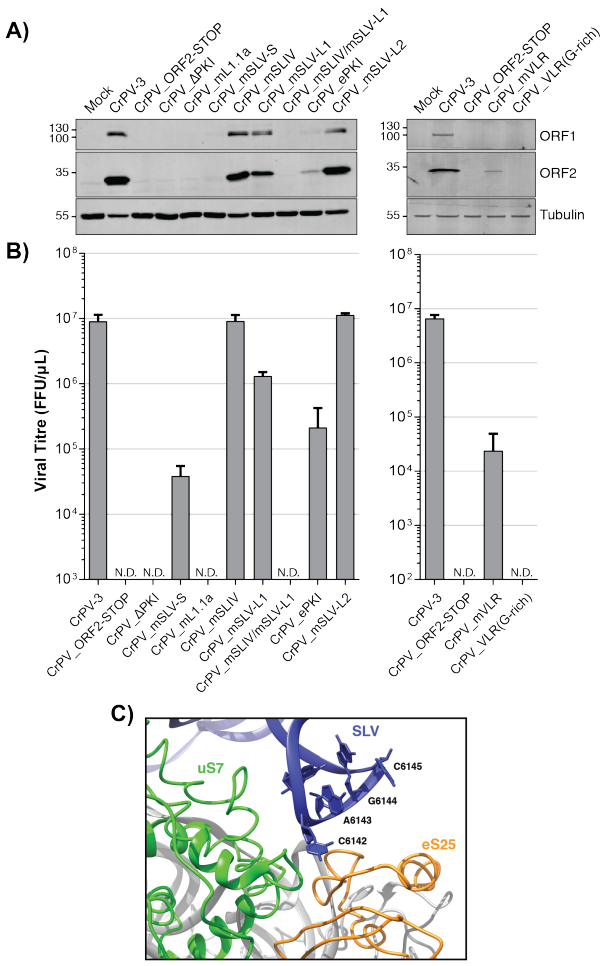


Figure 3.3 Viruses harboring mutations in the IGR IRES produce viable virus.

(A) Western blot analysis of viral 3CD (ORF1) and VP2 (ORF2) from lysates of cells transfected with the indicated CrPV genomic RNA at 48 h.p.t. Note that uncropped images are shown in Supplementary Figure S2 **(B)** Titres of transfected CrPV IGR IRES mutants. 2.5 x 10^6 S2 cells were transfected with 3 µg of the indicated *in vitro* transcribed genomic RNA. Titres were measured from the cell pellet as described in the Materials and Methods at 48 h.p.t. Shown are averages from at least 3 independent experiments (±S.D.). N.D. = Not detected. **(C)** Zoomed in segment from a cryo-EM structure of the 80S ribosome from *Kluveromyces lactis* bound to the CrPV IGR IRES (PDB ID 4V92) (Republished from Cell[14] licensed under CC by 3.0). SLV of the IRES is coloured in blue while ribosomal protein eS25 and uS7 are coloured in orange and green, respectively.

3.3.3 Viable viruses with mutated IGR IRESs are impaired in viral growth at the translation level

Since mutant CrPV-3 mSLV-S, mSLV-L1 and mSLIV resulted in a productive infection (Figure 3.3), we next investigated whether the observed reduction in viral loads was due to diminished viral translational activity or replication. To address this, we infected S2 cells with wild type CrPV-3 or mutant virus and monitored viral protein synthesis via [^{35}S]-met/cys metabolic pulse-labeling and viral RNA by Northern blot analysis (Figure 3.4). There was no observable difference in accumulation of viral RNA between wild type or any of the mutant viruses suggesting that RNA replication is unaffected (Figure 3.4). However, as early as 4 hours post-infection there is a noticeable reduction in viral structural protein synthesis when a mutation is present either in the loop (mSLV-L1) or stem (mSLV-S) of SLV with the latter being more detrimental, while no reduction is observed with a SLIV mutation (Figure 3.4). These data corroborated viral titres observed after RNA transfection (Figure 3.3B) and altogether indicate that IGR IRES-dependent translation is an essential step in promoting virus infection.

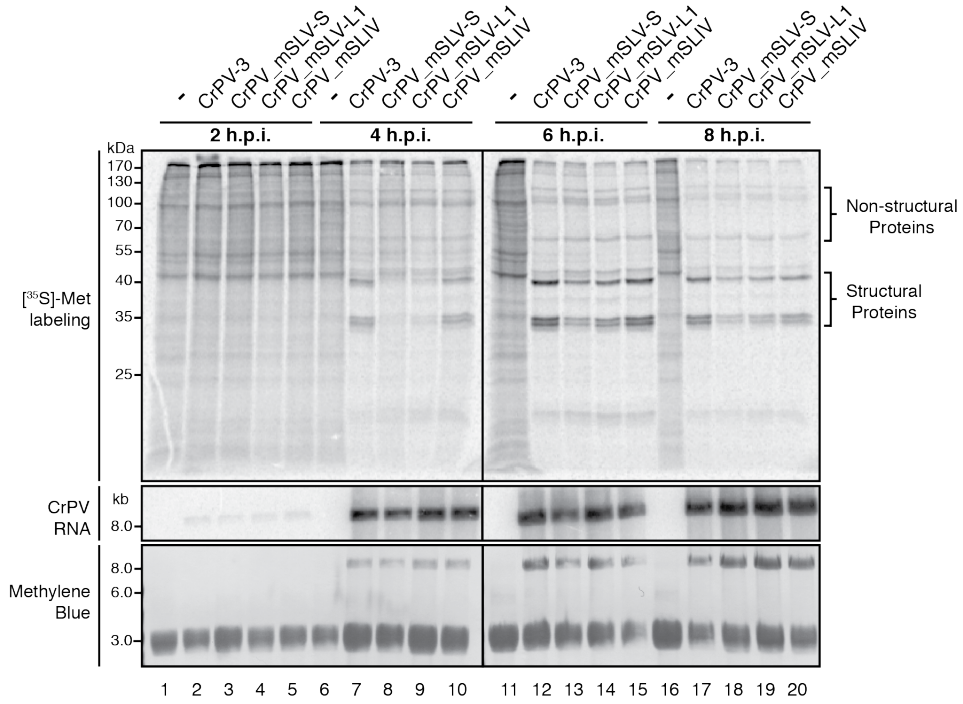


Figure 3.4 Viable viruses with mutated IGR IRESs are impaired in viral growth at the translational level.

S2 cells were mock-infected (-) or infected with CrPV-3 or the indicated mutant (MOI 10). Cells were metabolically labeled with [^{35}S]-methioine/cysteine for the last 30 minutes of infection. Lysates were subjected to SDS-PAGE and Northern blot analysis. Accumulation of viral RNA was monitored by probing for the CrPV RNA genome. Shown are representative gels from three independent experiments.

3.4 DISCUSSION

Structural and biochemical studies on the dicistrovirus IGR IRES have yielded enormous insights into the detailed mechanism of IRES-mediated translation, ribosome dynamics, and decoding. However, all of these studies are *in vitro* using reconstituted systems and translation extracts and thus, the context of IGR IRES translation in a virus system was lacking. Here, we now complement these studies by providing a physiological context of IGR IRES translation in a virus infection system, thus emphasizing the importance of specific structural elements of the IRES in its native context. It is possible that *in vitro* assays are not sensitive enough or only capture specific properties of these elements on IRES translation and that the full potential are only observed under more physiological conditions. For example, these RNA elements may sample more structural conformations under virus infection that allow viral protein synthesis to occur. Alternatively, there may be functional redundancy in these elements. Although structural studies indicate that SLIV and SLV interact with distinct regions of the 40S subunit, uS7 and eS25, biochemical data suggest that eS25 makes contacts with both loop regions of SLV and SLIV, which may explain how mutations in either stem-loop alone are not sufficient to abolish viral translation in infected cells, yet when combined can prevent viral protein synthesis and virus production (Figure 3.3B; see mSLIV/mSLV-L)(Muhs et al., 2011; Nishiyama et al., 2007). In the case of ePKI, it is possible that increasing the rigidity of PKI by strengthening the base-pairing is not favorable in this context as the IGR IRES may require innate flexibility in the tRNA-mimicry region to facilitate IRES translocation.

For some mutations such as ΔPKI, mL1.1a, and VLR(G-rich), the loss of IRES translation activity correlated with the lack of virus production, thus demonstrating the importance of these domains in directing IRES translation during virus infection. Surprisingly,

several mutations such as mSLIV, mSLV-S, and mSLV-L1 yielded robust virus production yet were IRES-translation compromised. For instance, while the $AUUU_{6111-14}$ loop sequence of SLIV is absolutely conserved throughout dicistroviruses, strongly suggesting nucleotide context importance, mutating AUUU to UAAA still yielded similar viral titres as the wild-type CrPV-3. It is possible that it is indeed the structure and not the sequence that is necessary for SLIV function.

Interestingly, both SLV loops mutants, SLV-L1 and SLV-L2, yielded different results. Mutating nucleotides $CA_{6142-43}$ decreased overall viral yield and ORF2 production while mutating nucleotides $GC_{6144-45}$ did not have a significant affect despite being within the same vicinity (Figure 3.3; compare mSLV-L1 and mSLV-L2). Cryo-EM structural data of the CrPV IGR IRES bound to the 80S ribosome of *Kluyveromyces lactis*, indicate that nucleotide C_{6142} of SLV is in a flipped out orientation that is in close proximity to ribosomal protein uS7, while nucleotides $GC_{6144-45}$ face inward into the loop (Figure 3.3C)(Fernandez et al., 2014; Murray et al., 2016). The position and conformation of C_{6142} when bound to the ribosome could explain the discrepancies observed between a detrimental mutation (mSLV-L1) and a mutation with no apparent affect (mSLV-L2). Alternatively, mutating C_{6142} to G may potentially lead to an additional base pair with C6146 in SLV, resulting in a lengthening of the helical stem and a reduction of the loop to 3 nucleotides, which may be insufficient for eS25 binding. At any rate, the nucleotide identity of SLV appears to be crucial for IRES function.

Finally, mutations within the VLR have been shown to affect ribosome positioning and translocation on the IRES (Au and Jan, 2012; Ruehle et al., 2015). Recent cryo-EM studies reveal that the dynamic and flexible VLR becomes ordered upon the PKI translocation from the A to P site, resembling a post-translocated ribosome (Abeyrathne et al., 2016; Murray et al.,

2016). At this state, the VLR interacts with R148 and R157 in β-hairpin of uS7, which resembles an E-site tRNA anticodon stem loop interaction (Abeyrathne et al., 2016). We show that mutations in VLR that likely disrupt these key ribosome-IRES interactions had a significant inhibitory effect on virus production (Figure 3.3), thus further supporting the importance of this domain in directing viral structural protein synthesis during infection.

Overall, the use of the CrPV infectious clone provides a powerful biological framework for pinpointing the relevant IGR IRES mechanistic details and this study underpins the importance of complementing *in vitro* based approaches with those in a more physiological virus system.

Chapter 4: IGR IRES-mediated translation of a +1 overlapping reading frame enhances CrPV infection in flies

4.1 INTRODUCTION

The ribosome mediates translation involving decoding the open reading frame codon by codon through delivery of the correct aminoacyl-tRNAs to the ribosomal A site. This fundamental process occurs with high fidelity for proper gene expression in all species. However, mechanisms exist that can alter the translational reading frame, thus producing alternative protein products from a single RNA (Baranov et al., 2015). In general, these mechanisms termed recoding, involve a specific RNA structure or element that interacts with the ribosome to cause the translating ribosome to shift reading frame by -1/+1/+2 allowing it to read through stop codons or bypass sequences and restart translation downstream (Dinman, 2012; Firth and Brierley, 2012; Jan et al., 2016). Study of these mechanisms has been enlightening; revealing key ribosome:RNA interactions that alter fundamental processes in the mechanics of ribosome decoding and reading frame maintenance. Importantly, recoding mechanisms are now appreciated as important regulatory processes that can impact the fate of protein expression in cells and virus infections (Au and Jan, 2014). Unlike these recoding mechanisms that involve an actively translating ribosome, the intergenic internal ribosome entry site (IRES) within a subset of dicistroviruses has the unusual property to directly recruit the ribosome and initiate translation from overlapping 0 and +1-frame codons to produce two distinct proteins (Ren et al., 2012). Here, we report a novel recoding mechanism and translational initiation pathway whereby a related dicistrovirus IRES directs the ribosome to initiate translation downstream.

In general, the dicistrovirus IGR IRESs are conserved at the structural, but not sequence level and are classified into two sub-groups (termed Type I and II) based on the presence of

distinct structural elements; the main distinction comes from a larger L1.1 loop and an additional stem-loop (SLIII) in Type II IRESs (Jan, 2006; Nakashima and Uchiumi, 2009). SLIII allows for the PKI domain of Type II IRESs to mimic the global shape of a tRNA in addition to assisting in reading frame selection and the larger L1.1 region functions to mediate 60S recruitment (Au et al., 2015; Jang et al., 2009; Pfingsten et al., 2006). The domains of Type I and II IGR IRESs function similarly to directly recruit 80S ribosomes and initiate translation (Cevallos and Sarnow, 2005; Hatakeyama et al., 2004; Hertz and Thompson, 2011a; Jang and Jan, 2010). Recent high resolution cryo-EM structures of the IGR IRES bound to the 80S ribosome have demonstrated that the IRES initially binds in the A site (Abeyrathne et al., 2016; Fernandez et al., 2014; Muhs et al., 2015). Movement of the IRES involves an eEF2-dependent pseudo-translocation event where the ribosome rotates up to 10° allowing PKI to move into the P site in an inchworm-like manner (Abeyrathne et al., 2016). This allows for the non-AUG initiation codon of the IRES to be presented in the A site for the incoming amino-acyl tRNA. The first pseudo-translocation event and delivery of the first amino-acyl tRNA are the rate-limiting steps of initiation on the IGR IRES (Zhang et al., 2016).

Biochemical, phylogenetic, and bioinformatics analyses have demonstrated that a subset of Type II IGR IRESs are able to direct synthesis of a hidden +1 open reading frame (ORF), termed ORFx, within ORF2 of the viral genome (Firth et al., 2009; Ren et al., 2012). The functional role of ORFx during viral infection remains elusive. Extensive mutagenesis of the PKI domain of the *Israeli acute paralysis virus* (IAPV) IGR IRES has revealed that 0 and +1 frame translation can be uncoupled, suggesting that the IGR IRES may adopt specific conformations that govern the translational reading frame (Au et al., 2015; Ren et al., 2014). Generally, Type I and II IGR IRESs are thought to operate similarly in mechanism. Specific domains between the

two types are functionally interchangeable (Jang and Jan, 2010). In the present study, we investigate the capacity of other IGR IRESs from Dicistroviruses to facilitate +1-frame translation. We show that the IGR IRES from Cricket paralysis virus (CrPV) can synthesize a 41 amino acid long ORFx protein using an unexpected mechanism that involves IRES-mediated ribosome bypassing. Furthermore, we provide insight into the role of ORFx during CrPV infection and show that mutants deficient in ORFx have impaired virulence in adult flies, thus uncovering a novel viral recoding strategy that is essential for viral pathogenesis.

4.2 MATERIALS & METHODS

4.2.1 Cell culture and virus.

Drosophila Schneider line 2 (S2; Invitrogen) cells were maintained and passaged in Shield's and Sang medium (Sigma) supplemented with 10% fetal bovine serum.

Propagation of CrPV in *Drosophila* S2 cells has been previously described (Garrey et al., 2010). CrPV-2 and mutant viruses were generated from *Drosophila* S2 cells using an adapted protocol (Krishna et al., 2003). Briefly, 5.0 x 10^7 S2 cells were transfected with *in vitro* transcribed RNA derived from pCrPV-3 or mutant plasmids and incubated for 48h. Cells were dislodged into the media, treated with 0.5% Igepal CA-630 (Nonidet P-40) and 0.1% 2-mercaptoethanol, and incubated on ice for 10 min. Cell debris was cleared by centrifugation at 13,800 RCF for 15 min at 4°C. Viral particles were then concentrated by ultracentrifugation at 141,000 RCF for 2.5 h at 4°C. The pellet was resuspended in PBS and sterilized through a 0.2 μM filter. Viral titres and yield were determined as previously described (Kerr et al., 2016). All viruses were sequence verified via RT-PCR with primers directed against the CrPV IGR IRES.

4.2.2 Plasmids and bicistronic reporter constructs.

Each IGR IRES with flanking upstream and downstream sequences was cloned between the EcoRI and the NcoI sites within the intergenic region of pRΔDEF as described (Johannes et al., 1999). The Firefly luciferase (FLuc) open reading frame was fused in frame to either the 0 or +1 frame. The sequences that were cloned are as follows: nucleotides 6372–6908 of Israeli acute paralysis virus (IAPV; NC_009025.1); nucleotides 6741-6969 of Taura syndrome virus (TSV; NC_003005.1); nucleotides 6860-7243 of Mud crab dicistrovirus (MCDV; HM777507.1); nucleotides 5974–6372 of Cricket paralysis virus (CrPV; NC_003924.1); nucleotides 6060-6422 of Drosophila C Virus (DCV; NC_001834); nucleotides 5982-6324 of Plautia stali intestine virus (PSIV; NC_003779.1); nucleotides 151-771 of Big Sloux River virus (BSRV; JF423197.1); nucleotides 5626-5917 of Black queen cell virus (BQCV; NC_003784.1); nucleotides 3453-3712 of Aphid lethal paralysis virus (ALPV) strain brookings non-structural polyprotein and capsid protein precursor (HQ871932.1); nucleotides 6546-7159 of Rhopalosiphum padi virus (RhPV; NC_001874.1). The PSIV_CrPV chimera contains the nucleotides 5981-6192 of PSIV and 6217-6371 of Cricket paralysis virus isolate CrPV-2 (KP974706.1)(Kerr et al., 2015).

For CrPV IRES-containing bicistronic constructs, the AUG start codon of the FLuc gene in the +1 frame was removed by PCR-based site-directed mutagenesis (Stratagene). For the T2A-containing constructs, the Thosea asigna virus (accession: AF062037) 2A sequence was inserted in frame preceding the FLuc gene into plasmid pEJ566 between EcoRI and NdeI restriction sites as described previously (Wang et al., 2013).

Hairpin structure sequences HP1 and HP2 were inserted between the 6th codon AAC and 7th codon AAG of CrPV (+1) frame. The HP1 sequence is 5'-CAGCGTTATCG GGAACGTCGAAGGGGTACGCCATTGTTCGACCACCGGCTCGAGGGGCCC-3' (ΔG = -

17.0 kcal/mole). The HP2 sequence is 5'- ACTGCAGCCACAACCGGCCCCAAGCTT GGGCCGTGGTGGCTGCAG-3' (ΔG = -32.70 kcal/mole).

Other dicistrovirus sequences used for analysis in in Table S1 are: Solenopsis invicta virus-1 (SINV-1) NC_006559; Kashmir bee virus (KBV) NC_004807; Acute bee paralysis virus (ABPV) NC_002548; Homalodisca coagulata virus-1 (HoCV-1) NC_008029; Himetobi P virus (HiPV) NC_003782; Cripavirus NB-1/2011/HUN (Bat guano dicistrovirus) NC_025219.1; Aphid lethal paralysis virus (ALPV) NC_004365; Aphid lethal paralysis virus strain brookings (ALPV-brookings) HQ871932; and Triatoma virus (TrV) NC_003783.

The constructs encoding HA-tagged CrPV ORFx were generated as follows. For the C-terminally tagged ORFx-HA, a 3X HA tag was inserted between XbaI and ApaI sites in pAc5.1/V5-HisB (Thermo). CrPV ORFx (nucleotides 6254-6376) was subsequently cloned in-frame using KpnI and XbaI sites. For the N-terminally tagged HA-ORFx a synthesized fragment of DNA (gBlock; IDT) containing a 3X HA tag followed by the CrPV Orfx sequence was cloned into the pAc5.1 vector with KpnI and ApaI. All constructs were verified via sequencing.

4.2.3 *In vitro* transcription and RNA transfection.

pCrPV-2 (Kerr et al., 2015) and derivative plasmids were linearized with Ecl136II. RNA was transcribed using a T7 RNA polymerase reaction and subsequently purified with a RNeasy kit (Qiagen). The integrity and purity of the RNA were confirmed on a 1.2% denaturing formaldehyde agarose gel.

Transfection of *in vitro* synthesized RNA into S2 cells was performed using Lipofectamine 2000 (Invitrogen) as per the manufacturer's instructions. 3 µg of RNA derived from either pCrPV-2 or its cognate mutants were used for transfection using 2.5 x 10^6 cells.

4.2.4 *In vitro* translation assays.

Plasmids were linearized with BamHI and then purified using QiaQuick Purification columns (Qiagen). The Sf-21 TnT Coupled Transcription/Translation System was utilized for all translation assays (Promega). In this system, linearized plasmid is added to the extracts and RNA is synthesized by T7 RNA polymerase and subsequently translated. Total reaction volume was 10 µL containing: 6.7 µL of Sf-21 cell extract (Promega), 0.3 µL of L-[^{35}S]- methionine/cysteine (PerkinElmer, >1000 Ci/mmol) and 1 µg linearized plasmid DNA. Each reaction mixture was incubated at 30°C for 2 hours and then resolved on a 15% SDS-PAGE. Gels were dried and radioactive bands were monitored via phosphoimager analysis. For quantitation of bands, the number of methionines and cysteines were accounted for and normalized for each protein. For T2A containing constructs (Wang et al., 2013), luciferase activity was monitored using a Dual-Luciferase reporter assay system (Promega) and an Infinite M1000 PRO microplate reader (Tecan).

4.2.5 Northern blot analysis.

Total RNA was isolated from cells using TRIzol reagent. Equal amounts of RNA were loaded on a denaturing agarose gel and subsequently transferred to Zeta-probe blotting membrane (Bio-Rad). DNA probes were radiolabelled with α[P^{32}]-ATP (MP Biomedicals) using a DecaLabel DNA labeling kit (Fermentas). Probes were incubated with membrane in hybridizatn buffer (7% SDS, 0.5 M Sodium phosphate, 1 mM EDTA) at 65°C overnight. Membranes were rinsed with H_2O and radioactive bands were detected via phosphoimager analysis (Storm; GE Healthcare). The DNA probe to detect CrPV RNA was generated by PCR amplification using primers 5'-TCCTCAAGCCATGTGTATAGGA-3' and 5'-GTGGCTGAAATACTATCTCTGG-3'.

4.2.6 Western blots.

Equal amounts of S2 protein lysates were resolved on a 12% SDS-PAGE gel and then transferred to a polyvinylidene difluoride Immobilon-FL membrane (Millipore). Membranes were blocked for 30 min at room temperature with 5% skim milk in TBST (50 mM Tris, 150 mM NaCl, 1% Tween-20, pH 7.4). Blots were incubated for 1 h at room temperature with the following antibodies: CrPV ORF1 (raised against CrPV RdRp) rabbit polyclonal (1:10,000), CrPV ORF2 (raised against CrPV VP2) rabbit polyclonal (1:10,000) (Garrey et al., 2010), mouse anti-tubulin (1:1000; Santa Cruz), mouse anti-KDEL (1:1000;), or mouse anti-cytochrome C (1:1000; AbCam). Membranes were washed 3 times with TBST and incubated with either goat anti-rabbit IgG-HRP (1:20,000; GE Healthcare), goat anti-mouse IgG-HRP (1:5000; Santa Cruz), IRDye 800CW goat anti-mouse (1:10,000; Li-Cor Biosciences), or IRDye 800CW goat anti-rabbit (1:20,000; Li-Cor Biosciences) for 1 h at room temperature. An Odyssey imager (Li-Cor Biosciences) or enhanced chemiluminescence (Thermo) were used for detection.

4.2.7 LC-MS/MS Analysis.

Cell pellets harvested from CrPV-infected cells at 6 hpi were solubilized in 1% sodium deoxy cholate and 50 mM NH_4HCO_3. Protein concentrations were determined via BCA assay (Thermo). Proteins (100 µg) were reduced (2 µg DTT, 37°C, 30 min) and alkylated (5 µg iodoacetamide, RT, 20 min). Samples were digested with trypsin overnight at room temperature. Peptides were acidified with 1% TFA to pH <2.5 and the precipitated deoxycholate was removed via centrifugation. Peptides were desalted and concentrated on C18 STAGE-tips, eluted in 80% acetonitrile, 0.5% acetic acid, and dried in a vacuum concentrator (Eppendorf)(Rappsilber, Ishihama, & Mann, 2003). Samples were resuspended in 20% acetonitrile and 0.1% formic acid before loading on an Agilent 6550 mass spectrometer.

Data were searched using MaxQuant (v1.5.3.30)(Cox and Mann, 2008). Parameters included: carbamidomethylated cysteine (fixed), methionine oxidation (variable), glutamine and asparagine deamidation (variable), and protein N-terminal acetylation (variable); trypsin specific; maximum 2 missed cleavages; 10 ppm precursor mass tolerance; 0.05 Da fragment mass tolerance; 1% FDR; +1 to +7 charge states; and common contaminants were included. Both the *Drosophila* and CrPV protein databases used were the most recent annotations downloaded from UniProt (www.uniprot.org).

4.2.8 Subcellular fractionation by differential centrifugation.

3 X 10^7 S2 cells were transfected with either Drosophila expression vectors (pAc5.1) containing ORFx-HA or HA-ORFx (pORFx-HA or pHA-ORFx) for 48 h at 25°C (XtremeGene; Roche). Cells were then collected by centrifugation at 200 RCF for 5 min and washed twice with ice cold PBS. Cell pellets were then resuspended in SF Buffer (20 mM HEPES [pH 7.4], 10 mM KCl, 1.5 mM $MgCl_2$, 1 mM EDTA, 1 mM EGTA) and incubated on ice for 10 min. Cells were subsequently disrupted by passaging through a 25 gauge needle 20 times. Whole cells were removed by centrifugation at 200 RCF for 5 min and nuclei were then collected by centrifugation at 600 RCF for 10 min. The mitochondria and endoplasmic reticulum (ER) were pelleted by centrifugation at 3000 RCF for 15 min. Finally, remaining membrane was pelleted at 16,000 RCF for 30 min. The remaining supernatant was collected as the cytosolic fraction.

4.2.9 Immunofluorescence.

S2 cells (1.5 x 10^6) were transfected for 48 h with 2 µg of pAc5.1/V5-HisB plasmid constructs expressing N- or C-terminally HA-tagged ORFx. Cells were seeded onto coverslips and fixed with 3% paraformaldehyde for 15 minutes. Subsequently, cells were permeabilized with 0.2% Triton X-100 in 1X PBS for 1 h before blocking coverslips with 3% BSA in 1X PBS

for 1 h. Coverslips were then incubated with rabbit anti-HA (1:1000; Cell Signalling) and either mouse anti-Lamin (1:1000; DSHB), mouse anti-Golgin84 (1:1000, DSHB), or mouse anti-Calnexin 99A (1:1000, DSHB) in blocking solution overnight at 4°C. Cells were then washed three times with 1X PBS and subsequently incubated with either goat anti-rabbit Texas Red (1:5000) or goat anti-mouse AlexaFluor 488 (1:5000)(Invitrogen). After washing twice with 1X PBS, cells were stained with Hoescht dye (1:10,000 in 1X PBS) for 10 min. Finally, slides were mounted and analyzed using a Leica SP5 confocal microscope with a 63x oil objective lens and a 2x digital zoom. Z-stacks of 15 slices each were taken of for each condition.

4.2.10 Fly stocks and viral injections.

Flies (*Isogenic* w^{1118}; Bloomington Drosophila Stock Center) were maintained on standard cornmeal food at 25°C and 70% humidity with a 12 h light-dark cycle. Freshly eclosed virgin males and females were separated and collected in groups of 10 each. Flies (10 males and 10 females) were injected with 200 nL of PBS, CrPV-2, CrPV-S12, or CrPV-S19 (5000 FFU) using a PV830 PicoPump (World Precision Instruments) and transferred to standard food. Mortality was monitored daily.

4.3 RESULTS

4.3.1 The CrPV IGR IRES-dependent supports +1-frame translation, is IGR IRES-dependent, and initiates downstream

It was previously demonstrated that a subset of dicistrovirus IGR IRESs can direct translation in the 0- and +1-frames and that a base pair adjacent to the PKI domain is important for initiation (Ren et al., 2014; Ren et al., 2012). The IGR IRESs are classified into two types: Type I, and II, with the main difference being an extra SLIII within the PKI domain of Type II

IRES. Since the honeybee and fire ant viruses harbour Type II IGR IRESs that can support +1 frame ORFx translation, we investigated whether Type I IRESs also had the capacity for +1-frame translation. This was indeed found to be the case for a subset of Type I IGR IRESs. Particularly, the CrPV IGR IRES could support +1-frame translation at 5% of the levels of 0-frame ORF2 translation (see Appendix A; Figure A.1).

To investigate the mechanism of CrPV +1-frame translation, an additional bicistronic reporter construct was created that contains the CrPV IRES where Renilla luciferase (36 kDa; RLuc) monitors scanning-dependent translation and a firefly luciferase gene was subcloned into the +1-frame downstream of the IRES with CrPV nucleotides 6217-6387 (Accession KP974707.1), which includes the predicted +1-frame CrPV ORFx (65 kDa; ORFx-FLuc). A shortened IRES-mediated 0-frame protein (~11 kDa; sORF2) is also translated, thus scanning-dependent and CrPV IRES-mediated 0/+1-frame translation can be simultaneously monitored from this dual luciferase reporter construct (Figure 4.1A; see Appendix A - Figure A.2). Synthesis of all three proteins is detected by incubating the bicistronic construct in a Sf21 translation extract in the presence of [^{35}S]-methionine/cysteine (Figure 4.1B, lane 1) (Ren et al., 2012). Additionally, bicistronic reporter constructs that contain a T2A 'stop-go' peptide were generated, which allow sensitive and accurate measurement of luciferase activity (Wang et al., 2013) (see Appendix A; Figure A.2).

We first determined whether the structural integrity of the CrPV PKI domain is important for +1-frame translation. As observed previously, the CrPV IRES +1-frame ORFx translation is approximately 5% of 0-frame translation. Mutating CC_{6214-5} to GG, which disrupts PKI base pairing and abolishes CrPV IRES activity, resulted in negligible or diminished 0- and +1 frame translation whereas a compensatory mutation that restores PKI base pairing rescued translation

(Jan and Sarnow, 2002) (Figures 4.1B), indicating that the integrity of the IRES and the PKI domain is required for CrPV IRES +1-frame translation.

The adjacent U-G base pairing of the IAPV IRES is important for +1-frame translation (Ren et al., 2012). CrPV also has the capacity to form a wobble base pair adjacent to the PKI domain through nucleotides U_{6186} and G_{6217}, potentially directing translation from the first +1-frame CUA leucine codon. To determine if this base pair is necessary to drive CrPV IRES +1-frame translation, we mutated U_{6186} and G_{6217} to other bases. Mutating U_{6186} to C or G led to an approximate 18%-23% reduction in +1 frame activity and mutating G_{6217} to C, A, or U resulted in roughly a 39%, 24%, and 36% reduction in +1 frame activity, respectively. Although each mutation reduced +1 frame translational activity to some degree, none of the mutations abolished it (Figures 4.1B, lanes 4-9). These results suggest that unlike with IAPV, base pairing between nucleotides 6186 and 6217 is not absolutely required for CrPV IRES +1-frame translation.

To determine the initiation site of CrPV ORFx, we systematically replaced codons downstream of the IRES with a stop codon and monitored 0 and +1-frame translation *in vitro* using the bicistronic reporter construct (Figure 4.1B, C). Overall, stop codons placed in the +1-frame did not significantly affect 0-frame translation, indicating that IRES activity was not compromised (Figure 4.1B, C). Replacing individual codons between the 1st and the 12th +1-frame codon with a stop codon inhibited to varying extents (between 36%-71% reduction compared to wild-type) but did not completely abolish +1 frame translation (Figure 4.1B, lanes 11-16; 4.1C). Conversely, +1-frame translation was completely inhibited when the 13th +1-frame codon and codons thereafter was replaced with a stop codon (Figure 4.1B, lanes 15-16; 4.1C). Replacing both the 3rd and 5th +1-frame codons with stop codons reduced +1-frame translation by 33% but did not eliminate it, ruling out ribosome read-through. It is possible that IRES translation initiates

in the 0-frame and then a fraction of translating ribosomes shift into the +1-frame. To address this, we inserted a stop codon in the 0-frame downstream of the IRES. As expected, a stop codon in the 0-frame 1st to 4th codons downstream of the IRES abolished 0-frame translation (Figure 4.1B, lanes 18-21). However, the 0-frame stop codon insertions reduced by 52-77% but did not eliminate +1 frame translation indicating that a subset of ribosomes do not shift from the 0 to +1 reading frame. Further supporting this, introducing stop codons in both the 0 and +1 frames did not abolish +1 frame translation, though inhibited +1 frame translation by ~80% (Figure 4.1C, 0 S4/+1 S5). We noted that the adjacent 14th codon is an AUG methionine. Mutating the AUG to a CGG (M14R) decreased by 63% but did not abolish +1-frame translation (Figure 4.1C). In summary, the mutational analysis indicated that CrPV ORFx translation requires an intact IRES to initiate translation from the +1-frame AAA lysine 13th codon. Mutating the 13th AAA codon to UUU slightly reduced but did not abolish +1-frame translation (Figure 4.1D), suggesting that there may be flexibility in the start codon identity for CrPV ORFx +1-frame translation initiation.

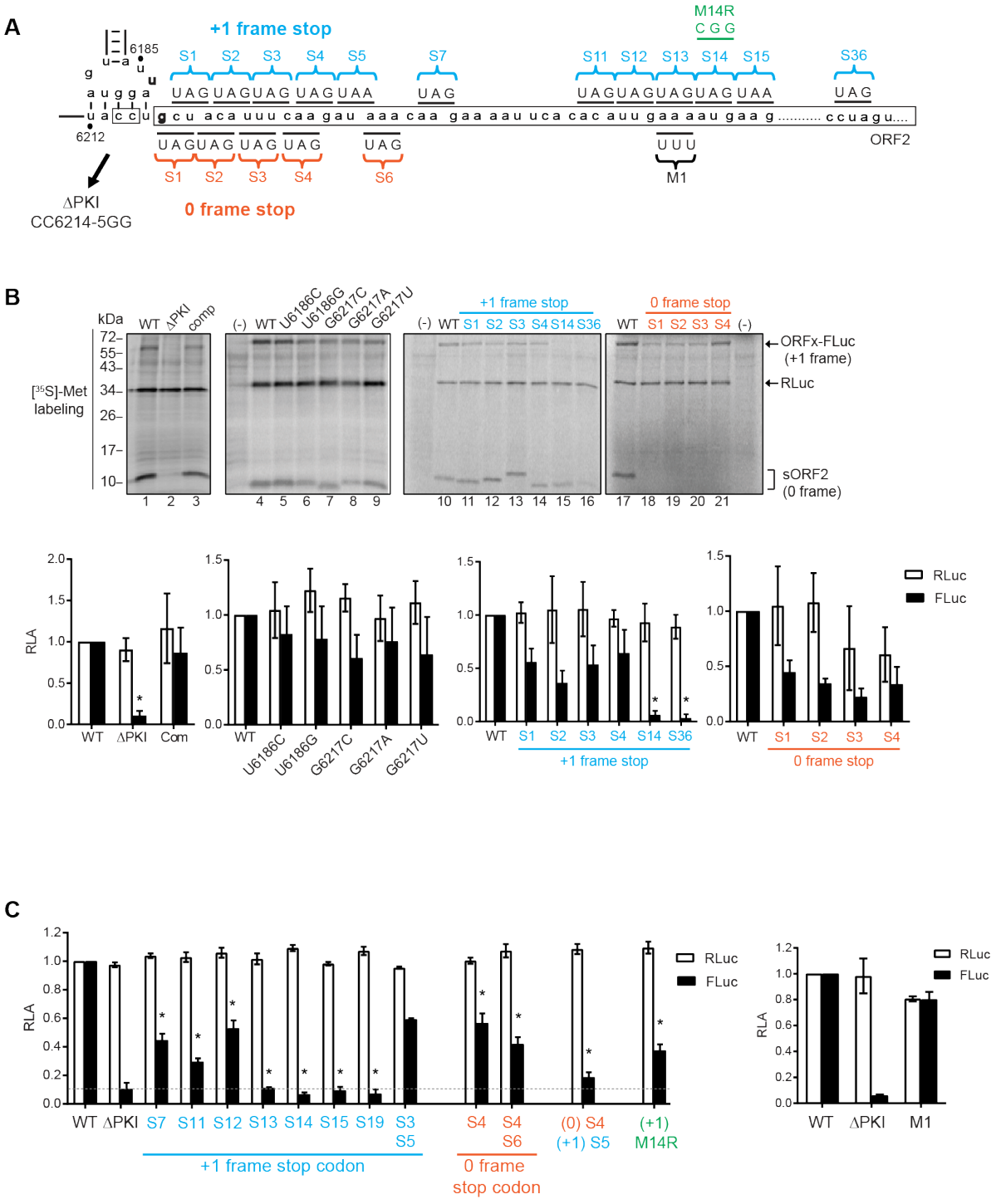

A
+1 frame stop
M14R
C G G
6185
S1 S2 S3 S4 S5 S7 S11 S12 S13 S14 S15 S36
U A G U A G U A G U A G U A A U A G U A G U A G U A G U A A U A G
g c u aca uuu caa gau aaa caa gaa aau uca cac auu gaa aau gaa g ..... ccuagu....
6212
ORF2
U A G U A G U A G U A G U A G U U U
S1 S2 S3 S4 S6 M1
ΔPKI
CC6214-5GG
0 frame stop
B
kDa
WT ΔPKI comp
(-) WT U6186C U6186G G6217C G6217A G6217U
+1 frame stop
(-) WT S1 S2 S3 S4 S14 S36
0 frame stop
WT S1 S2 S3 S4 (-)
[35S]-Met labeling
72–
55–
43–
34–
26–
17–
10–
ORFx-FLuc (+1 frame)
RLuc
sORF2 (0 frame)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21
RLA
RLuc
FLuc
WT ΔPKI Com
WT U6186C U6186G G6217C G6217A G6217U
WT S1 S2 S3 S4 S14 S36
+1 frame stop
WT S1 S2 S3 S4
0 frame stop
C
RLA
RLuc
FLuc
WT ΔPKI S7 S11 S12 S13 S14 S15 S19 S3 S5 S4 S4 S6 (0) S4 (+1) S5 (+1) M14R
+1 frame stop codon
0 frame stop codon
WT ΔPKI M1

Figure 4.1. The translation of +1 frame CrPV ORFx is IGR IRES-dependent and initiates downstream of the PKI region

A) Schematic of the various mutations introduced downstream of the IGR IRES. The PKI region and downstream 'spacer' sequence of CrPV IGR IRES are shown. A series of mutations and stop codons were introduced either on the 0-frame or +1-frame. **B)** Analysis of the IGR IRES-dependence and initiation site of +1-frame ORFx. Linearized reporter constructs are incubated in Sf21 translation extract at 30°C for 2 hours in the present of [^{35}S]-methionine/cysteine. Translation of FLuc and RLuc was monitored by autoradiography after resolving on a 16% SDS PAGE. A representative gel from at least three independent biological experiments is shown. * = p-value < 0.05. **C)** Quantitation of IRES-mediated translation. Translation of FLuc and RLuc monitoring FLuc/RLuc enzymatic activity were quantified and normalized to wild-type (WT) IAPV IRES. RLuc monitors scanning-dependent translation acting as an internal control. RLA- relative light units. Shown are averages from at least three independent biological experiments (± SD). * = p-value < 0.05. Experiments here performed by Dr. Q.S. Wang.

4.3.2 The spacer region downstream of the IGR IRES is necessary for +1 frame translation

The results demonstrated that inserting stop codons in the 'spacer region' (Figure 4.1A) between the IRES PKI domain and the +1-frame 13th codon of ORFx inhibited but did not completely abolish +1-frame translation, suggesting that an element within this spacer region may promote +1-frame translation. We first addressed whether the CrPV spacer region is sufficient to direct +1-frame translation by generating a chimeric construct whereby the PSIV IRES is fused with the CrPV spacer region (Figure 4.2). The PSIV IRES directs strong 0-frame but no or relatively weak +1-frame translation (see Appendix A). The PSIV-CrPV spacer chimeric reporter resulted in 0-frame translation (Figure 4.2A), indicating that the spacer region does not affect the activity of the PSIV IRES. In contrast to the PSIV IRES construct, the chimeric PSIV-CrPV reporter resulted in +1-frame translation, implying that the CrPV spacer region is sufficient to drive +1-frame translation in the presence of a functional IGR IRES

(Figure 4.2A).

To examine more closely the molecular determinants of the CrPV spacer region that direct +1-frame translation, we deleted or inserted a randomized sequence within the spacer region to test whether the length was important (see Appendix A; Figure A.3). Insertions within or deletion of the entire spacer region resulted in loss of +1-frame but not 0-frame translation, suggesting that sequence context is important for +1-frame translation. Additionally, we examined the possibility of the spacer region to form intramolecular interactions. *In silico* analysis using RNA folding algorithms did not reveal any obvious stable secondary structures in the spacer region (data not shown), however we noticed that nucleotides $AUUU_{6222\text{-}25}$ and $AAAU_{6240\text{-}43}$ have the potential to interact. Mutating $UUU_{6223\text{-}25}$ to AAA (A/U) or $AAAA_{6240\text{-}43}$ to UUUU (U/A) drastically reduced +1-frame translation (see Appendix A; Figure A.3). However, compensatory mutations did not restore +1-frame translation. To delineate a functional element within the spacer region, we systematically tested deletions from the 3' end of the spacer region (Figure 4.2B). Interestingly, deleting 3 to 27 nucleotides did not affect the +1 frame activity. However, deleting 30 nucleotides and leaving seven nucleotides adjacent to the IRES abolished +1 translation (Figure 4.2B). Interestingly, Δ3, Δ6 and Δ27 mutants appear to have much higher Fluc/Rluc activity when compared to WT, an observation that requires further examination. In summary, these results suggest that the specific sequences within the spacer region, in particular the sequences immediately downstream of the PKI domain, are important for mediating IRES-dependent +1 frame translation.

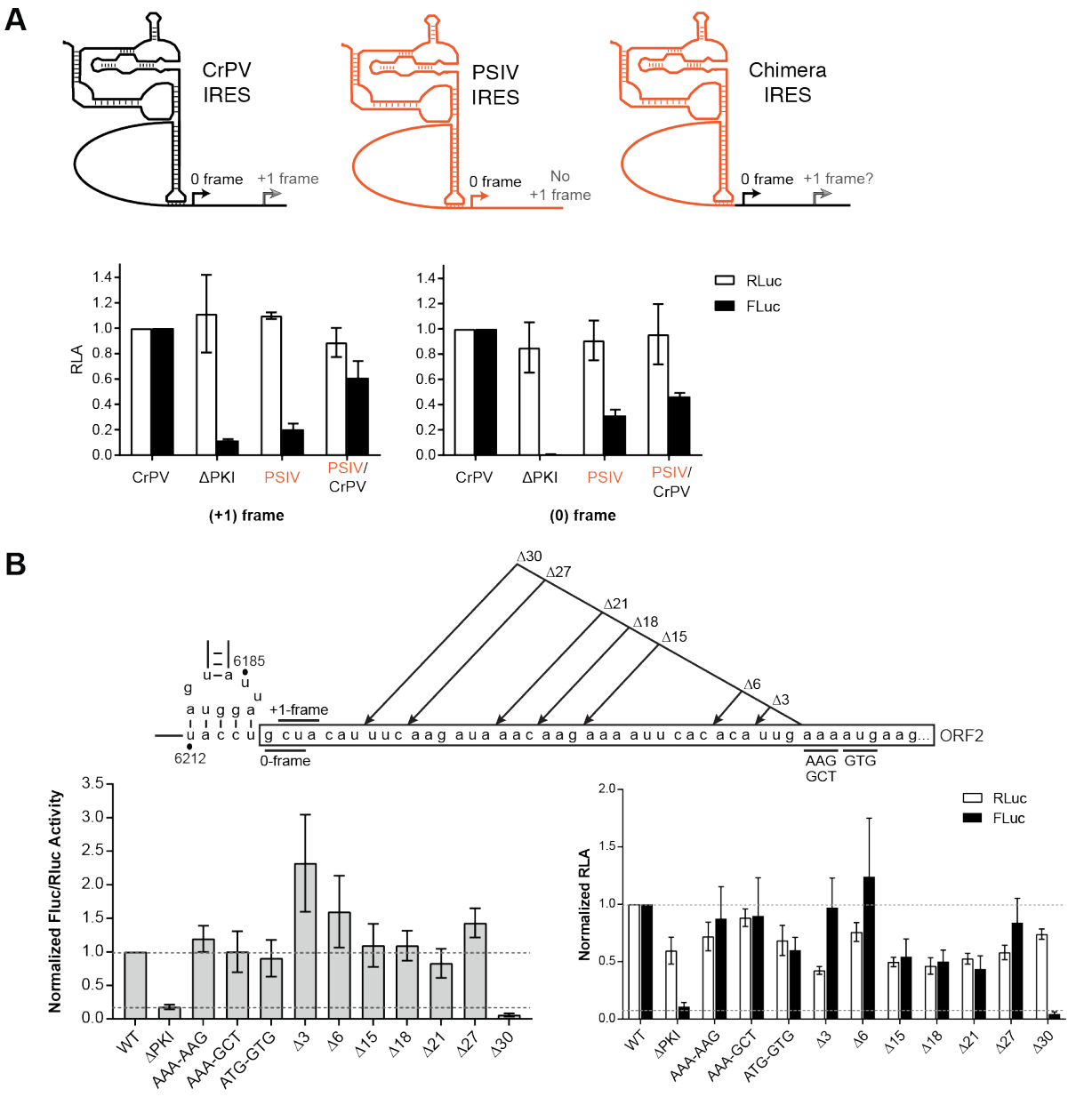


Figure 4.2. The downstream sequence of CrPV IGR IRES promotes +1-frame translation.

A) A chimeric IRES displays +1-frame translation activity. The PSIV IGR IRES, which shows no +1-frame translation activity (Figure 1 and Figure 3A), is combined with the downstream sequence of CrPV IGR IRES PKI region. The luciferase activities are normalized to the wild-type CrPV constructs in both +1- and 0-frame. The PSIV-CrPV chimera in the 0-frame is used as control to demonstrate that PSIV IGR IRES functions under this context. Experiments performed by Dr. Q.S. Wang. **B)** Truncation of the spacer region impairs +1-frame translation from the CrPV IGR IRES. Schematic displaying truncations within the spacer region between the PKI domain and the +1 frame start site is shown above. *In vitro* translation assay of the CrPV IRES bicistronic construct with the

respective mutations were monitored by luciferase activities. Both RLuc and FLuc activities are normalized to that of WT constructs. Results are shown as a normalized ratio of FLuc/RLuc activity (left) and as separated RLuc and Fluc activities (right) Shown are averages from at least three independent biological experiments (± SD). RLA = relative luciferase activity.

4.3.3 Molecular determinants of CrPV +1-frame translation

The current data led us to generate two hypotheses: i) a subset of 40S subunits recruited to the CrPV IRES scan downstream to start translation at the 13th codon (scanning hypothesis) or ii) 40S or 80S ribosomes recruited to the IGR IRES bypass the spacer region to the 13th codon for initiation (bypass hypothesis). To address the scanning hypothesis, we took a three-pronged approach. First, we utilized mutants of the CrPV IRES in the L1.1 loop that is known to be deficient in recruitment of the 60S subunit (Jang et al., 2009; Pfingsten et al., 2006). If scanning is occurring, 40S subunits recruited to the L1.1 IRES may scan downstream to the downstream +1-frame initiation codon. Reporter constructs harbouring mutations in the L1.1 loop were deficient in 0 and +1-frame translation (see Appendix A; Figure A.4), suggesting that 60S recruitment by the IGR IRES is necessary for translation in the +1-frame.

Secondly, we utilized the translational inhibitor edeine to assess if scanning was occurring (see Appendix A; Figure A.4). Edeine prevents the 40S ribosomal subunit from recognizing an AUG start codon (Dinos et al., 2004; Kozak and Shatkin, 1978). Both 0- and +1-frame translation were resistant to edeine relative to that of scanning-dependent translation. This result further supports that the +1-frame 14th AUG codon is not utilized for +1-frame translation.

Finally, we introduced a stable hairpin (HP1 ΔG = -17.0 or HP2 ΔG = -37.0 kcal/mole) into the spacer region between the CrPV IRES and the downstream 13th codon to prevent 40S scanning (see Appendix A; Figure A.4). Introducing HP1 or HP2 into the spacer region did not

affect 0-frame translation showing that these mutations did not affect IRES activity (Figure 4C). In contrast, +1-frame translation was decreased by ~50%. However, in these constructs, the hairpins are likely unwound due to 80S ribosomes translating in the 0-frame, thus confounding the interpretation of the +1-frame translation result (see Appendix A; Figure A.4). To uncouple these effects, we mutated the 0-frame 3rd codon to a stop codon upstream of the hairpin (0S3), thus blocking translating ribosomes from reaching the hairpin and maintaining hairpin integrity. Surprisingly, placing a stop codon upstream of the hairpin abolished +1-frame translation (see Appendix A; Figure A.4), suggesting that ribosomes may scan downstream to the 13th codon. However, given the insertion mutations that inhibit +1-frame, it is possible that these hairpin mutations are disrupting a functional RNA element within the spacer region that mediates ribosomes initiating at the downstream +1 frame initiation codon.

4.3.4 The integrity of the VLR and pseudo-translocation of the IGR IRES through the ribosome is required for ORFx expression

How do 80S ribosomes recruited to the IRES start translation at the downstream 13th +1-frame codon? Normally, the IGR IRES, which occupies the ribosomal A site upon ribosome binding, undergoes pseudotranslocation in order for delivery of the first aminoacyl-tRNA to the A site. We reasoned that the ribosome bound to the CrPV IGR IRES must have an empty ribosomal P or A site in order to reposition downstream and accommodate the +1 frame start codon. To address whether translocation of the IGR IRES through the ribosome is required for +1 frame translation, we introduced mutations in the VLR, which has been shown to disrupt the pseudotranslocation event (Ruehle et al., 2015).

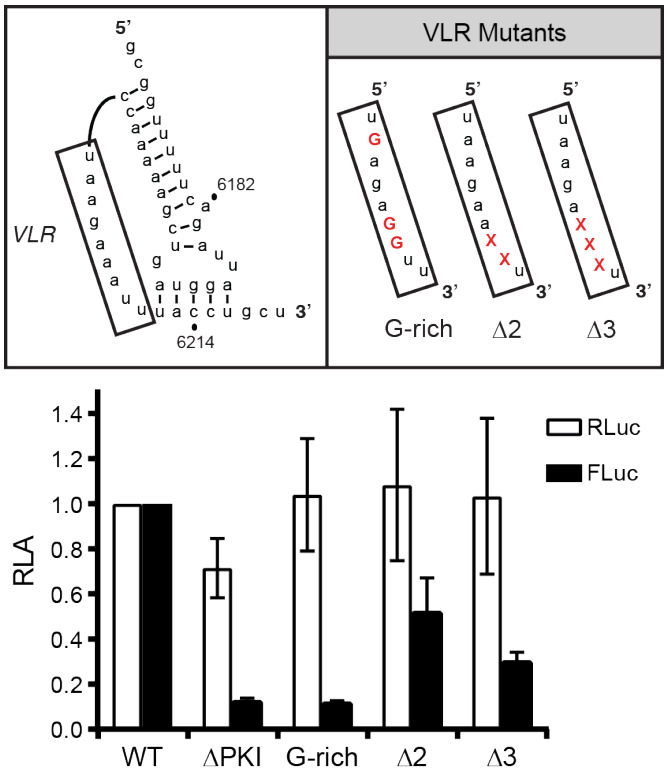


Figure 4.3. Pseudotranslocation is necessary for +1-frame translation.

Constructs containing mutations in the variable loop region (VLR) of the IGR IRES known to disrupt pseudotranslocation were monitored for translation activity. Luciferase activities were compared between WT IGR IRES and VLR mutants after *in vitro* translation assays. Shown are averages from at least three independent biological experiments (± SD).

Specifically, shortening the length of the VLR by 2 or 3 nucleotides (Δ2 and Δ3, respectively) inhibits the first pseudo-translocation event from the A site to the P site whereas altering the identity of nucleotides A_{6204}, and $AA_{6208\text{-}9}$ to guanosines (G-rich) inhibited IRES translocation from the P site to the E site (Ruehle et al., 2015). Interestingly, all three VLR mutants decreased +1-frame activity (Figure 4.3); both the G-rich and Δ3 mutants demonstrated little to no activity, while the Δ2 mutant still exhibited roughly 50% activity to that of WT. This result is consistent with previous data that the Δ2 mutant IRESs are still able to accommodate a fraction of aminoacyl-tRNA in the A site (~25%), allowing translation to occur (Ruehle et al., 2015). Altogether, these results indicate that the pseudotranslocation event of the IGR IRES through the ribosome contributes to +1-frame translation downstream.

4.3.5 +2 frame translation mediated by the CrPV IRES

We next asked whether repositioning of ribosomes bound to the IGR IRES to the downstream initiation codon is frame-dependent. To address this, we inserted a series of nucleotides into our bicistronic construct that shifts only the ORFx into the +2 frame (Figure 4.4). Specifically, we inserted either a single nucleotide or up to 7 proceeding the 13^{th} AAA codon. As a control, we inserted either 6 or 9 nucleotides in the same position that does not introduce an additional frameshift. Insertion of a U creates a stop codon in the +1-frame while an inserted C does not. If the ribosome merely begins translation in the 0-frame before slipping into the +1-frame, then we expect to see no ORFx translation. To our surprise, we observed ORFx expression with all insertions (Figure 4.4). These results suggest that repositioning of the ribosome from the IRES to a downstream 13th +1 or +2 frame codon is reading frame-independent.

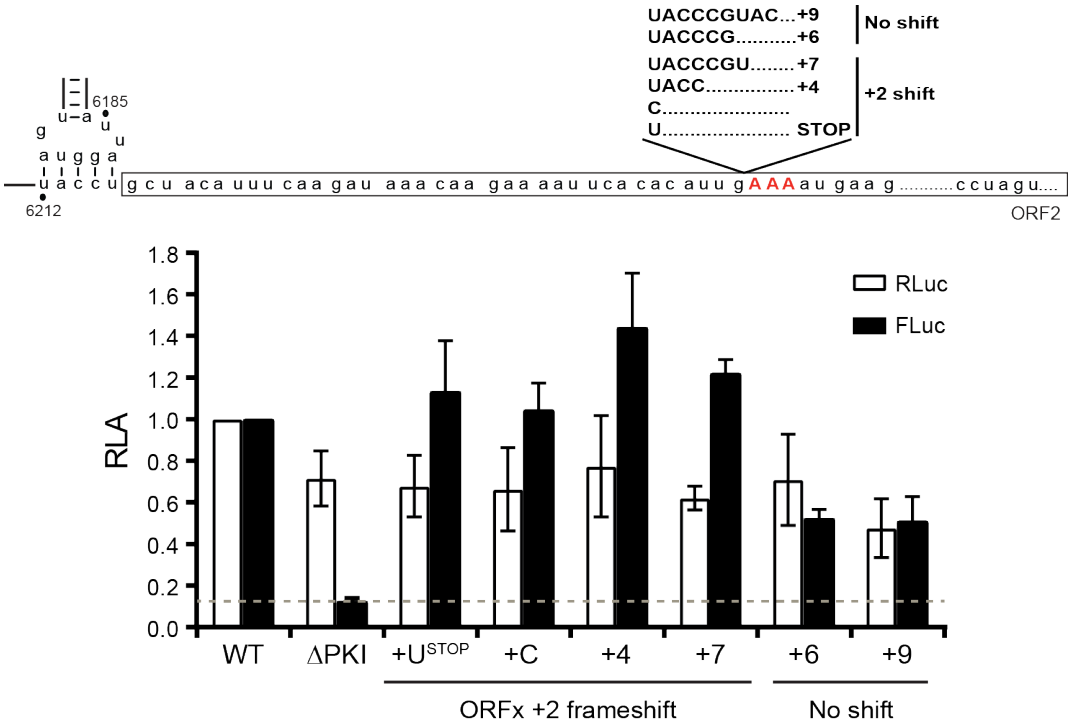


Figure 4.4. CrPV IRES-mediated +2 frame translation.

Mutations and insertions that shift the ORFx into the +2 frame are shown in the schematic above. Translation assays were monitored and normalized to that of the wild-type CrPV IRES. Shown are averages from at least three independent biological experiments (± SD).

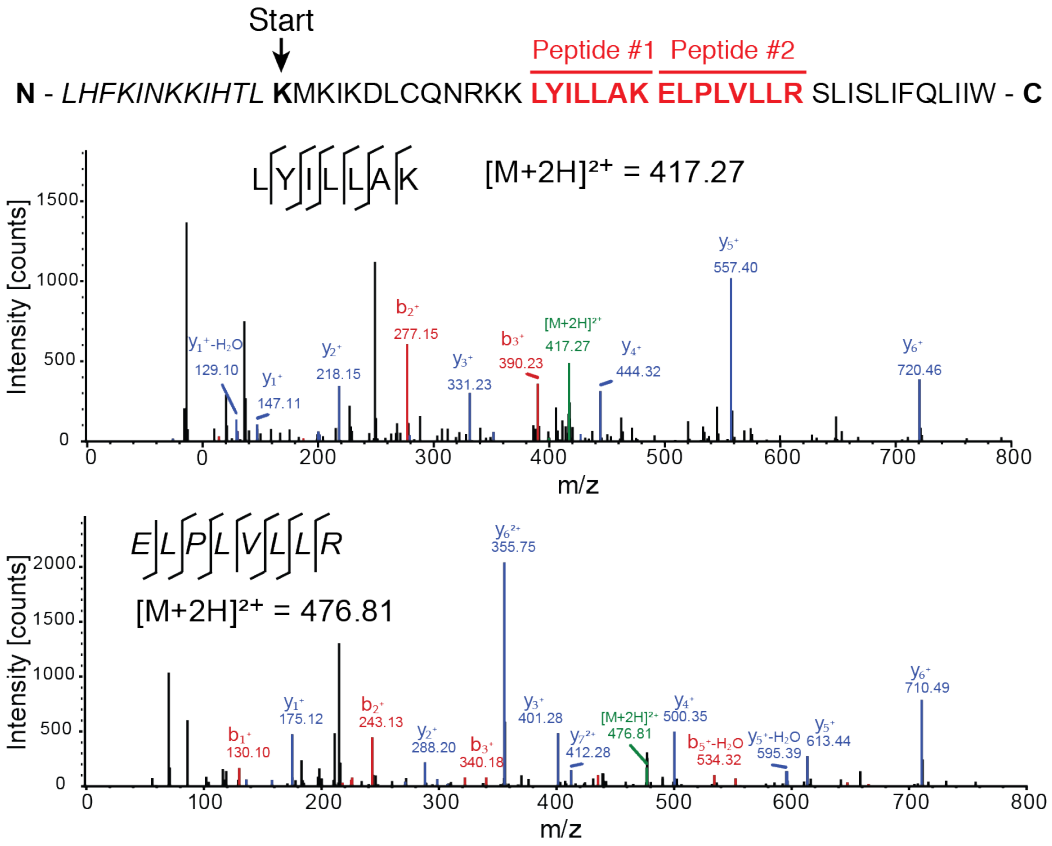


Figure 4.5. ORFx is expressed in CrPV infected S2 cells.

The predicted +1 frame amino acid sequence is shown above. Residues that are italicized represent the amino acid sequence if initiation occurred adjacent to the IGR IRES, whereas "Start" signifies the position of the 13th amino acid. *Drosophila* S2 cells were infected with CrPV (MOI 10). Proteins were extracted, digested with trypsin and subjected to LC-MS/MS analysis. Peptide fragment spectra were searched against a *Drosophila* uniprot database plus CrPV proteins. Two peptides (highlighted) were detected from the trypsin digestion of S2 cell lysate at 6 hours post infection. Individual fragment ions are annotated in the spectra and in the sequence representation.

4.3.6 CrPV +1-frame ORFx is expressed yet not required for infection in *Drosophila* S2 cells

Our *in silico* and biochemical data indicate that the CrPV IRES directs +1-frame translation from the 13th codon to synthesize a predicted 41 amino acid protein, termed CrPV ORFx (Figure 4.5). To determine whether CrPV ORFx is synthesized during infection, we used a proteomics approach to identify ORFx peptides in CrPV-infected S2 cells. Specifically, *Drosophila* S2 cells infected with CrPV (MOI 10) were harvested at 6 hours post infection and lysed. Proteins were subsequently digested with trypsin and peptides were analyzed by LC-MS/MS. We identified two peptides that correspond to CrPV ORFx both of which were located downstream of the +1-frame 13^{th} codon (Figure 4.5). Importantly, these peptides were not identified in mock-infected S2 cells. Thus, ORFx is expressed in CrPV-infected S2 cells.

We next sought to determine if CrPV ORFx expression influences the outcome of viral infection. Using a recently developed CrPV infectious clone, termed CrPV-2, we introduced mutations that would abolish ORFx expression (Kerr et al., 2015). To this end, we created two separate mutant clones: the +1-frame 12^{th} codon (UUG_{6251-3}) was altered to an amber stop codon (UAG; CrPV-S12) and the +1-frame 19^{th} codon (UUA_{6272-4}) changed to an ochre stop codon (UAA; CrPV-S19) (Figure 4.6A). Both mutations are synonymous in the 0 frame. Based on our translation data (Figure 4.1), the +1-frame S19 would inhibit ORFx expression whereas the +1-frame S12 would inhibit but not abolish ORFx expression.

We first examined whether ORFx influences viral protein synthesis *in vitro*. Incubation of *in vitro* transcribed CrPV-2 RNA in Sf21 translation extracts led to synthesis and processing of all viral proteins as reported previously (Kerr et al., 2015). Introduction of stop codons in the 0 frame of ORF1 or ORF2 inhibited synthesis of either viral non-structural or structural proteins,

respectively (Kerr et al., 2015). Both CrPV-S12 and CrPV-S19 RNAs resulted in viral protein synthesis that was indistinguishable compared to wild-type CrPV -2 *in vitro* (see Appendix A; Figure A.5). Note that CrPV ORFx is predicted to be ~4.9 kDa and is undetectable in this analysis. These results demonstrate that CrPV ORFx is not necessary for viral protein synthesis *in vitro*.

We next assessed the viability of the CrPV-S12 and -S19 viruses in cell culture. To determine if ORFx plays a role in the first replication cycle of CrPV, we transfected S2 cells with *in vitro* transcribed RNA from wild-type or mutant S12 or S19 CrPV-2 to generate virus. After 48 hours, virus was harvested and used to infect naïve S2 cells at an MOI of 10 to follow the first round of infection. We monitored viral protein synthesis via metabolic [^{35}S]-Met/Cys labeling as well as viral protein and RNA accumulation by Western and Northern blotting, respectively. Infection with wild-type CrPV-2, CrPV-S12 or CrPV-S19 all resulted in accumulation of viral proteins and RNA and shutdown of host translation in a similar manner (see Appendix A; Figure A.5). Furthermore, there was no visible difference in viral RNA and protein levels between wild-type and mutant virus infections. These results indicated that ORFx does not play a role in the first round of CrPV replication in S2 cells.

To determine if ORFx influences subsequent rounds of infection, we infected S2 cells at an MOI of 1 and monitored viral titres (see Appendix A; Figure A.5). Neither infection at an MOI of 1 produced significantly different titres between wild type CrPV-2 and either stop codon mutant at any time point, apart from both CrPV-S12 and CrPV-S19 resulting in higher titres than CrPV-2 after 24 hours post infection at MOI 1. A similar result was observed in CrPV-infected S2 cells at an MOI 0.1 (data not shown). Taken together, ORFx has no observable effect on the life cycle of CrPV in S2 cells.

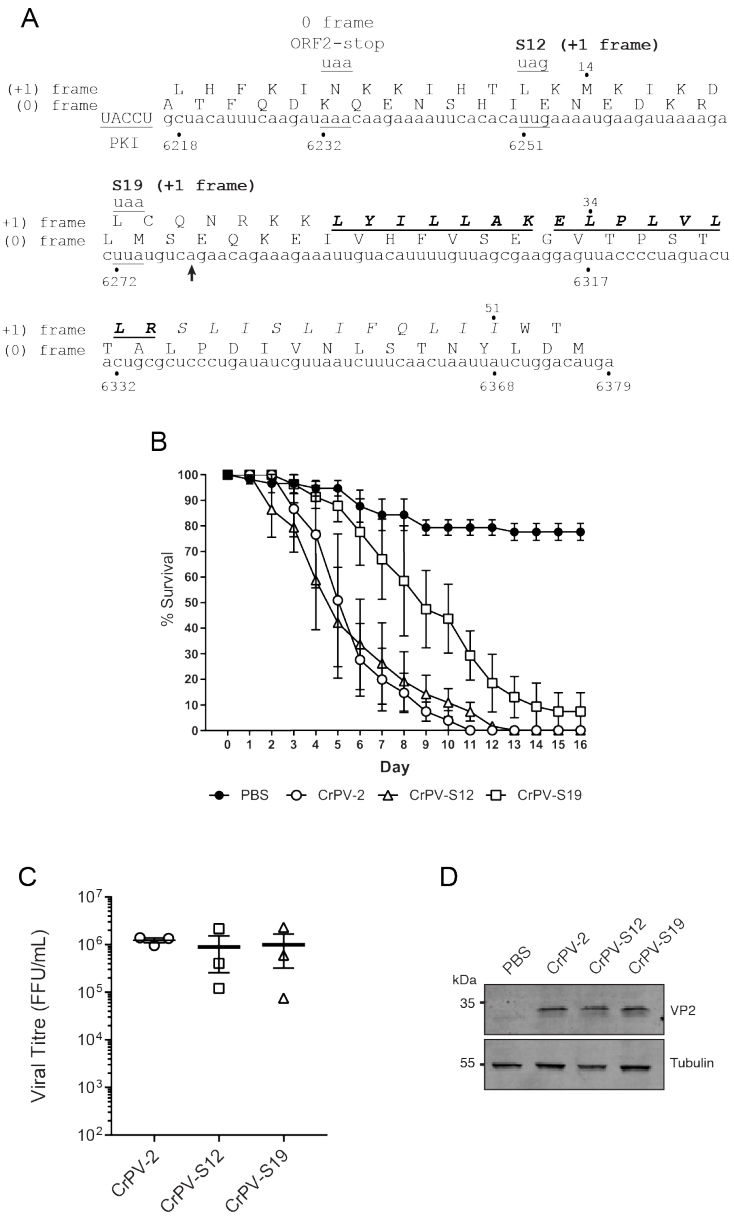


Figure 4.6. Pathogenesis of a CrPV mutant lacking ORFx is attenuated in adult *Drosophila melanogaster* flies. **A)** Schematic of mutations in the CrPV infectious clone, CrPV-2. The downstream nucleotide sequence of CrPV-2 IGR IRES and its potential amino acid sequence of +1-frame ORFx are shown. The ORF2-Stop mutant depletes the synthesis of 0-frame viral structural proteins. Mutants S12 and S19 place a stop codon on the +1-frame with a synonymous mutation in the 0-frame. Note that CrPV-2 (Accession: KP974706) sequence has a $C_{6279}A$ mutation (denoted with a black arrowhead) compared to the original CrPV (Accession: AF218039) sequence. Residues that are italicized, underlined and in bold are the peptides identified by mass spectrometry in Figure 4.5. **B)** Adult flies (Iso w^{1118}; 10 male, 10 female) were injected intrathoracically with 5000 FFU of CrPV-2, CrPV-S12, CrPV-S19, or

PBS. Subsequently, flies were flipped onto standard media and survival was monitored daily. Shown is a graph representing the average from three separate biological experiments (±SEM). **C)** Flies were harvested at 5 days post injection (dpi) and homogenized in PBS. Viral titres were determined as described in materials and methods. Error bars represent the standard error of the mean. **D)** Western blot against CrPV VP2 on lysates from flies 5 dpi. Shown is a representative gel from at least three independent biological experiments.

4.3.7 ORFx contributes to CrPV pathogenesis in adult flies and associates with membranes

We next addressed whether ORFx contributes to CrPV pathogenesis in adult fruit flies. To test this, we injected adult flies intrathoracically with PBS, CrPV-2, CrPV-S12, or CrPV-S19 and monitored fly mortality daily. By day 5, flies injected with CrPV-2 or CrPV-S12 exhibited 50% mortality and were 100% dead by day 11 and 12, respectively (Figure 4.6B). By contrast, flies injected with CrPV-S19 did not reach 50% mortality until day 9 and reached 100% mortality at day 14 (Figure 4.6B). These results demonstrate that ORFx contributes to CrPV pathogenesis in adult flies.

To determine if the effect seen on CrPV pathogenesis is a result of defects in viral replication, we measured viral titres and assessed viral protein levels in adult flies. At 5 days post infection, flies were homogenized and lysates were analyzed via Western blot for viral proteins and viral titres were determined. Interestingly, viral titres showed no significant differences between wild type CrPV-2, CrPV-S12, and CrPV-S19 (Figure 4.6C). As well, protein levels in infected flies were unaffected by the S19 mutation compared to wild-type or the S12 mutant (Figure 4.6C). Finally, to ensure that the S12 and S19 mutations in the mutant CrPV clones are stable, we performed RT-PCR analysis and sequenced this region. Our results showed that the S12 and S19 mutations are stable during virus propagation in S2 cells (data not shown). In

summary, our results indicate that the defect in viral pathogenesis observed in the S19 mutant virus is not due to a defect in viral replication.

From our data, CrPV ORFx initiates translation from the 13th +1-frame AAA codon to express a 41 amino acid protein. Comparing the sequence of ORFx from CrPV and other species shows no appreciable homology to other proteins. Nevertheless, *in silico* topology predictions suggest that ORFx has the propensity to form an alpha helical transmembrane segment at its C-terminus (amino acids 22-39; Figure 4.7). To examine ORFx function, we generated Drosophila expression constructs containing either N- or C-terminal HA-tagged ORFx. We first assessed whether expression of HA-ORFx affected cell viability by a trypan blue exclusion assay. Transfection for 10 or 24 hours with the ORFx-HA and HA-ORFx construct did not result in a dramatic decrease in cell viability (Figure 4.7B). However, by 48 hours, transfection of ORFx-HA and HA-ORFx led to a statistically significant but relatively minor reduction (13% and 15% decrease, respectively) in cell viability, suggesting that ORFx expression has a slight toxic effect in S2 cells. Immunoblotting for HA showed that HA-ORFx is stably expressed in S2 cells (Figure 4.7). To determine the subcellular localization in S2 cells, we transfected the HA-tagged ORFx constructs and monitored ORFx localization by HA-antibody immunofluorescence staining in comparison with cytoplasmic, ER, Golgi, and nuclear marker protein antibodies. Co-staining showed that HA-tagged ORFx overlaps mainly with ER protein marker Calnexin, and that a subset of HA staining overlapped with Golgi- and nuclear-associated proteins, Golgin84 and nuclear Lamin, respectively (Figure 4.8). By contrast, HA-ORFx displayed little to no overlap with α-Tubulin staining.

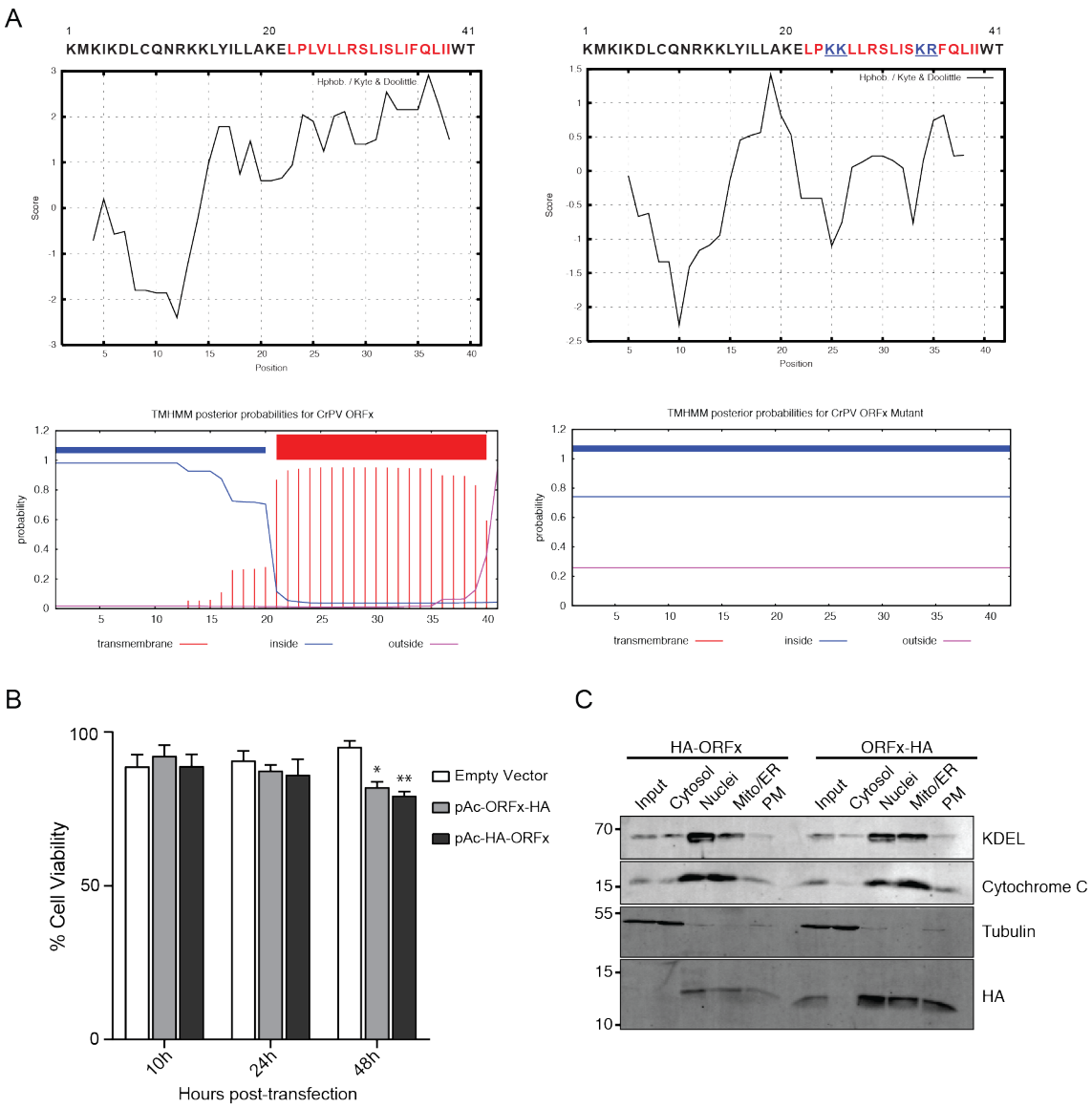
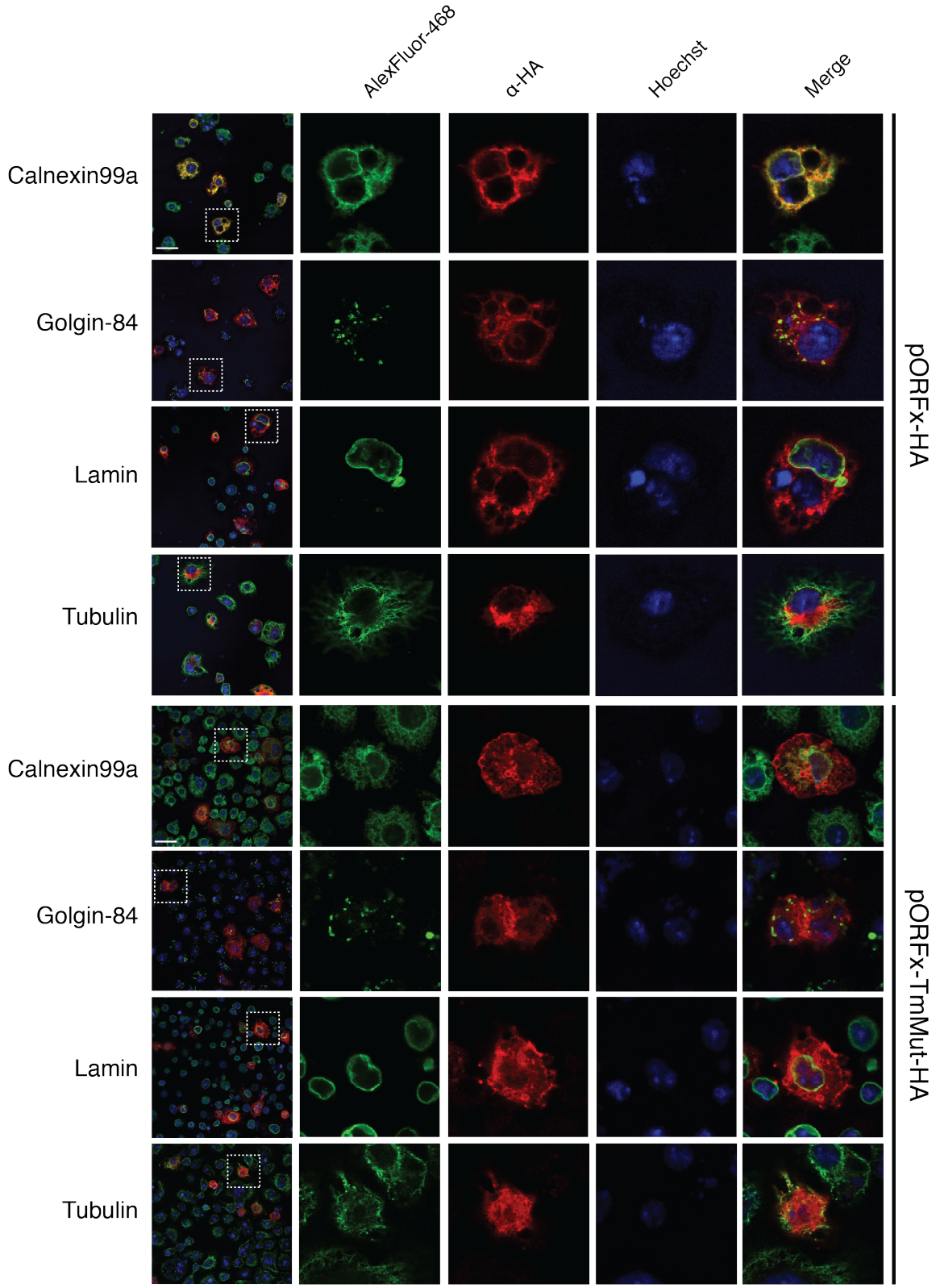


Figure 4.7. Cytotoxicity of CrPV ORFx expression in Drosophila S2 cells.

A) The transmembrane helix property of CrPV ORFx. The predicted CrPV ORFx sequence stemming from the 13th codon was analyzed for transmembrane helices using ProtScale and HMMTOP. Shown above is the hydropathy plot output from ProtScale for both the wild-type ORFx (left) and TmMut (right). The red amino acid region indicates the putative transmembrane helix predicted by HMMTOP (below). **B)** Ectopic expression of CrPV ORFx causes slight cytotoxicity in S2 cells. S2 cells (1.5 x 10^6) were transfected with expression constructs containing N- or C-terminally HA-tagged ORFx or empty vector and monitored over the course of 48 hours for cell death via trypan blue staining. * = p-value < 0.05. **C)** CrPV ORFx associates with membrane fractions. S2 cells were transfected with expression constructs encoding HA-tagged ORFx at the N- or C-terminus for 48 hours. Cells were then lysed and

fractionated by differential centrifugation (see material and methods). Fractions were then analyzed by Western blot to detect the presence of CrPV ORFx or cellular markers. Shown is a representative gel from two independent biological experiments.

To ensure that this localization was specific to propensity of the C-terminal of ORFx to form a transmembrane segment we generated mutations in this region that abolished the transmembrane segment (Figure 4.8). As expected, the mutant ORFx (ORFx-TmMut) no longer co-localized with Calnexin99a or Golgin84 and displayed a dispersed localization throughout the cell consistent with a non-membrane bound protein (Figure 4.8). Collectively, these results suggest that ORFx associates with membranous organelles and potentially the ER specifically (Figure 4.8). To examine this further using another approach, we transfected S2 cells with the HA-tagged ORFx expression plasmid and used differential centrifugation on cells to separate subcellular components. As controls, we monitored protein markers that should be enriched in different subcellular fractions. As expected, tubulin is found in the cytoplasmic fraction and the ER-associated KDEL protein and cytochrome C are enriched in nuclei and ER fractions (Figure 4.7). Both HA-tagged ORFx are detected within membranous fractions and not within the cytoplasmic fraction (Figure 4.7), implying that ORFx resides within the membrane of cells that may be important for its function.


AlexFluor-468
α-HA
Hoechst
Merge
Calnexin99a
Golgin-84
Lamin
Tubulin
pORFx-HA
Calnexin99a
Golgin-84
Lamin
Tubulin
pORFx-TmMut-HA

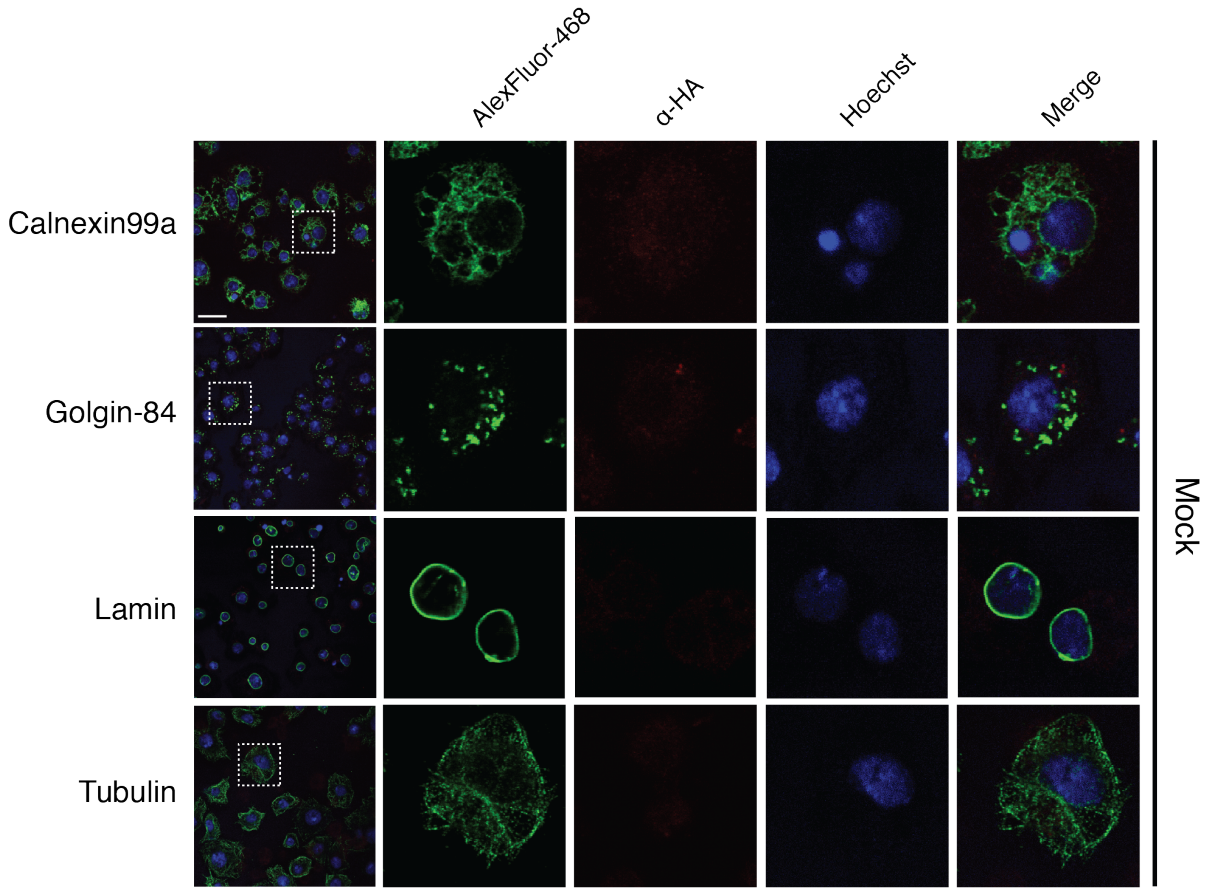


Figure 4.8. Subcellular localization of CrPV ORFx.

S2 cells were transfected with a construct expressing C-terminally HA-tagged ORFx, or a transmembrane mutant of ORFx and incubated for 48 h. Following incubation, cells were fixed, permeabilized, and co-stained with HA antibody with antibodies against either the nuclear lamina (lamin), the Golgi-associated protein Golgin84, ER-associated protein, Calnexin 99A, or alpha-tubulin. Shown are representative Z-stack micrographs from three independent experiments. Scale bars represent 15 µm.

4.4 DISCUSSION

Recoding mechanisms have illuminated diverse RNA structural elements that interact with the ribosome to affect reading frame maintenance (Dinman, 2012). Here, we have demonstrated a novel translational initiation recoding mechanism by which an IRES can promote ribosome repositioning to initiate at a downstream codon. In contrast to the IAPV IGR IRES, which directs ribosome reading frame selection (Ren et al., 2012), the CrPV IGR IRES can facilitate the expression of a downstream +1 overlapping frame, ORFx. We provide extensive analysis using mutagenesis that translation initiation of CrPV ORFx occurs 37 nucleotides downstream at the 13th AAA (Lys) codon. Moreover, we show that ORFx is expressed in CrPV-infected cells by mass spectrometry and that ORFx is required for promoting CrPV pathogenesis in a *Drosophila* injection model. Our data suggest a model whereby 40S and 60S ribosomal subunits bind to the CrPV IGR IRES, the IRES then undergoes a pseudotranslocation step placing PKI in the P site of the ribosome leaving the ribosomal A site empty. From here, the majority of the ribosomes continue translating in the 0-frame by delivery of the incoming Ala-tRNA^{Ala} to the 0-frame GCU codon whereas a small fraction of ribosomes bypasses or "slides" 37 nucleotides downstream to initiate translation at the +1-frame AAA (Lys) initiator codon (Figure 4.9). From our results, the following rules appear to apply to +1-frame translation in CrPV: i) the PKI must be intact, ii) both 60S and 40S ribosomal subunits are required to bind to the IRES, iii) pseudotranslocation of the IRES through the ribosome is necessary, iv) the spacer region between PKI and the 13th AAA codon is essential, and v) the nucleotide identity of the spacer region is crucial for efficient +1 frame translation.

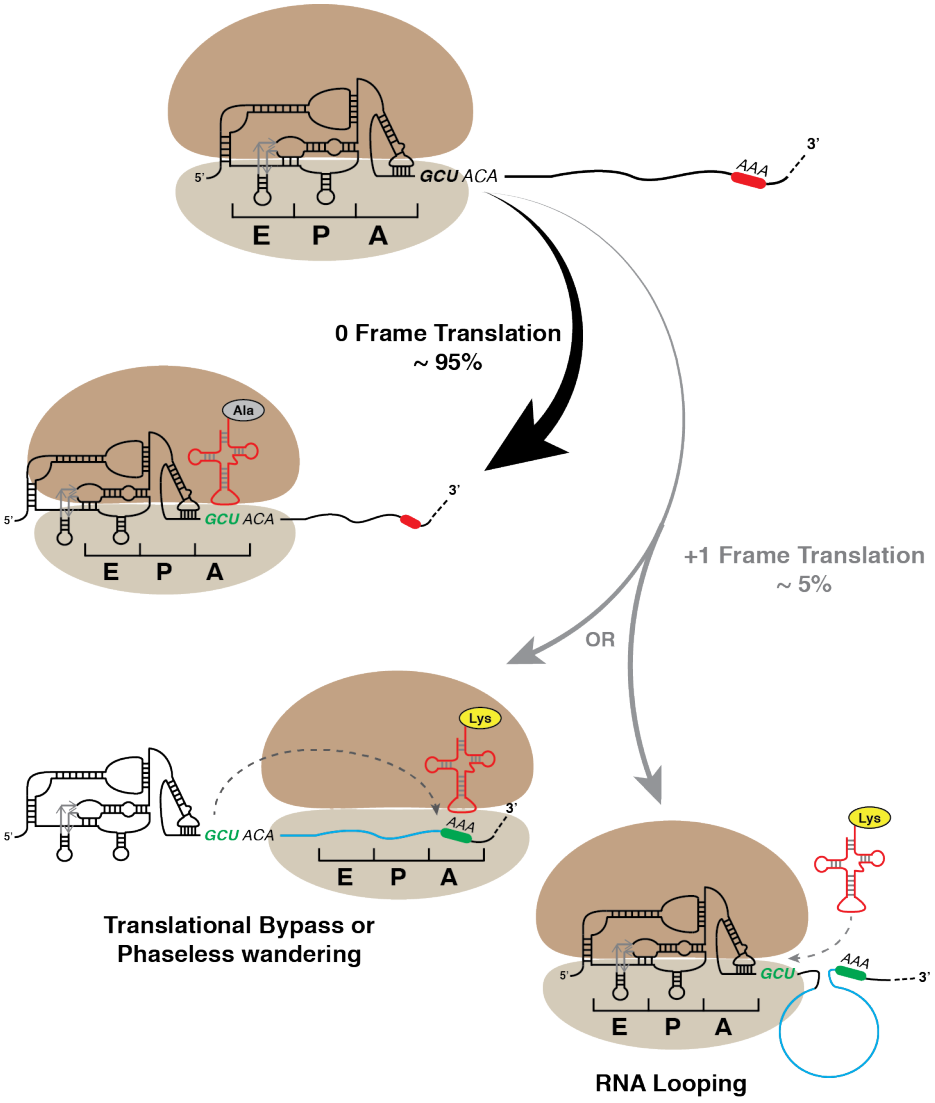


Figure 4.9. Models of CrPV IGR IRES mediated +1-frame ORFx translation initiation

The majority of ribosomes bounds to the IGR IRES initiate translation in the 0-frame from the GCU start codon. A subset of ribosomes bound to the CrPV IGR IRES (~5%) repositions downstream at the initiating +1 frame AAA codon. Possible models of IRES-mediated repositioning are shown and are not mutually exclusive. Whether a lys-tRNA^{lys} is utilized for +1 frame translation remains to be investigated.

A recent *in vitro* study using single molecule fluorescence spectroscopy by the Puglisi group demonstrated that the CrPV IGR IRES can facilitate +1-frame translation approximately 5% of the time compared to 0-frame translation whereby reading frame selection is dictated by the kinetics of tRNA binding in the first 0- or +1-frame codon (Petrov et al., 2016). In contrast, our results using more physiologically relevant approaches show that the CrPV IRES directs +1-frame translation using a different mechanism. Based on insect translation assays using bicistronic constructs and stop codon insertions in the downstream 'spacer' sequence, ribosomes that are recruited to the IRES to initiate translation in the +1-frame downstream of the initial codons (Figure 4.1). In support of this, insertion of a stop codon at the 12th +1-frame codon, but not the 19th, attenuates CrPV-mediated death in a Drosophila injection model, thus providing biological significance of CrPV +1-frame translation (Figure 4.7). We also note that the downstream authentic ‘spacer’ sequence is necessary for +1-frame translation, which is partially absent and/or altered from previous studies, which may preclude detection of bypass (Petrov et al., 2016).

A unique feature of this repositioning mechanism is that an intact IRES is essential for +1-frame translation (Figure 4.1). The current model is that the PKI domain of the IRES occupies the ribosomal A site upon ribosome binding to the IRES, followed by a pseudotranslocation event in order to vacate the A site to allow delivery of the first aminoacyl-tRNA in the 0-frame (Fernandez et al., 2014). Given this model, it is difficult to envision how the IRES can direct the ribosome to bypass 37 nucleotides to initiate translation at a +1-frame AAA Lys codon. A potential clue comes from our mutagenesis analysis that suggests that translocation of the CrPV IRES through the ribosome is a prerequisite for downstream +1-frame translation (Figure 4.3). This is intuitive as the PKI domain must vacate the A site to allow

translation in both the 0 and +1 frames. At this point, it is unclear whether the CrPV IGR IRES ribosome repositioning event requires delivery of an aminoacyl-tRNA to the A site of the ribosome prior to the ribosome sliding downstream to initiate translation at the 13th +1 frame codon or that a vacant ribosome bypasses from the IRES to the initiating +1 frame codon.

Another key feature is that the downstream spacer region is necessary for efficient CrPV IRES-mediated +1-frame synthesis (Figures 4.2, and Appendix A; Figure A.3). Interestingly, the majority of mutations within the spacer region causes a reduction ranging from 10-80% in the amount of CrPV IRES-mediated +1-frame synthesis (Fig 2), suggesting that nucleotide or codon identity is crucial for triggering the bypass event. Although our results altogether suggest a bypass event, frameshifting of elongating ribosomes may still contribute to +1 frame translation. There are no obvious RNA secondary structures within the spacer region, however, we cannot rule out any long distance RNA:RNA interactions that may contribute to +1-frame translation. Furthermore, sequences immediately downstream of the IRES PKI domain are important for +1 frame translation (Figure 4.2). Future investigation into how the spacer sequence influences ribosomes bound to the IRES to start translation in the 0 frame or at the downstream +1 frame initiating codon should shed light onto this mechanism.

Our results indicate that the CrPV IRES +1-frame starts translation at an AAA Lys codon, however, this is not a strict requirement as appreciable +1-frame translation still occurs when the AAA is altered to other codons (Figure 4.2). Although there is an AUG codon at the 14th codon, altering this codon did not eliminate +1-frame translation. In addition, +1-frame translation is edeine-insensitive and requires an intact IRES PKI domain and 60S joining, thus ruling out a 40S scanning mechanism per se (Figures 4.1 and Appendix A; Figure A.4). Although so far elusive and technically challenging, the identification of the aminoacyl-tRNA

that is delivered to the ribosomal A site for +1 frame ORFx translation should provide insights into the mechanism. Lastly, another interesting feature is that once ribosomes 'leave' the IRES, translation of the downstream codon is frame-independent, similar to that observed in the ribosome bypass event in *gene 60* of T4 bacteriophage (Samatova et al., 2014).

How does the ribosome reposition to the downstream +1-frame initiation codon after IRES binding? It is possible that the repositioning of the ribosomes occurs via a mechanism similar to that observed with prokaryotic 70S 'scanning' (Yamamoto et al., 2016) or 70S 'sliding' that occurs in coupled translational reinitiation (Adhin and van Duin, 1990; Nomura et al., 1984) (Figure 4.9). Moreover, it has been reported and proposed that energy-independent scanning or diffusion of the ribosome or ribosomal subunits (i.e. phaseless wandering) can occur to locate an AUG codon (Abaeva et al., 2016; Adhin and van Duin, 1990; Sarabhai and Brenner, 1967; Skabkin et al., 2013)(Figure 4.9). Our study has suggested that CrPV IRES +1-frame translation requires 80S ribosome binding to the IRES and is edeine-insensitive (see Appendix A; Figure A.4), thus we favour a model that 80S ribosomes repositions to the 13th +1-frame codon, however we cannot formally exclude a 40S repositioning event. Another possibility is that the IRES is anchored to the 60S subunit of the 80S via the L1.1. region, thus allowing the mRNA to freely move within the mRNA cleft of the ribosome until the downstream +1 frame start codon is reached. Furthermore, a related dicistrovirus PSIV IRES can direct translation using a scanning-like mechanism in prokaryotes, which may suggest a similar property observed with our studies on the CrPV IRES (Colussi et al., 2015). One model that warrants comparison is the translational bypassing observed in *gene 60* of T4 bacteriophage (Wills et al., 2008). In *gene 60*, translating ribosomes stall in a non-canonical rotated state at a 'take-off' Gly codon with a peptidyl-tRNA^{Gly}, which dissociates from the anticodon, and 'lands' at a matching Gly codon 50 nucleotides

downstream and allowing translation to resume. This bypass mechanism is dependent on a post-translocation step requiring a downstream 5' 'take-off' stem-loop structure and a nascent translated peptide (Chen et al., 2015a; Samatova et al., 2014). Although CrPV does not have an obvious RNA structure within the spacer region, it is possible that the highly complex structure of the IGR IRES itself may contribute to downstream +1 frame translation (Figure 4.9). Moreover, it is known that the ribosome bound to the IGR IRES is in a rotated state and that the first pseudotranslocation step is rate limiting, which may contribute to CrPV IRES reading frame selection (Fernandez et al., 2014; Zhang et al., 2016). Finally, it is possible is that CrPV IRES bypass could be occurring through an RNA looping event (Paek et al., 2015) (Figure 4.9); the downstream RNA is brought into close proximity with the 80S ribosome allowing it to transition to 13th AAA codon potentially by an unknown protein factor or a long-range RNA:RNA interaction. Taken together, it is likely that it is a combination of tRNA kinetics and conformational changes of the IRES and the downstream spacer region that lead to bypassing, of which the contributions of each element require further investigation.

The biological relevance of CrPV +1-frame translation was initially evidenced by the detection of ORFx peptides in CrPV-infected S2 cells (Figure 4.5). To our surprise, disruption of ORFx synthesis by stop codon insertion in the +1-frame did not perturb viral infection in tissue culture cells but showed retarded mortality in adult flies even though viral load remained similar between wild type and mutant viruses (Figure 4.7 and Appendix A; Figure A.5). Naturally, CrPV infection is thought to cause death through paralysis, subsequently leading to dehydration or starvation of the host (Chtarbanova et al., 2014; Reinganum, 1975). CrPV can infect several tissues in the fly including the trachea, midgut, and central nervous system although the latter has not been demonstrated directly (Bonning and Miller, 2010; Chtarbanova et al., 2014). How

ORFx contributes to CrPV pathogenesis is an outstanding question. Our results indicate that ORFx is membrane associated and does not contribute directly to viral replication but rather to the pathogenesis of CrPV infection in fruit flies. Furthermore, expression of ORFx is slightly cytotoxic in Drosophila cells, a property that may also contribute to pathogenesis of CrPV infection. Further studies into the localization of ORFx and tissue tropism of wild-type versus mutant virus infection in the fly should provide insights into the role of ORFx.

Viruses continue to surprise us with their ability to manipulate the ribosome in remarkable ways. Here, in addition to the previous findings on the honey bee dicistrovirus IRES, we have revealed another recoding mechanism utilizing an IRES, thus highlighting the strong selection to increase the coding capacity of the dicistroviridae genome. Furthermore, an IRES that can direct ribosome repositioning to facilitate the translation of a hidden +1 overlapping ORFx adds to the growing list of diverse pathways of ribosome translational initiation. The CrPV IGR IRES may serve as a model for other IRESs such as those from hepatitis C virus and Theiler's murine encephalomyelitis virus that can direct translation of alternative reading frames (Vassilaki and Mavromara, 2009; Yamasaki et al., 1999). It will be of considerable interest to investigate whether other IRESs direct reading frame selection by a similar CrPV IGR IRES-mediated ribosome repositioning mechanism. Ribosome repositioning or bypass is not specific to bacteria or mitochondria (Herr et al., 2000; Lang et al., 2014) but may be a more general phenomenon in eukaryotes that initially thought.

Chapter 5: CrPV may hijack exosomes to facilitate cell-to-cell transmission in *Drosophila* cell culture

5.1 INTRODUCTION

Traditionally, viruses have been categorized into one of two classes: enveloped or nonenveloped (Flint and American Society for Microbiology., 2009). Peplomers embedded within an envelope render it indispensable for those viruses that harbor one as it provides key determinants for cell tropism and mechanism of entry into host cells. Envelopes are acquired through non-lytic release from the host cell whereby the viral typically 'buds' from a distinct membrane in the cell. For example, hepatitis C virus, hantavirus, and influenza virus bud from the ER, Golgi, and plasma membrane, respectively (Romero-Brey et al., 2012; Rossman and Lamb, 2011; Vaheri et al., 2013). By contrast, the absence of an envelope bestows its own advantages that are not realized by enveloped viruses. Nonenveloped viruses are characteristically more resistant to chemical treatments and have greater environmental stability, thus allowing them to persist outside of a host for longer. Viruses belonging to this class, such as adenovirus and picornaviruses, assemble and accumulate in nonlumenal compartments until the host cell membrane is dismantled leading to release of the progeny virions. Whether a virus has evolved to possess an envelope or not has a substantial impact on how it is transmitted and recognized by the host immune system.

This central paradigm of viruses existing as enveloped or nonenveloped has been challenged by surmounting evidence demonstrating that non-lytic spread of several nonenveloped viruses occurs both in tissue culture and animal hosts (Bird et al., 2014; Chen et al., 2015b; Feng et al., 2013; Takahashi et al., 2010). Recent work showed that poliovirus, coxsackievirus B3, and rhinovirus virions are packaged *en bloc* into phosphatidylserine vesicles

that are non-lytically released from the host cell (Chen et al., 2015b). It is presumed that this mode of transmission increases infection efficiency and may permit genetic complementation between quasi-species. Although the exact mechanism of how these viruses are packaged into vesicles is unclear, the vesicles themselves are seemingly of autophagosomal origin and envelopment is dependent on the presence of PS and the viral receptor (Chen et al., 2015b). It is proposed that a double-membraned vesicle wraps the virions and then fuses with the plasma membrane, therefore releasing a single-membraned virion. This suggests that these viruses usurp the cellular autophagy pathway for their own benefit.

hepatitis A virus (HAV), another member of the picornaviridae, has also proven to acquire an envelope from host cell membranes (Feng et al., 2013). In contrast to PV and CVB3, during infection HAV hijacks the exosomal pathway for non-lytic release. HAV interacts with ALIX and thereby the ESCRT-III complex which drives the sorting of HAV virions into multivesicular bodies (MVBs) leading to membranous virions that are secreted from the host cell. The envelope acquired by HAV offers the advantage of cloaking it from neutralizing antibodies that circulate in the host blood leaving the virus able to spread from cell-to-cell. Nonenveloped HAV virions are found in the stool of infected chimpanzees suggesting that the acquisition of an envelope is not a passive mechanism and may be cell type specific (Feng et al., 2013).

In general, exosomes are small (~30 nm) vesicles of endocytic origin that are secreted by all mammalian cell types. They have been shown to contain active enzymes, metabolites, mRNAs, and small RNAs that are transported to adjacent cells (Raposo and Stoorvogel, 2013; Thery et al., 2002). Thus, exosomes are thought to participate in cell-to-cell communication. Exosomes have shown to play a significant role during viral infection by facilitating both the

host immune response and viral pathogenesis (Meckes and Raab-Traub, 2011; Robbins and Morelli, 2014). Interestingly, exosome-like vesicles have also been found to be secreted by insect cells, suggesting an ancestral cell-to-cell communication pathway (Koppen et al., 2011). In fact, a recent report demonstrated that exosome-like vesicles are secreted by haemocytes in virus-infected fruit flies that contain virally-derived siRNAs (Tassetto et al., 2017). These siRNAs are delivered to adjacent cells throughout the fly eliciting a long-lasting systemic RNAi response. Still, little is known if insect viruses, much like their mammalian counterparts, can utilize extracellular vesicles to their own advantage.

Here, we investigate the role of exosome-like vesicles in *Drosophila* S2 cells during CrPV infection. During infection, translation of the CrPV first open reading frame (ORF1) is driven by the 5'UTR IRES and encodes the viral non-structural proteins such as the RNA-dependent RNA polymerase (RdRp) and the 3C-like protease (Bonning and Miller, 2010). Conversely, the second open reading frame (ORF2) encodes the viral structural proteins and is translated via the intergenic region (IGR) IRES (Bonning and Miller, 2010). Although little is known about the full replication cycle of dicistroviruses, we demonstrate that CrPV, in addition to lytic release, can hijack exosome-like vesicles for non-lytic release to facilitate viral infection within *Drosophila* cells.

5.2 MATERIALS & METHODS

5.2.1 Cell culture and virus.

Drosophila Schneider line 2 (S2) cells were maintained and passaged in Shield's and Sang medium (Sigma) supplemented with 10% fetal bovine serum depleted in exosomes. Exosomes were depleted by spinning FBS at 120,000 RCF for 18 h at 4°C. Propagation and purification of CrPV in *Drosophila* S2 cells has been previously described.

5.2.2 Exosome isolation.

Exosomes were isolated from S2 cells as previously described (Koppen et al., 2011). $5.0x10^7$ cells were cultured in 50 mL of exosome-depleted media for 24 h. Cells and media were harvested and cells were pelleted at 300 RCF for 5 min. Cells were kept for further analysis downstream. Media was cleared of cellular debris by serial centrifugation at 5000 RCF for 10 min followed by 10,000 RCF for 30 min both at 4°C. The supernatant was then under laid with a 5 mL 45% sucrose cushion and ultracentrifuged at 100,000 RCF for 2 h at 4°C. The interphase containing the exosomes was kept by removing 45 mL of the supernatant and 2.5 mL of the sucrose cushion. Exosomes were then pelleted by diluting the interphase with 50 mL of 1X PBS and ultracentrifuging at 100,000 RCF for 2h at 4°C.

5.2.3 Sucrose and iodixanol gradients.

For density estimation of exosomes from S2 cells, exosomes were first isolated as above and then layered onto an 11 mL linear sucrose gradient (0.25 – 2 M sucrose). The gradients were then centrifuged at 100,000 RCF for 18 h at 4°C. Fractions were collected using an ISCO fractionator and density was determined by comparing fractions to a standard curve using a Brix/RI-Check refractometer (Reichart).

To separate exosomes and enveloped CrPV from non-enveloped virus we used isopycnic iodixanol, or Optiprep, gradients (Invitrogen) as previously described (Feng et al., 2013). Briefly, cell-culture supernatants were clarified by serial centrifugation at 1000 RCF for 10 minutes at 4°C, followed by 10,000 RCF for 30 minutes at 4°C. Virus and exosomes were then pelleted by centrifuging at 100,000 rCF for 1h at 4°C. The resulting pellet was resuspended in 500 µL of 1X PBS and layered onto an 8-40% iodixanol step gradient. Gradients were centrifuged at 141,000

RCF in a SW 41 Ti rotor (Beckman) for 48 h at 4°C. Fractions were collected from the top by hand.

5.2.4 Protein digestion and duplex demethylation labeling.

Cell or exosome pellets harvested from CrPV-infected cells at 6 hpi were solubilized in 6M urea and 2M thiourea. Protein concentrations were determined via BCA assay (Thermo). Equal amounts of protein (50 μg) from each sample was reduced by adding 1 μg of dithiothretiol and incubating for 30 min at room temperature. Proteins were alkylated with the addition of 5 μg iodoacetamide and allowed to incubate for 20 min at room temperature. Samples were digested with LysC before dilution with 4 volumes of 50 mM ammonium bicarbonate and digestion overnight with trypsin. Digested peptides were purified and concentrated on C18 STAGE-tips, eluted in 80% acetonitrile, 0.5% acetic acid, and dried in a vacuum concentrator (Eppendorf). Dried peptides were resuspended in 100 mM triethylammonium bicarbonate and chemical demethylation labeling was performed using light (CH_2O; mock-infected samples) or heavy ($^{13}CD_2O$; CrPV-infected samples) isotopologues of formaldehyde. Labeled samples were combined and STAGE-tip purified. Eluted samples were dried and resuspended in 20% acetonitrile and 0.1% formic acid.

5.2.5 Liquid chromatography tandem mass spectrometry (LC-MS/MS) analysis.

Digested peptides were analyzed by LC-MS/MS using a nanoflow HPLC (Thermo easy-nLC1000) coupled to a Q-Exactive mass spectrometer (Thermo). For each sample, 5.0 μg of peptides (based on the protein quantitation) were injected into the LC and loaded onto an in-house packed fused-silica (5 μm Aqua C18 particles (Phenomenex)) fritted trap column (2 cm, 100 μm I.D., 360 μm O.D., 5 μL/min flow rate, Buffer A = 0.5% acetic acid), then resolved on a reverse phase 75 μm inner diameter fused silica, in-house packed 30 cm analytical column

(ReproSil C18, 3 μm particle size (Dr. Maisch)) using a 75 min linear gradient run at 250 μL/min from 5% to 35% Buffer B (acetonitrile, 0.5% acetic acid), followed by a 15 min wash at 95% Buffer B. Instrument acquisition parameters included a 1% underfill ratio, 70 000 precursor mass resolution, 17 500 fragment mass resolution, normalized collision energy (NCE) of 28%, +1 and unassigned charges were excluded, "exclude isotopes" was turned on, intensity-dependent MS/MS at 1.7e5 intensity threshold, and the instrument was set to scan from 300 to 2000 m /z with a 30 s dynamic exclusion time.

Data were searched using MaxQuant (v1.5.3.30). Parameters included: carbamidomethylated cysteine (fixed), methionine oxidation (variable), glutamine and asparagine deamidation (variable); trypsin specific; maximum 2 missed cleavages; 10 ppm precursor mass tolerance; 0.05 Da fragment mass tolerance; 1% FDR; +1 to +7 charge states; match between runs and re-quantify; and common contaminants were included. Protein groups required a minimum of 1 peptide to be identified and a minimum of 2 peptides for quantification. Both the *Drosophila* and CrPV protein databases used were downloaded from UniProt (www.uniprot.org; April 23rd, 2014). Using Perseus (v1.5.2.6), contaminants and reverse hits were filtered out, protein ratios were Log_2 transformed, and ratios were averaged between replicates.

For GO term enrichment analysis, we employed the Gene Score Resampling (GSR) in 4ErmineJ v3.02, using the Log2 transformed protein ratios for "protein score" (Gillis et al., 2010). We considered a GO term to be significantly enriched if the Bejamini Hochberg-corrected GSR p-value was less than 0.05.

5.2.6 RT-PCR.

RNA from whole cells was extracted using TriZol reagent (Invitrogen). Alternatively, RNA was isolated from gradient fractions using phenol:chloroform extraction. Briefly, each fraction was brought up to 250 µL with RNase-free H_2O. An equal amount of phenol was then added to each fraction, samples were vortex and centrifuged at 4°C at 13.2K RPM for 30 minutes. The aqueous layer was removed and placed into a new tube with and equal volume of chloroform. The sample was vortexed and centrifuged for 5 minutes at 4°C at 13.2K RPM. The aqueous layer was once again removed and an equal volume of 100% ispropanol was added in addition to 400 mM of ammonium acetate. Samples were incubated at -20°C for at least 2 h. RNA was then pelleted by centrifugation for 10 minutes at 13.2K RPM. The RNA pellet was washed 3 times with 75% ethanol before resuspending in 10 µL RNase-free H_2O.

For RT-PCR on gradient fractions, the entirety of the extracted RNA was used and 100 ng of *in vitro* synthesized firefly luciferase RNA was added as a standard to each fraction. RT-PCR was performed using Superscript Reverse Transcriptase III (Invitrogen) at 50°C using a random hexamer primer. CrPV cDNA was amplified using primers P1 (5' – TCCTCAAGCCATGTGTATAGGA – 3') and P2 (5' – GTGGCTGAAATACTATCTCTGG – 3') while FLuc cDNA was amplified using primers P3 (5' – ATGAACGTGAATTGCTCAAC – 3') and P4 (5' – CCGGATTGTTTACATAACC – 3').

5.2.7 Western blots.

For cell and exosomal lysates, equal amounts of protein (10 µg) were resolved on a 12% SDS-PAGE gel and then transferred to a polyvinylidene difluoride Immobilon-FL membrane (Millipore). Membranes were blocked for 30 min at room temperature with 5% skim milk in TBST. Blots were incubated for 1 h at room temperature with the following antibodies: CrPV

ORF1 (raised against CrPV RdRp) rabbit polyclonal (1:10,000), CrPV ORF2 (raised against CrPV VP2) rabbit polyclonal (1:10,000), mouse anti-syntaxin 1A (1:5000; DSHB), mouse anti-ubiquitin (1:5000; AbCam), or mouse anti-Actin (1:5000; AbCam). Membranes were washed 3 times with TBST and incubated with either goat anti-rabbit IgG-HRP (1:20,000; GE Healthcare), goat anti-mouse IgG-HRP (1:5000; Santa Cruz), IRDye 800CW goat anti-mouse (1:10,000; Li-Cor Biosciences), or IRDye 800CW goat anti-rabbit (1:20,000; Li-Cor Biosciences) for 1 h at room temperature. An Odyssey imager (Li-Cor Biosciences) or enhanced chemiluminescence (Thermo) with ECL were used for detection.

For westerns performed on gradient fractions, proteins were first extracted using trichloroacetic acid (TCA). To each fraction an equal amount of H_2O was added for a total volume of 400 µL. 100 µL of 100% TCA was added and samples were incubated on ice for 10 minutes. Protein was pelleted by centrifugation for 10 minutes at 4°C at 13.2K RPM. Supernatant was removed and the pellet was washed twice with 200 µL of ice cold acetone. The final pellet was then dried by incubating at 95°C for 5 minutes and protein was resuspended in 20 µL of SDS-PAGE loading buffer.

5.3 RESULTS

5.3.1 Isolation of exosome-like vesicles from S2 cells during CrPV infection

To characterize exosome-like vesicles (ELVs) under CrPV infection, we isolated ELVs from *Drosophila* S2 cells that were mock-infected or infected with CrPV using differential ultracentrifugation as previously described (Koppen et al., 2011). Both the S2 cell pellet and the ELV pellet retrieved were analyzed by Western blot to assess the presence of exosomal protein markers (Figure 5.1A)(Beckett et al., 2013). As expected, both pellets contained β-actin, and ubiquitin. For ubiquitin, differential banding patterns were observed between the cell pellet and

ELV fraction (Figure 5.1A). Ubiquitin is thought to aid in protein sorting to exosomes on the cytosolic side of the multi-vesicular body, therefore an enrichment of specific proteins would be expected in the ELV pellet. Moreover, syntaxin-1A showed a strong enrichment in the ELV fraction compared to the cell pellet as seen previously (Figure 5.1A). In addition to Western blot analysis, we employed negative stain electron microscopy to assess the integrity of the vesicles. Electron micrographs of the vesicles displayed a range in size of 30-100 nm, typical of exosomes isolated from mammalian cells (Figure 5.1B). These results indicate that bona fide exosome-like vesicles can be isolated from S2 cells.

We next investigated ELVs from S2 cells infected with CrPV. We chose to infect cells for 6 h to mitigate contamination of the ELVs by cellular debris due to lysis. As seen previously, syntaxin-1A was enriched in the ELV fraction compared to the cell pellet and differences were observed in the distribution of ubiquitin between the cell pellet and the ELV fraction (Figure 5.2). Interestingly, distinct ubiquitylated proteins could be detected when comparing ELVs isolated from mock or CrPV-infected cells suggesting that there may be changes in protein content in the vesicles (Figure 5.2A). Using an antibody that recognizes the RNA-dependent RNA polymerase (RdRP) of CrPV, CrPV RdRP was present in the CrPV-infected cell pellets, but not in the ELV fractions suggesting that it does not get packaged into the ELVs.

To further validate that exosome-like vesicles were isolated, we measured their density via sucrose gradients. ELVs were isolated as before and then layered onto a 0.25 – 2 M sucrose density gradient. After centrifugation, the gradients were fractionated and the distribution of syntaxin-1A across the gradient was monitored by immunoblotting (Figure 5.2B). From both mock- or CrPV-infected cells, syntaxin-1A was present in fractions of a density range from 1.09

– 1.16 g/mL, consistent with the range of ELVs (Thery et al., 2002)(Figure 5.2B). Altogether these results demonstrate that exosome-like vesicles can be isolated from CrPV-infected cells.

5.3.2 Proteomic analyses of CrPV-infected cells and derivative ELVs.

Our Western blot analyses hinted that the protein composition of ELVs is different between mock- and CrPV-infected S2 cells. To address this, we utilized quantitative mass spectrometry to identify changes in the proteomes of ELVs. Proteins of ELVs from mock- or CrPV infected S2 cells (6 h) were isolated and digested with trypsin to produce peptides that were subsequently labeled with a 'light' formaldehyde reagent (Mock-infected) or 'heavy' formaldehyde reagent (CrPV-infected) before analyzing via LC-MS/MS. This analysis was also performed on cell pellets in parallel to determine if proteins enriched in ELVs was due to an increased abundance within the cell or specific packaging of host proteins. From this analysis, we identified and quantified 1428 and 802 proteins from the cell and ELV pellets, respectively, in 2 of 3 biological replicates (Figure 5.3). Of these, 913 proteins were specific to the cell pellet while 287 protein were detected only in the ELV samples. As well, 515 proteins were shared between both the cell and ELV pellets (Figure 5.3A). For proteins quantified in the cell pellet, only 3 displayed a >2-fold increase in protein abundance compared to mock-infected cells. These included both viral polyproteins (ORF1 and ORF2) and the uncharacterized protein, CG31731-RC (Figure 5.3B). The majority of proteins quantified in the cell pellet at 6 hpi decreased in protein abundance with 16 proteins being down-regulated more than 2-fold compared to mock-infected cells (Figure 5.3B). This agrees with the fact that CrPV shutdowns host protein synthesis during infection and synthesizes viral proteins in molar excess. In the ELV pellet, 40 proteins increased >2-fold in abundance at 6 hpi while 8 proteins decreased >2-fold (Figure 5.3B).

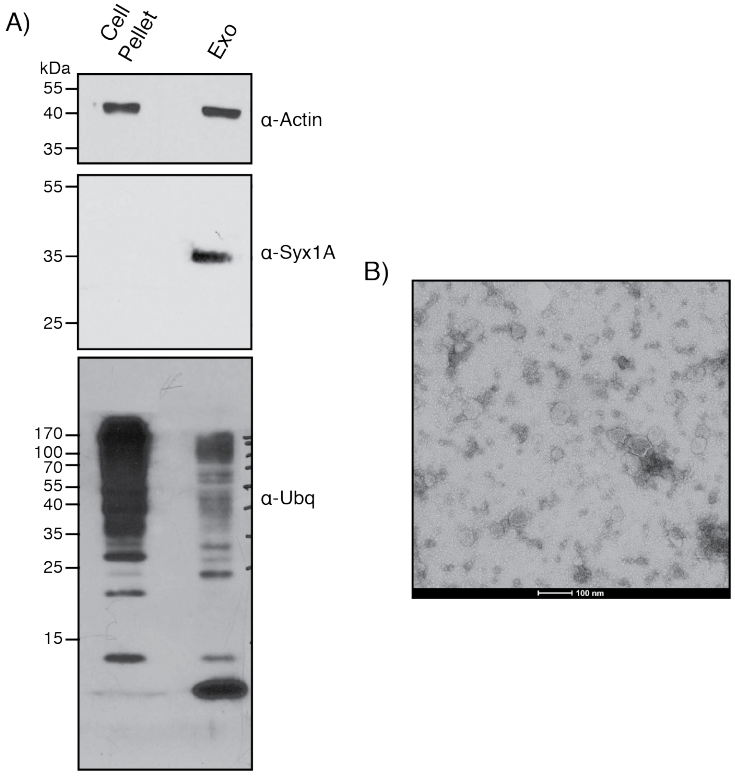


Figure 5.1 Validation of exosome-like vesicle isolation from *Drosophila* S2 cells.

(A) Western blot analysis of different ELV markers between the cell pellet and ELV pellet (Exo) after isolation. Syx1A = Syntaxin-1A; Ubq = Ubiquitin. **(B)** Transmission electron micrograph of ELVs isolated using differential ultracentrifugation.

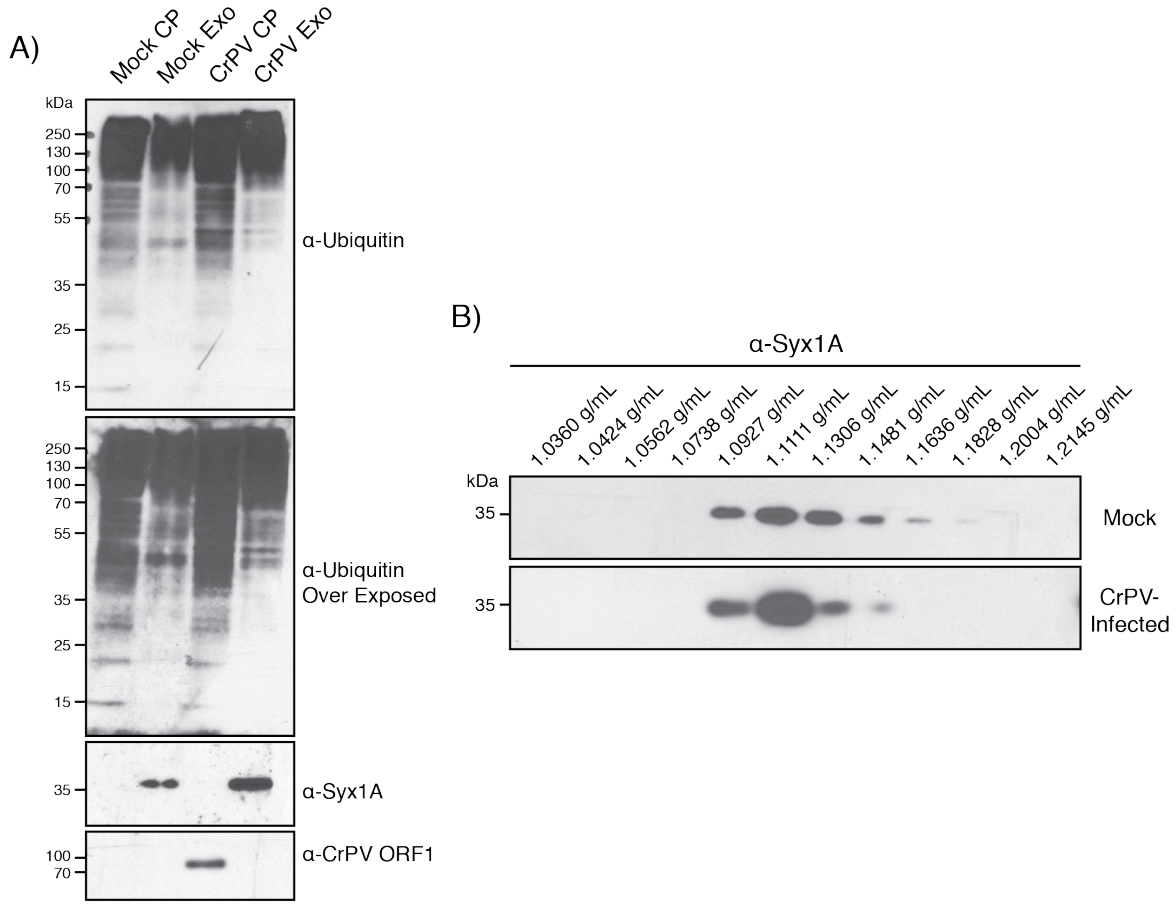


Figure 5.2 Exosome-like vesicle can be isolated from *Drosophila* S2 cells during CrPV infection.

(A) Western blot analysis of the cell pellet versus the ELV pellet (Exo) in either mock- or CrPV-infected S2 cells. Cells were infected at an MOI of 10 and exosomes were harvested at 6 hpi. Syx1A = Syntaxin-1A; Ubq = Ubiquitin.

(B) Sucrose density gradient separation of exosome-like vesicles isolated from mock- or CrPV-infected cells at 6 hpi.

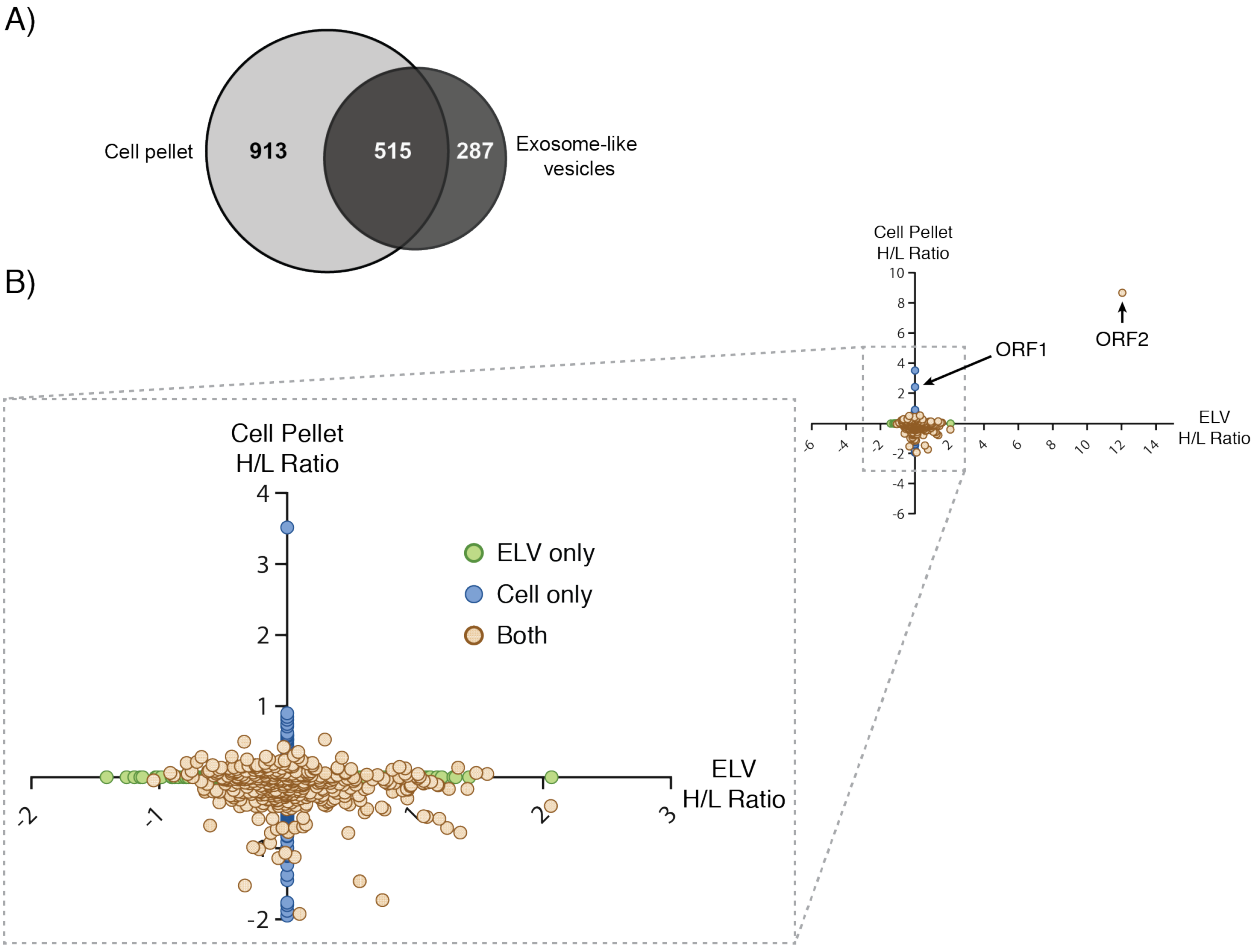


Figure 5.3 Quantitative proteomic analysis of ELVs isolated from mock- and CrPV-infected cells at 6 hpi. **(A)** Comparison of proteins identified in S2 cell pellets and ELVs. Cells and ELVs were collected from both mock- and CrPV-infected S2 cells at 6 hpi. Proteins were collected from all samples and equal amounts were subjected to trypsin digestion. Peptides were dimethylated with mock samples labeled with a light (L) formaldehyde and CrPV-infected samples with a heavy (H) formaldehyde. Peptides were then pooled and analyzed by LC-MS/MS. **(B)** Scatter plots of proteins quantified in S2 cell pellets in comparison to those in ELVs. Plotted is the Log2 transformed abundance ratios comparing the heavy to light labeled samples (H/L ratio). Boxed region is zoomed-in area of the graph lacking viral proteins.

We also identified multiple markers that have been associated with exosomes in previous studies in the ELV pellet such as ALIX, Syntaxin-1A, Rab35, and Flotillin-1, supporting that we were isolating bona fide exosomes. Surprisingly, ORF2 of CrPV was found in the ELV pellet ($Log_2 = 12.04$; Figure 5.3B and 5.4). However, there were no peptides mapped to the ORF1 of CrPV in the ELV pellet; congruent with our previous Western blot analysis (Figure 5.2).

To gain further insight into the cellular processes underlying proteins in the cell pellet and ELVs in CrPV-infected cells, we performed a gene score resampling (GSR) analysis using ErmineJ. Unsurprisingly, there were no significantly enriched terms found within the cell pellet protein list since CrPV infection results in global translation shutdown of the host and not inhibition of specific mRNAs. On the other hand, many GO terms were significantly enriched in ELVs such as those linked to transport (e.g. GO:0006818; GO:0006811; GO:0055085) and metabolic processes (e.g. GO:0006091; GO:0006793; GO:0051186) (see Appendix B). Given the biogenesis of ELVs occurs through invagination of the endosomal membrane, it is unsurprising that a large portion of proteins are annotated with metabolic functions as these processes largely occur in the cytoplasm. Altogether, these results suggest that the protein composition of ELVs during CrPV-infection is unique compared to that of the cell pellet and ELVs derived from mock-infected cells.

5.3.3 CrPV may hijack ELVs during infection

The observation that ORF2 (i.e. structural), but not ORF1 proteins from CrPV are found in ELVs suggested that CrPV could be commandeering these vesicles during infection. However, it is possible that CrPV virions are associated with the ELVs and not necessarily within the vesicles.

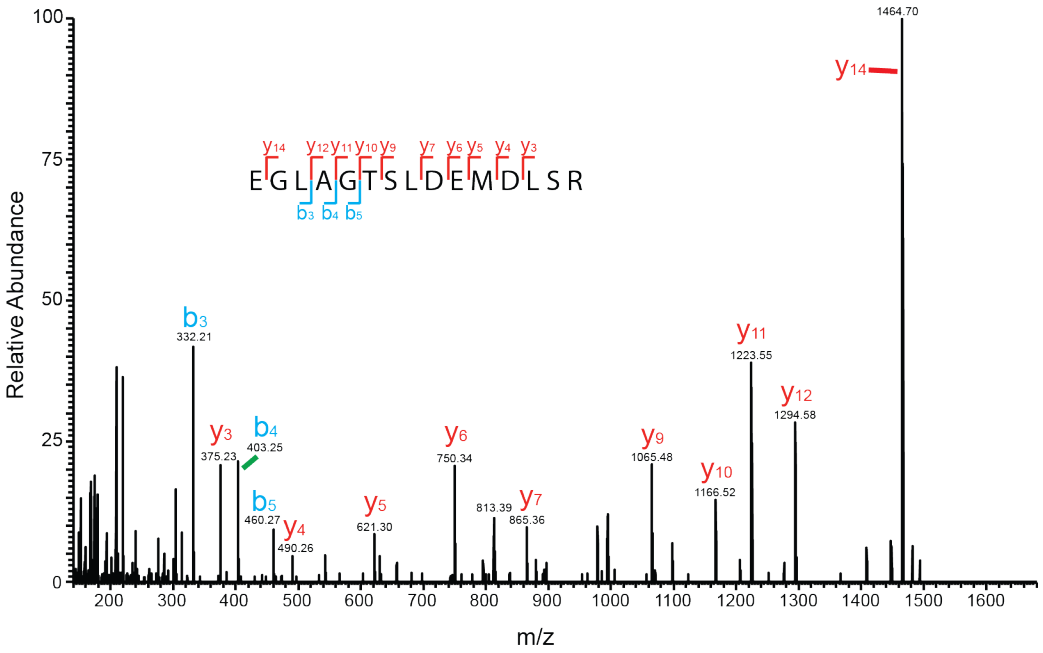


Figure 5.4 MS/MS spectra of one peptide identified from the CrPV ORF2 enriched in ELVs derived from CrPV-infected cells.

Fragment spectra from the 813.40 m/z doubly charged, formaldehyde labeled (+32 Da) precursor ion of EGLAGTSLDEMDLSR from trypsin digested ELVs. In total, 21 peptides were mapped to CrPV ORF2. Individual fragment ions are annotated in the spectrum and in the sequence representation. In red are y-ions, while b-ions are represented in blue.

To distinguish these possibilities, we performed iodixanol gradients on S2 cells infected with CrPV for 24 h. Iodixanol gradients are more sensitive to changes in sample densities and thus exosomes and virions fractionate differentially. After infection, supernatants were collected and ELVs (and virus) were pelleted via ultracentrifugation before layering onto an 8 – 40% iodixanol gradient. Each gradient was subsequently fractionated and analyzed by Western blot and RT-PCR for the presence of viral proteins or RNA, respectively. As suggested from our mass spectrometry data, both CrPV structural proteins and RNA distributed binomially into two distinct populations; one that co-fractionates with syntaxin-1A and one that fractionates at the bottom of the gradient (Figure 5.5), which likely represents assembled virions. To determine if the syntaxin-1A-containing fractions contain virions enveloped with a lipid, we treated the isolated ELVs with NP-40 before layering onto the gradients. NP-40 treatment abolished the bimodal distribution of CrPV VP2 and RNA resulting in an enriched population near the bottom of the gradient, thus suggesting that the syntaxin-1A-containing fractions that contained CrPV structural proteins and RNA are enveloped with a membrane (Figure 5.5). The disruption of syntaxin-1A is likely due to its association with lipid rafts. Lipid rafts are known to be resistant to many detergents and can result in aggregation of the rafts leading to a denser moiety in the gradient.

We next determined whether the light and dense fractions containing CrPV are infectious. We pooled fractions containing ‘naked’ CrPV (fractions 15-17) or ‘enveloped’ CrPV (eCrPV; fractions 5-9) together and extracted the virus through centrifugation (Figure 5.6A). We then infected S2 cells for 48 h and measured the resulting viral titre. Supernatants from cells infected with either CrPV or eCrPV resulted in an approximately equal amount of viral titre after 48 h, indicating that eCrPV is as infectious as its non-enveloped form (Figure 5.6).

To determine if the resulting CrPV-containing ELVs were a result of apoptosis/lysis of infected S2 cells we examined the integrity of the cell membrane during viral infection. We infected S2 cells with CrPV at an MOI of 1 and monitored cell membrane integrity via trypan blue alongside viral titres in the supernatant over the course of 8 hours (Figure 5.6B). At 2 hpi virus is detectable albeit at low levels which do not begin to increase until 6 hpi as to be expected. By 8 hpi we observed a 10-fold more virus present in the supernatant compared to the 2-hour time point (Figure 5.6B). Interestingly, even at 8 hpi the membrane of S2 cells remains intact, suggesting that the virus being released into the supernatant is not due to lysis. Taken together, these data suggest that CrPV can commandeer ELVs to facilitate its infection throughout *Drosophila* cells.

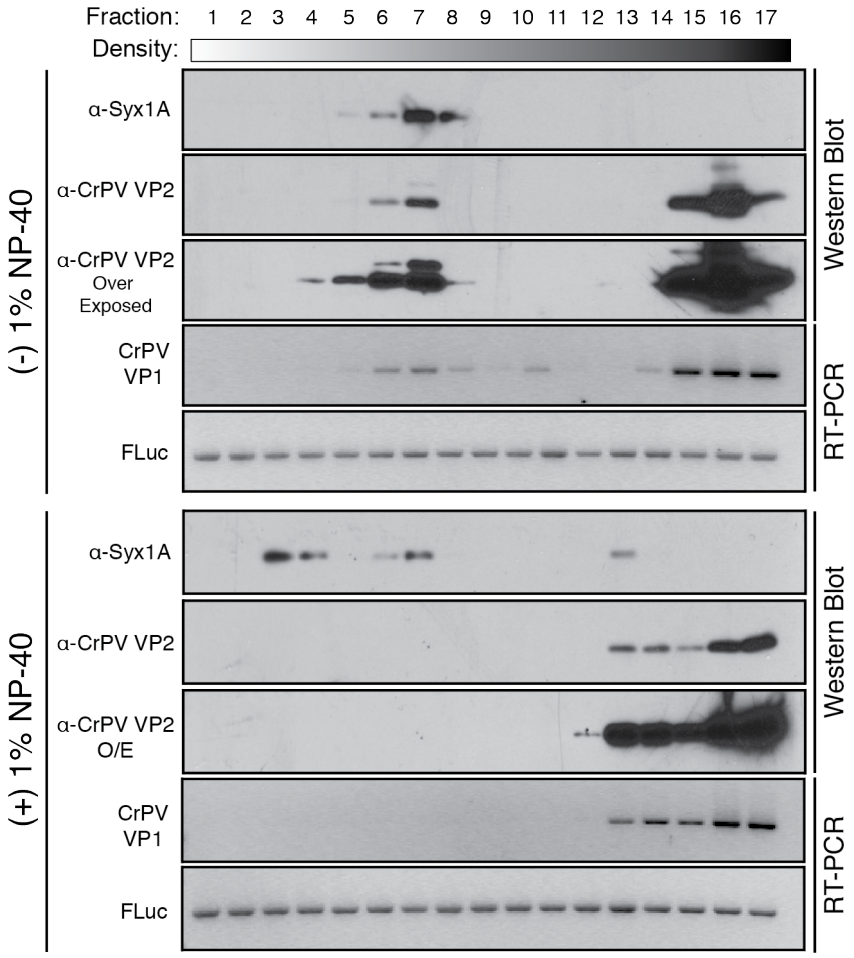


Figure 5.5 CrPV may hijack ELVs during infection of *Drosophilia* S2 cells.

S2 cells were infected with CrPV at an MOI of 10 for 24 hours. Supernatant from infected cells was collected and ELVs were isolated as described in the Materials and Methods. Isolated ELVs were then treated with or without 1% NP-40 and layered onto an 8-40% iodixanol gradient before centrifuging at 141,000g for 48 hours. Fractions were then collected, split, and protein was extracted via trichloroacetic acid precipitation while RNA was extracted by phenol-chloroform. Syx-1A was used as a marker for ELVs. *In vitro* synthesized firefly luciferase RNA was doped into each fraction to control for PCR efficiency. Shown is a representative gel from three independent experiments.

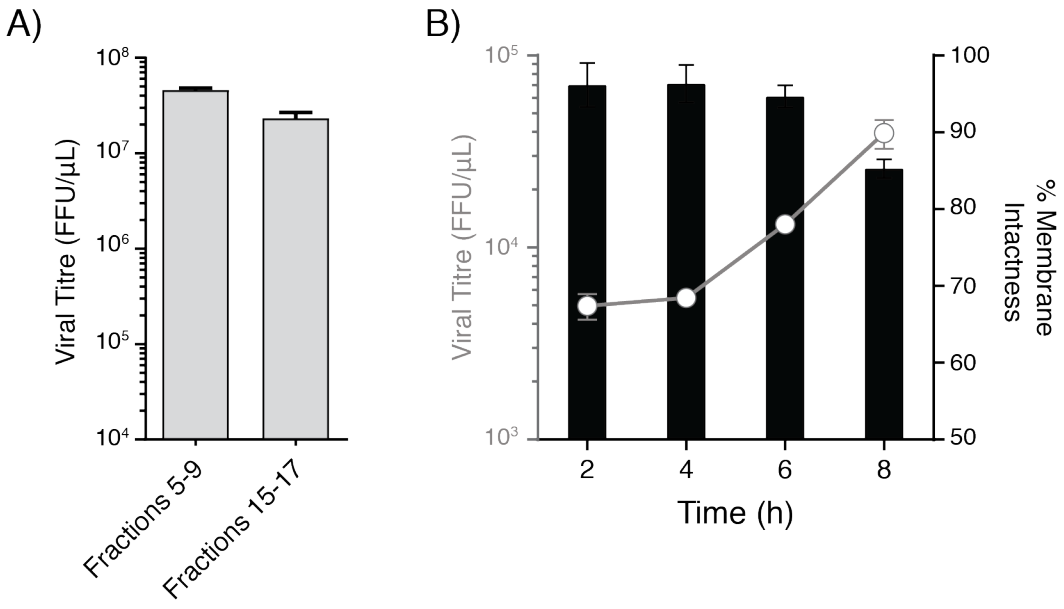


Figure 5.6 ELVs from infected S2 cells elicit subsequent infection and release from cells without lysis.

(A) ELVs isolated from CrPV-infected cells at 24 hpi were separated on iodixanol gradients. Fractions were then collected, pooled, and added to naïve S2 cells. After 48 hours, viral titres were determined using a fluorescence foci forming assay. Shown is the average from three biological replicates (±SD). **(B)** Membrane integrity remains as extracellular viral loads increase. S2 cells were infected with CrPV at an MOI of 10 and viral loads were measured in the supernatant at various time points throughout infection via a fluorescence foci forming assay. Trypan blue staining of S2 cells was used as a proxy for membrane integrity at each time point. Shown is the average from three biological replicates (±SD).

5.4 DISCUSSION

Viruses continue to evolve novel strategies to evade the host immune system. In recent years it has been demonstrated that non-enveloped viruses have the ability to co-opt cellular processes to acquire an envelope (Bird and Kirkegaard, 2015; Feng and Lemon, 2014). In this study, we utilized proteomic approaches to evaluate the differences in protein content between exosomes derived from mock- or CrPV-infected cells. In doing so, we identified 1408 and 802 proteins from the cell and exosome-like vesicles, respectively (Figure 5.3). Only 64% of the ELV-associated proteins overlapped with the cell pellet suggesting that there may be a distinct set of proteins enriched within ELVs compared to the cell pellet. Moreover, during CrPV infection, 40 proteins increased and 8 protein decreased in abundance compared to mock-infected cells (Figure 5.3). This suggests that there is an active mechanism for the sorting and secretion of proteins in these cells during infection. The best-described mechanism for the formation of exosomes in mammalian cells is through the endosomal sorting complex required for transport (ESCRT)(Colombo et al., 2014). This complex is composed of roughly thirty proteins that assemble into four complexes (ESCRT-0, -I, -II, and –III) along with other associated proteins. ESCRT-0 binds and sequesters ubiquitylated transmembrane proteins in the endosomal membrane. HRS (a component of ESCRT-0) then recruits TSG101 of the ESCRT-I complex. In tandem with ESCRT-II, ESCRT-I acts to cause endosomal membrane deformation. ESCRT-I is then involved in the recruitment of ESCRT-III via ESCRT-II or ALIX. Finally, ESCRT-III facilitates vesicle scission into the endosomal lumen (Colombo et al., 2014; Hurley and Hanson, 2010). The exact mechanism of how cytosolic proteins are sorted into exosomes is not well understood. It is thought that mono-ubquitinylated proteins containing ‘late’ domain with a sequence motif P(S/T)XP, $YPX_{1\text{-}3}L$, or PPXYY are targeted for delivery to MVBs by

interacting with proteins such as TSG101 or ALIX (Buschow et al., 2005; Feng and Lemon, 2014; Freed, 2002). Alternatively, a recent role for HSC70 has been proposed where it binds proteins containing a KFERQ sequence and to phosphatidylserine on the MVB outer membrane (Sahu et al., 2011). However, not all proteins in exosomes are ubquitinylated nor do they all contain KFERQ sequences. It appears a passive mechanism may be involved in protein sorting to MVBs that involves lipid (e.g. ceramide or cholesterol) and/or tetraspannin associations (Edgar et al., 2014; Stuffers et al., 2009; Trajkovic et al., 2008; van Niel et al., 2011). Currently there is limited knowledge on how CrPV alters host cell processes during infection and although numerous other viruses have been shown to exploit exosome biogenesis pathway including human immunodeficiency virus-1 (HIV-1), hepatitis B virus (HBV), and Epstein Barr virus (EBV), there remains no consensus view of a distinct mechanism that leads to changes in cargo content of exosomes (Raab-Traub and Dittmer, 2017).

From our proteomic analysis, the most highly enriched protein(s) in ELVs derived from CrPV-infected cells was the CrPV structural proteins. We identified 21 peptides that mapped across the CrPV ORF2 polyprotein resulting in a $\log_2$ fold change of >12 in ELVs (Figure 5.4). Interestingly, we were unable to identify any of the CrPV non-structural proteins encoded by ORF1 either by mass spectrometry or Western blot within the ELV fraction (Figure 5.2 and 5.3). Together with our density gradient analysis, viral titres, and membrane integrity assay (Figures 5.4 and 5.5), these results indicate that, like hepatitis A virus (HAV) and hepatitis E virus (HEV), CrPV may co-opt exosomes for its own advantage during infection. Akin to our observations with CrPV, recent work demonstrated that during HAV infection there arises two distinct populations of virus. Both populations remain infectious; however one acquires an envelope (eHAV) while the other remains non-enveloped (Feng et al., 2013). HAV encodes two tandem

'late' domains in its VP2 capsid protein that are proposed to facilitate interaction of the viral capsid with ALIX and be subsequently sorted into exosomes. Intriguingly, eHAV contains an unprocessed version of the viral capsid protein VP1, termed VP1pX, that is not present in the non-enveloped virus suggesting there could an additional role played by the pX extension to facilitate envelopment of HAV (Feng et al., 2013). Similarly, HEV also contains a viral protein (ORF3) that is only found in enveloped forms of the virus which directs interacts with TSG101 through a PSAP motif (Nagashima et al., 2011a; Nagashima et al., 2011b). Thus, it appears both HAV and HEV usurp the host ESCRT system to acquire an envelope. Whether CrPV follows a similar mechanism remains to be determined, but further studies investigating the role of the ESCRT complex and how viral proteins contribute to this process in CrPV infection would provide valuable insights into its mode of transmission.

Why would CrPV require an envelope for cell-to-cell transmission? With HAV, enveloped virions circulate in the blood of infected patients while non-enveloped virions are found in the feces (Feng et al., 2013). For decades virologists have understood that enveloped and non-enveloped viruses have distinct advantages over on another. Enveloped viruses can readily manipulate the contents of their cognate envelope allowing for flexibility in receptors, evasion of the host immune response, and shuttling viral proteins to enhance infection. Nevertheless, envelopes are typically susceptible to bile salts and retain lower stability in the environment (e.g. due to desiccation or chemical resistance)(Smith, 1939). Thus, a biphasic viral life cycle provides the security of an envelope with the robustness of a capsid. In the case of eHAV, the envelope is thought to cloak the virus from neutralizing antibodies in the bloodstream. Unlike humans, *Drosophila* lacks any sort of antibody defense to viruses; however, it is understood that hemocytes circulating in the hemolymph mediate phagocytosis of invading

pathogens to prevent infection (Stuart and Ezekowitz, 2008). Furthermore, *Drosophilia* cells secrete the Ig-domain containing protein, Dscam, that has been shown to have over 18,000 isoforms and may act in opsonizing pathogens for phagocytosis (Stuart and Ezekowitz, 2008; Watson et al., 2005). It is plausible that CrPV hijacks the exosomal pathway to shroud itself from the phagocytic machinery within the host organism.

Overall, our data offers the first look into how extracellular vesicles contribute to the pathogenesis of CrPV and how these viruses are transmitted between host cells. Moreover, it suggests that this virus can co-opt these vesicles for its own advantage to potentially evade the host immune response. Further research investigating the exact mechanisms contributing to this phenomenon will provide insight to the currently unknown life cycle of dicistroviruses.

Chapter 6: Summary and future directions

Positive single stranded RNA viruses exist as some of the most intriguing systems to study. These tiny complex biological assemblies subsist as some of the most detrimental pathogens to humans, as key influencers of ecosystems, as major players in agriculturally relevant crops and animals, and as model systems to understand fundamental biological processes. Despite their importance, a comprehensive understanding of how these viruses replicate remains poorly established. Included in this list are members of the *Dicistroviridae*; viruses with dicistronic genomes that infect arthropods. Achieving an understanding of dicistroviruses is an important step as they are thought to be ubiquitous in nature in addition to being industrially and medically relevant (Bonning and Miller, 2010; Culley et al., 2003; Shi et al., 2016; Victoria et al., 2009). Specifically, investigations into dicistroviruses could lead to potential antiviral therapeutics for agriculturally important arthropods (Cox-Foster et al., 2007a; Lightner, 1996; Lightner and Redman, 1998). Moreover, elements encoded within the genomes of these viruses have helped to unravel fundamental properties of cellular processes (e.g. translation, micro RNAs, and non-sense mediated decay)(Butcher and Jan, 2016; Isken et al., 2008; Petersen et al., 2006). Finally, with the emergence and spread of viruses transmitted through arthropod vectors (e.g. dengue virus and Zika virus), studying the host-virus interactions of insect-borne viruses has become even more pertinent. Since dicistroviruses, namely CrPV, infect the genetically amenable *Drosophila melanogaster*, they provide a unique model to discern the mechanisms governing viral replication in an insect host. To this end, my thesis has focused on the characterization of how the dicistrovirus CrPV usurps the host cell machinery to facilitate gene expression, viral transmission, and cause disease.

As mentioned, there remain many unresolved questions pertaining to the dicistrovirus field. This is, in part, due to the lack of necessary tools to study their biology. Much of the insights gained into the mechanisms of (+)ssRNAs viruses has come from the use of reverse genetic tools, such as infectious clones, whereby specific mutations can be made in the virus to discern the function of a given protein or RNA element. Chapter 2 follows the development of an infectious clone of CrPV, the first of the *Dicistroviridae*. Utilizing purified CrPV genomic RNA and a series of RT-PCR and cloning steps, we constructed two fully infectious clones, termed CrPV-2 and CrPV-3. *In vitro* synthesized RNA derived from these clones can replicate, support synthesis of all known viral proteins *in vitro* and *in vivo*, generate infectious virus when transfected into naïve cells, and produce virus that causes mortality in adult fruit flies. Each clone has distinct differences in its nucleotide sequence from that which was first published. This likely reflects the quasi-species nature of RNA viruses and may contribute to why an infectious clone for the *Dicistroviridae* could not be previously established. Interestingly, CrPV-2 harbours a 196 nt duplication in its 5'UTR. Our data demonstrate that the presence of this duplication attenuates CrPV, presumably due to the inability to package its RNA or release the RNA upon subsequent infection. Recent work elucidating that structure of the CrPV 5'UTR showed that it is highly structured with two domains (Gross et al., 2017). Domain II encompasses the IRES element that drives translation of the downstream ORF, while Domain I, and its six stem-loops, is dispensable for IRES activity. Although the function of this domain is unclear, the CrPV-2 duplication almost entirely copies SL3 of Domain I. Many viruses contain RNA structures in their 5'UTR that are utilized in RNA replication or as RNA packaging signals (Liu et al., 2009; Patel et al., 2015). For CrPV-2, the 5'UTR duplication did not appear to affect viral RNA replication; thus, it is plausible that the RNA structures found in the 5'UTR are involved in

packaging. It would be valuable to perform pull-down experiments on biotinylated 5'UTR RNA, with or without Domain I, followed by label-free quantitative LC-MS/MS comparing mock- and infected-cells at different time points. Specific protein targets could then be validated using individual-nucleotide resolution crosslinking affinity purification (iCLAP)(Li et al., 2014a). To elaborate, ribonucleoprotein complexes are isolated from transfected or infected cells, crosslinked via UV light, and immuno-precipitated (IP) using an antibody towards the tagged protein of interest. The protein is then degraded and the RNA is sequenced to determine the precise binding site of the target protein to the RNA. iCLAP experiments in conjunction with dsRNA knockdowns and mutations in the infectious clone would prove valuable in understanding the dynamics of RNA:protein interactions during CrPV infections and would shed further light into how dicistroviruses replicate.

Studies utilizing the CrPV infectious clone are not restricted to the 5'UTR. This tool creates possibilities to study a plethora of viral-host interactions of dicistroviruses in *Drosophila*. Some of the remaining major questions include the function of the 3'UTR, IGR IRES, and viral proteins *in vivo*. The 3'UTR plays a crucial role during RNA replication in multiple RNA viruses. For example, the 3'UTR in HCV aids in circularization of the viral RNA and acts retain ribosome complexes at the termination of translation to ensure efficient initiation of subsequent rounds of translation (Bai et al., 2013). Addressing the role of the CrPV 3'UTR in infection could be done in a similar fashion as the 5'UTR; application of iCLAP, dsRNA knockdowns, and mutational analysis in the infectious clone. Moreover, the infectious clone has been used to construct a functional replicon of CrPV where the structural proteins encoded in ORF2 were replaced by a Fluc gene (Khong et al., 2016). Introduction of a mutation in the viral RdRp uncouples RNA replication and translation in this system; this allows for the contributions of

various viral components to these processes to be determined. A similar system has been used previously to investigate the contributions of the 3'UTR to dengue virus replication (Alvarez et al., 2005).

Although the roles of few CrPV proteins are understood (e.g. 3C protease and RdRp), the function of most proteins is not well defined. Specifically, the role of proteins 1A, 2B, and 3A. CrPV 1A is a multifunctional protein involved in the shutdown of the host RNAi via binding AGO2 and inhibiting the formation of stress granules (SG) during viral infection (Khong et al., 2017; Nayak et al., 2010b). To determine the function of 1A in SG disruption, mutations were generated in its C-terminus within the infectious clone which led to a reduction in viral loads during infection. Interestingly, 1A was observed to stimulate translation from the 5'UTR of CrPV, suggesting that there are additional roles played by this small protein (Khong et al., 2017). Unlike 1A, CrPV proteins 2B and 3A are completely uncharacterized. In poliovirus, 2B acts as a pore-forming channel to increase host cell membrane permeability to small solutes and ions during the late phase of infection (Nieva et al., 2012). Meanwhile, poliovirus 3A aids in RNA replication to anchor the viral RdRp and RNA to replication vesicles (Jackson, 2014; Xiang et al., 1995). Whether CrPV 2B and 3A have similar roles as the poliovirus counterparts is unknown; however, insertion of epitope tags or fused fluorescent proteins into the infectious clone would allow for tracking via fluorescence microscopy and give a better understanding of the localization and potential function of these proteins during CrPV infection. Conversely, the function of CrPV proteins can be further defined by examining their cognate protein-protein interaction network utilizing techniques such as BioID (Roux et al., 2013). Briefly, a promiscuous biotin ligase is fused to a protein of interest and upon addition of biotin any interacting partners, or proteins in close proximity, will be biotinylated. Putative interacting

partners are then analyzed via LC-MS/MS. Identifying interacting partners of each viral protein would help uncover the breadth of pathways circumvented during CrPV replication.

Apart from FMDV most picornaviruses encode only a single copy of VPg. By contrast, CrPV encodes four copies of VPg each of which has slight amino acid differences (Nakashima and Shibuya, 2006). In picornaviruses VPg is translated in equimolar amounts to all other viral proteins due to translation occurring as a single polyprotein. As CrPV proteins are encoded in two differentially regulated ORFs, it may be that these viruses have evolved multiple VPg copies to ensure a correct VPg to genomic RNA stoichiometric ratio for an optimal viral life cycle. In FMDV, all naturally isolated strains encode three copies of VPg (VPg_1, VPg_2, VPg_3). Interestingly, not every copy of VPg is required for infectivity, suggesting a strong selective pressure to maintain them (Pacheco et al., 2003). All three VPg proteins can be uridylyated, however, only deletion of VPg_3 abolishes FMDV RNA replication while deletion of VPg_1 or VPg_2 merely reduce it (Falk et al., 1992; Nayak et al., 2005; Pacheco et al., 2010). Deletions, duplications, or mutational experiments of the CrPV VPg sequences in the infectious clone would help to determine the function of each copy and why these viruses evolved this unique mechanism.

Considering the development of the infectious clone, in Chapter 3 we utilized it to investigate the role of the CrPV IGR IRES during viral infection. Previous studies have provided detailed structural, biochemical, and biophysical information into IGR IRES translation. However, these studies have been limited to *in vitro* approaches and little is known about how these IRESs behave *in vivo*. By introducing mutations known to perturb the IGR IRES into the infectious clone we established the physiological relevance of various structural regions during infection. As expected, a subset of the mutations failed to produce viable virus (e.g. ΔPKI and

mL1.1a). Intriguingly, mutations in SLIV and SLV that were previously shown to abolish IRES activity *in vitro* still supported virus production *in vivo*. While this study provides a complement to previous *in vitro* work, why these mutations remain functional during infection is unclear. Investigations into the overall structure of the IGR IRES using *in vivo* SHAPE could perhaps reveal dynamic structural conformations that this IRES adopts. Additionally, structural studies employing cryo-EM on *ex-vivo* wild-type or mutated IRES-80S ribosome complexes isolated from infected cells would expose *in vivo* functional states adopted by the IGR IRES. A similar approach has been performed on translating human ribosomes to understand the dynamics that occur during canonical eukaryotic translation (Behrmann et al., 2015).

In vivo IRES studies are not limited to the CrPV IGR IRES. Recently the IAPV IGR IRES was introduced in lieu of the CrPV IGR IRES in the infectious clone to generate a chimeric virus. As there is no current means to study IAPV in its native context, the CrPV/IAPV chimera was utilized to understand the role of SLIV in IRES-mediated translation (Au et al., 2017). It would be of interest to examine if other elements within dicistroviruses (or other viral families) can readily compensate for one another. For example, swapping the CrPV structural proteins for those of DCV or IAPV may reveal insights into determinants of host and tissue tropisms. Overall, these examples and those described above highlight the power of the CrPV infectious clone as a reverse genetic tool for understanding RNA and protein functional relationships in dicistroviruses.

In Chapter 4 we explored the unusual mechanism of CrPV IGR IRES-mediated +1-frame translation. Previous work in our lab demonstrated that a subset of dicistroviruses IGR IRESs have the capacity for +1-frame translation of a hidden overlapping open reading frame, ORFx (Ren et al., 2012). These viruses appear to utilize a U:G wobble base pair adjacent to PKI to set

the ribosome into the +1 reading frame. CrPV, on the other hand, uses a unique mechanism where it initiates at a downstream lysine codon. Extensive mutational analysis and biochemical assays established the following rules for +1-frame translation: (1) translation is IGR IRES-dependent; (2) it requires pseudotranslocation of the IRES; (3) it initiates at the 13th AAA (Lys) codon downstream; (4) the spacer region is necessary; and (5) the nucleotide identity is crucial for efficient +1-frame translation. Although the exact mechanism is elusive, we propose that the IRES employs a translational bypass-like mechanism to facilitate +1-frame translation. Truncation mutants suggested that the translation of the first few codons adjacent to PKI are required for +1-translation. Here, we envision that the 80S ribosome engages the IRES and begins translation; following delivery of the first few aminoacyl-tRNAs there is a pausing event. From here, much like with bacteriophage T4 *gene 60* bypassing, the IRES may block back-pedaling of the ribosome that may force it to disengage from the RNA and resume translation downstream (Chen et al., 2015a). Interestingly, recent data demonstrated that delivery of the third aminoacyl-tRNA is a kinetically 'slow' step which may represent the necessary pause for the ribosome to skip downstream (Zhang et al., 2016). However, convoluting this model is that the introduction of stop codons in the 0- or +1-frames adjacent to PKI do not disrupt +1-frame translation. Additional experiments using single molecule fluorescence resonance energy transfer with labeled IRES and ribosome would likely illuminate the structural rearrangements occurring during +1-frame translation. Furthermore, verification of the ORFx N-terminus via Edman degradation and reading frame assessment using ribosomal profiling would solidify the translation start site and help prove the precise mechanism.

An outstanding question is simply: what is the role of ORFx? In the latter half of Chapter 4 we begin to address this question by examining the contribution of ORFx to CrPV infection.

Again, with the CrPV infectious clone, we generated viruses containing synonymous +1-frame stop codons at either the 12th (CrPV-S12) or 19th (CrPV-S19) position that presumably express or do not express ORFx, respectively. No differences in were observed in viral loads when infecting adult flies or tissue culture cells. The sole difference seen was an attenuation of fly morality in viruses unable to produce ORFx. This suggests that ORFx contributes to pathogenesis of CrPV in adult flies. Additionally, *in silico* predictions, sub-cellular fractionation, and immunofluorescence data of the 41 amino acid form of ORFx leads to the conclusion that it resides as a membrane protein in the ER. Intriguingly, *Drosophila* and humans encode membrane inserted micropeptides named sarcolamban and myoregulin, respectively (Anderson et al., 2015; Magny et al., 2013). These micropeptides act to bind the protein SERCA (sarcoplasmic reticulum Ca^{+2}-ATPase; Ca-P60A in *Drosophila*) in the lipid bilayer, inhibiting its function. SERCA functions as the main ATP-dependent transporter to pump Ca^{+2} into the sarcoplasmic reticulum resulting in muscle relaxation (Anderson et al., 2015; Magny et al., 2013). Since CrPV infection leads to paralysis and eventual death of its host, it is tempting to speculate that ORFx may bind to SERCA in the membrane, inhibiting its function. Behavioral experiments on fruit flies infected with CrPV-S12 or CrPV-S19, such as climbing assays, would help support this model. Moreover, IP-MS of tagged ORFx would identify potential interacting partners and give a better understanding of the function of ORFx. Care needs to be taken when performing the above experiment as false positives could arise due to the membrane associated nature of ORFx, thus rigorous validation must be done. Overall, this work demonstrates a novel programmed IRES-mediated recoding strategy to increase viral coding capacity and impact viral replication.

Chapter 5 investigates a different aspect of the CrPV life cycle: release and transmission of progeny virions. Traditionally thought to only escape the host cell via lysis, we have demonstrated that CrPV may also hijack the exosomal pathway to facilitate viral release. It is not outside the realm of possibility that this occurs. Mammalian viruses have been shown to usurp this pathway for budding and envelopment (Feng et al., 2013; Fujii et al., 2007). As well, recent evidence has shown that *Drosophila* cells not only secrete exosomes, but also exploit them for a systemic anti-viral response (Tassetto et al., 2017). Circulating hemocytes take up viral dsRNA and virus, which is subsequently reverse transcribed into viral DNA molecules. These DNA molecules are then used for *de novo* synthesis of secondary virus-derived siRNAs (vsRNAs). vsRNAs are then incorporated into exosome-like vesicles for delivery to adjacent cells (Tassetto et al., 2017). Since *Drosophila* has evolved this pathway for its own use, it is conceivable that a viral agent could co-opt it. Although, how CrPV manages to commandeer the machinery from this pathway for its own benefit is unresolved. With HAV, there are late domains within the capsid proteins that facilitate envelopment of the virus. Indeed, brief inspection of the capsid proteins of CrPV shows some putative binding sites for ALIX and TSG101. Follow-up experiments mutating these sites in the infectious clone of CrPV would provide insight into the mechanism of envelopment. Conversely, knockdown experiments using dsRNA targeting ESCRT components would additionally support this model.

As discussed previously, CVB3 and poliovirus are released *en bloc* in phosphatidylserine vesicles to facilitate infection (Chen et al., 2015b). It would be of interest to determine if CrPV utilizes a similar mechanism of release or if it tows closer to the lines of HAV, pointing towards the evolutionary path of these viruses. Experiments implementing live cell imaging and single molecule fluorescent *in situ* hybridization against the viral RNA could help address these

questions. Additionally, a major remaining inquiry is why CrPV would evolve to acquire an envelope. In Chapter 5, we propose that it may do so to evade DsCam and the phagocytic machinery of the host organism. Previous studies have demonstrated that depletion of phagocytic cells by treatment of adult *Drosohpila* flies with polystyrene beads enhances CrPV infection (Costa et al., 2009). Additional experiments such as infecting flies deficient in DsCam and determining if hemolymph from infected flies reacts with enveloped vs non-enveloped CrPV (i.e. probe for DsCam bound to the virus), akin to serum from infected patients and HAV, would help establish if this pathway is involved in the response to CrPV infection.

In conclusion, the work presented in this thesis has helped to provide a deeper understanding of the CrPV life cycle and its host interactions. The establishment of a dicistrovirus infectious clone provides a foundation for future studies investigating the fundamentals of both dicistroviruses and biological processes. An exciting aspect of modern virology is that ongoing studies enrich our understanding of individual viruses, while also revealing unifying principles that link many aspects of virus infection, replication, and host-virus interactions across many viral families. Insights derived from virological studies put forth a central order within the vast and seemingly chaotic diversity of known viruses, with significant implications in viral evolution and function. As the full extent of underlying parallels are still developing, continuing research will extend and refine the core connections between diverse viral families. Indeed, the resulting information generated from these studies will continue to have important ramifications in our abilities to combat the current and emerging threat of viral disease while furthering development of the beneficial uses of viruses.

REFERENCES

Abaeva, I.S., Pestova, T.V., Hellen, C.U., 2016. Attachment of ribosomal complexes and retrograde scanning during initiation on the Halastavi arva virus IRES. Nucleic Acids Res 44, 2362-2377.

Abeyrathne, P.D., Koh, C.S., Grant, T., Grigorieff, N., Korostelev, A.A., 2016. Ensemble cryo-EM uncovers inchworm-like translocation of a viral IRES through the ribosome. eLife 5.

Adams, M.J., Lefkowitz, E.J., King, A.M.Q., Harrach, B., Harrison, R.L., Knowles, N.J., Kropinski, A.M., Krupovic, M., Kuhn, J.H., Mushegian, A.R., Nibert, M., Sabanadzovic, S., Sanfacon, H., Siddell, S.G., Simmonds, P., Varsani, A., Zerbini, F.M., Gorbalenya, A.E., Davison, A.J., 2017. Changes to taxonomy and the International Code of Virus Classification and Nomenclature ratified by the International Committee on Taxonomy of Viruses (2017). Archives of Virology 162, 2505-2538.

Adhin, M.R., van Duin, J., 1990. Scanning model for translational reinitiation in eubacteria. J Mol Biol 213, 811-818.

Agirre, A., Barco, A., Carrasco, L., Nieva, J.L., 2002. Viroporin-mediated membrane permeabilization. Pore formation by nonstructural poliovirus 2B protein. J Biol Chem 277, 40434-40441.

Alexopoulou, L., Holt, A.C., Medzhitov, R., Flavell, R.A., 2001. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. Nature 413, 732-738.

Alvarez, D.E., De Lella Ezcurra, A.L., Fucito, S., Gamarnik, A.V., 2005. Role of RNA structures present at the 3'UTR of dengue virus on translation, RNA synthesis, and viral replication. Virology 339, 200-212.

Anderson, D.M., Anderson, K.M., Chang, C.L., Makarewich, C.A., Nelson, B.R., McAnally, J.R., Kasaragod, P., Shelton, J.M., Liou, J., Bassel-Duby, R., Olson, E.N., 2015. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. Cell 160, 595-606.

Andreev, D.E., Hirnet, J., Terenin, I.M., Dmitriev, S.E., Niepmann, M., Shatsky, I.N., 2012. Glycyl-tRNA synthetase specifically binds to the poliovirus IRES to activate translation initiation. Nucleic Acids Res 40, 5602-5614.

Au, H.H., Cornilescu, G., Mouzakis, K.D., Ren, Q., Burke, J.E., Lee, S., Butcher, S.E., Jan, E., 2015. Global shape mimicry of tRNA within a viral internal ribosome entry site mediates translational reading frame selection. Proc Natl Acad Sci U S A 112, E6446-6455.

Au, H.H., Jan, E., 2012. Insights into factorless translational initiation by the tRNA-like pseudoknot domain of a viral IRES. PLoS One 7, e51477.

Au, H.H., Jan, E., 2014. Novel viral translation strategies. Wiley interdisciplinary reviews. RNA 5, 779-801.

Au, H.H.T., Elspass, V.M., Jan, E., 2017. Functional insights into the adjacent stem-loop in honey bee dicistroviruses that promotes IRES-mediated translation and viral infection. J Virol.

Aylett, C.H., Boehringer, D., Erzberger, J.P., Schaefer, T., Ban, N., 2015. Structure of a yeast 40S-eIF1-eIF1A-eIF3-eIF3j initiation complex. Nat Struct Mol Biol 22, 269-271.

Bai, Y., Zhou, K., Doudna, J.A., 2013. Hepatitis C virus 3'UTR regulates viral translation through direct interactions with the host translation machinery. Nucleic Acids Res 41, 7861-7874.

Baltimore, D., 1971. Expression of animal virus genomes. Bacteriol Rev 35, 235-241.

Ban, N., Beckmann, R., Cate, J.H., Dinman, J.D., Dragon, F., Ellis, S.R., Lafontaine, D.L., Lindahl, L., Liljas, A., Lipton, J.M., McAlear, M.A., Moore, P.B., Noller, H.F., Ortega, J., Panse, V.G., Ramakrishnan, V., Spahn, C.M., Steitz, T.A., Tchorzewski, M., Tollervey, D., Warren, A.J., Williamson, J.R., Wilson, D., Yonath, A., Yusupov, M., 2014. A new system for naming ribosomal proteins. Curr Opin Struct Biol 24, 165-169.

Baranov, P.V., Atkins, J.F., Yordanova, M.M., 2015. Augmented genetic decoding: global, local and temporal alterations of decoding processes and codon meaning. Nature reviews. Genetics 16, 517-529.

Barton, D.J., Black, E.P., Flanegan, J.B., 1995. Complete replication of poliovirus in vitro: preinitiation RNA replication complexes require soluble cellular factors for the synthesis of VPg-linked RNA. J Virol 69, 5516-5527.

Bawden, F.C., Pirie, N.W., 1937. The isolation and some properties of liquid crystalline substances from solanaceous plants infected with three strains of tobacco mosaic virus. Proc R Soc Ser B-Bio 123, 274-320.

Beckett, K., Monier, S., Palmer, L., Alexandre, C., Green, H., Bonneil, E., Raposo, G., Thibault, P., Le Borgne, R., Vincent, J.P., 2013. Drosophila S2 cells secrete wingless on exosome-like vesicles but the wingless gradient forms independently of exosomes. Traffic 14, 82-96.

Behrmann, E., Loerke, J., Budkevich, T.V., Yamamoto, K., Schmidt, A., Penczek, P.A., Vos, M.R., Burger, J., Mielke, T., Scheerer, P., Spahn, C.M., 2015. Structural snapshots of actively translating human ribosomes. Cell 161, 845-857.

Belcourt, M.F., Farabaugh, P.J., 1990. Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. Cell 62, 339-352.

Belov, G.A., Altan-Bonnet, N., Kovtunovych, G., Jackson, C.L., Lippincott-Schwartz, J., Ehrenfeld, E., 2007. Hijacking components of the cellular secretory pathway for replication of poliovirus RNA. J Virol 81, 558-567.

Belov, G.A., Evstafieva, A.G., Rubtsov, Y.P., Mikitas, O.V., Vartapetian, A.B., Agol, V.I., 2000. Early alteration of nucleocytoplasmic traffic induced by some RNA viruses. Virology 275, 244-248.

Benjeddou, M., Leat, N., Allsopp, M., Davison, S., 2002. Development of infectious transcripts and genome manipulation of Black queen-cell virus of honey bees. J Gen Virol 83, 3139-3146.

Bird, S.W., Kirkegaard, K., 2015. Escape of non-enveloped virus from intact cells. Virology 479-480, 444-449.

Bird, S.W., Maynard, N.D., Covert, M.W., Kirkegaard, K., 2014. Nonlytic viral spread enhanced by autophagy components. Proc Natl Acad Sci U S A 111, 13081-13086.

Bonami, J.R., Hasson, K.W., Mari, J., Poulos, B.T., Lightner, D.V., 1997. Taura syndrome of marine penaeid shrimp: characterization of the viral agent. J Gen Virol 78 (Pt 2), 313-319.

Bonning, B.C., Miller, W.A., 2010. Dicistroviruses. Annu Rev Entomol 55, 129-150.

Boyapalle, S., Beckett, R.J., Pal, N., Miller, W.A., Bonning, B.C., 2008. Infectious genomic RNA of Rhopalosiphum padi virus transcribed in vitro from a full-length cDNA clone. Virology 375, 401-411.

Boyer, J.C., Haenni, A.L., 1994. Infectious transcripts and cDNA clones of RNA viruses. Virology 198, 415-426.

Brasey, A., Lopez-Lastra, M., Ohlmann, T., Beerens, N., Berkhout, B., Darlix, J.L., Sonenberg, N., 2003. The leader of human immunodeficiency virus type 1 genomic RNA harbors an internal ribosome entry segment that is active during the G2/M phase of the cell cycle. J Virol 77, 3939-3949.

Brierley, I., Dos Ramos, F.J., 2006. Programmed ribosomal frameshifting in HIV-1 and the SARS-CoV. Virus Res 119, 29-42.

Brierley, I., Jenner, A.J., Inglis, S.C., 1992. Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. J Mol Biol 227, 463-479.

Buschow, S.I., Liefhebber, J.M., Wubbolts, R., Stoorvogel, W., 2005. Exosomes contain ubiquitinated proteins. Blood Cells Mol Dis 35, 398-403.

Butcher, S.E., Jan, E., 2016. tRNA-mimicry in IRES-mediated translation and recoding. RNA Biol 13, 1068-1074.

Carillo-Tripp, J., Bonning, B.C., Miller, W.A., 2015. Challenges associated with research on RNA viruses of insects. Current Opinion in Insect Science 6, 1-7.

Carrasco, L., 1995. Modification of membrane permeability by animal viruses. Advances in virus research 45, 61-112.

Castano, A., Ruiz, L., Hernandez, C., 2009. Insights into the translational regulation of biologically active open reading frames of Pelargonium line pattern virus. Virology 386, 417-426.

Cevallos, R.C., Sarnow, P., 2005. Factor-independent assembly of elongation-competent ribosomes by an internal ribosome entry site located in an RNA virus that infects penaeid shrimp. J Virol 79, 677-683.

Chen, C.Y., Sarnow, P., 1995. Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs. Science 268, 415-417.

Chen, J., Coakley, A., O'Connor, M., Petrov, A., O'Leary, S.E., Atkins, J.F., Puglisi, J.D., 2015a. Coupling of mRNA Structure Rearrangement to Ribosome Movement during Bypassing of Non-coding Regions. Cell 163, 1267-1280.

Chen, Y.H., Du, W., Hagemeijer, M.C., Takvorian, P.M., Pau, C., Cali, A., Brantner, C.A., Stempinski, E.S., Connelly, P.S., Ma, H.C., Jiang, P., Wimmer, E., Altan-Bonnet, G., Altan-Bonnet, N., 2015b. Phosphatidylserine vesicles enable efficient en bloc transmission of enteroviruses. Cell 160, 619-630.

Cherry, S., Doukas, T., Armknecht, S., Whelan, S., Wang, H., Sarnow, P., Perrimon, N., 2005. Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition. Genes Dev 19, 445-452.

Cherry, S., Kunte, A., Wang, H., Coyne, C., Rawson, R.B., Perrimon, N., 2006. COPI activity coupled with fatty acid biosynthesis is required for viral replication. PLoS Pathog 2, e102.

Cherry, S., Perrimon, N., 2004. Entry is a rate-limiting step for viral infection in a Drosophila melanogaster model of pathogenesis. Nature immunology 5, 81-87.

Cho, M.W., Teterina, N., Egger, D., Bienz, K., Ehrenfeld, E., 1994. Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells. Virology 202, 129-145.

Christian, P.D., Scotti, P.D., 1998. Picornalike viruses of insects, The insect viruses. Springer, pp. 301-336.

Chtarbanova, S., Lamiable, O., Lee, K.Z., Galiana, D., Troxler, L., Meignin, C., Hetru, C., Hoffmann, J.A., Daeffler, L., Imler, J.L., 2014. Drosophila C virus systemic infection leads to intestinal obstruction. J Virol 88, 14057-14069.

Colbere-Garapin, F., Christodoulou, C., Crainic, R., Pelletier, I., 1989. Persistent poliovirus infection of human neuroblastoma cells. Proc Natl Acad Sci U S A 86, 7590-7594.

Colombo, M., Raposo, G., Thery, C., 2014. Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. Annu Rev Cell Dev Biol 30, 255-289.

Colussi, T.M., Costantino, D.A., Zhu, J., Donohue, J.P., Korostelev, A.A., Jaafar, Z.A., Plank, T.D., Noller, H.F., Kieft, J.S., 2015. Initiation of translation in bacteria by a structured eukaryotic IRES RNA. Nature 519, 110-113.

Costa, A., Jan, E., Sarnow, P., Schneider, D., 2009. The Imd Pathway Is Involved in Antiviral Immune Responses in Drosophila. PLoS ONE 4, e7436.

Costantino, D.A., Pfingsten, J.S., Rambo, R.P., Kieft, J.S., 2007. tRNA–mRNA mimicry drives translation initiation from a viral IRES. Nature Structural & Molecular Biology 15, 57-64.

Costantino, D.A., Pfingsten, J.S., Rambo, R.P., Kieft, J.S., 2008. tRNA-mRNA mimicry drives translation initiation from a viral IRES. Nat Struct Mol Biol 15, 57-64.

Coulibaly, F., Chiu, E., Ikeda, K., Gutmann, S., Haebel, P.W., Schulze-Briese, C., Mori, H., Metcalf, P., 2007. The molecular organization of cypovirus polyhedra. Nature 446, 97-101.

Cox, J., Mann, M., 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nature biotechnology 26, 1367-1372.

Cox-Foster, D.L., Conlan, S., Holmes, E.C., Palacios, G., Evans, J.D., Moran, N.A., Quan, P.L., Briese, T., Hornig, M., Geiser, D.M., Martinson, V., vanEngelsdorp, D., Kalkstein, A.L., Drysdale, A., Hui, J., Zhai, J., Cui, L., Hutchison, S.K., Simons, J.F., Egholm, M., Pettis, J.S., Lipkin, W.I., 2007a. A metagenomic survey of microbes in honey bee colony collapse disorder. Science 318, 283-287.

Cox-Foster, D.L., Conlan, S., Holmes, E.C., Palacios, G., Evans, J.D., Moran, N.A., Quan, P.L., Briese, T., Hornig, M., Geiser, D.M., Martinson, V., vanEngelsdorp, D., Kalkstein, A.L., Drysdale, A., Hui, J., Zhai, J., Cui, L., Hutchison, S.K., Simons, J.F., Egholm, M., Pettis, J.S., Lipkin, W.I., 2007b. A Metagenomic Survey of Microbes in Honey Bee Colony Collapse Disorder. Science 318, 283-287.

Culley, A.I., Lang, A.S., Suttle, C.A., 2003. High diversity of unknown picorna-like viruses in the sea. Nature 424, 1054-1057.

Danthi, P., Tosteson, M., Li, Q.H., Chow, M., 2003. Genome delivery and ion channel properties are altered in VP4 mutants of poliovirus. J Virol 77, 5266-5274.

de Breyne, S., Yu, Y., Unbehaun, A., Pestova, T.V., Hellen, C.U., 2009. Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites. Proc Natl Acad Sci U S A 106, 9197-9202.

de Miranda, J.R., Cordoni, G., Budge, G., 2010. The Acute bee paralysis virus–Kashmir bee virus–Israeli acute paralysis virus complex. Journal of Invertebrate Pathology 103, S30-S47.

De Sena, J., Mandel, B., 1977. Studies on the in vitro uncoating of poliovirus. II. Characteristics of the membrane-modified particle. Virology 78, 554-566.

Dinman, J.D., 2012. Mechanisms and implications of programmed translational frameshifting. Wiley interdisciplinary reviews. RNA 3, 661-673.

Dinos, G., Wilson, D.N., Teraoka, Y., Szaflarski, W., Fucini, P., Kalpaxis, D., Nierhaus, K.H., 2004. Dissecting the ribosomal inhibition mechanisms of edeine and pactamycin: the universally conserved residues G693 and C795 regulate P-site RNA binding. Mol Cell 13, 113-124.

Edery, I., Sonenberg, N., 1985. Cap-dependent RNA splicing in a HeLa nuclear extract. Proc Natl Acad Sci U S A 82, 7590-7594.

Edgar, J.R., Eden, E.R., Futter, C.E., 2014. Hrs- and CD63-dependent competing mechanisms make different sized endosomal intraluminal vesicles. Traffic 15, 197-211.

Etchison, D., Milburn, S.C., Edery, I., Sonenberg, N., Hershey, J.W., 1982. Inhibition of HeLa cell protein synthesis following poliovirus infection correlates with the proteolysis of a 220,000-dalton polypeptide associated with eucaryotic initiation factor 3 and a cap binding protein complex. J Biol Chem 257, 14806-14810.

Falk, M.M., Sobrino, F., Beck, E., 1992. VPg gene amplification correlates with infective particle formation in foot-and-mouth disease virus. J Virol 66, 2251-2260.

Fang, Y., Treffers, E.E., Li, Y., Tas, A., Sun, Z., van der Meer, Y., de Ru, A.H., van Veelen, P.A., Atkins, J.F., Snijder, E.J., Firth, A.E., 2012. Efficient -2 frameshifting by mammalian ribosomes to synthesize an additional arterivirus protein. Proc Natl Acad Sci U S A 109, E2920-2928.

Farabaugh, P.J., Zhao, H., Vimaladithan, A., 1993. A novel programed frameshift expresses the POL3 gene of retrotransposon Ty3 of yeast: frameshifting without tRNA slippage. Cell 74, 93-103.

Feng, Z., Hensley, L., McKnight, K.L., Hu, F., Madden, V., Ping, L., Jeong, S.H., Walker, C., Lanford, R.E., Lemon, S.M., 2013. A pathogenic picornavirus acquires an envelope by hijacking cellular membranes. Nature 496, 367-371.

Feng, Z., Lemon, S.M., 2014. Peek-a-boo: membrane hijacking and the pathogenesis of viral hepatitis. Trends in microbiology 22, 59-64.

Fernandez, I.S., Bai, X.C., Murshudov, G., Scheres, S.H., Ramakrishnan, V., 2014. Initiation of Translation by Cricket Paralysis Virus IRES Requires Its Translocation in the Ribosome. Cell.

Fernandez, J., Yaman, I., Sarnow, P., Snider, M.D., Hatzoglou, M., 2002. Regulation of internal ribosomal entry site-mediated translation by phosphorylation of the translation initiation factor eIF2alpha. J Biol Chem 277, 19198-19205.

Firth, A.E., Brierley, I., 2012. Non-canonical translation in RNA viruses. J Gen Virol 93, 1385-1409.

Firth, A.E., Wang, Q.S., Jan, E., Atkins, J.F., 2009. Bioinformatic evidence for a stem-loop structure 5'-adjacent to the IGR-IRES and for an overlapping gene in the bee paralysis dicistroviruses. Virol J 6, 193.

Flanegan, J.B., Petterson, R.F., Ambros, V., Hewlett, N.J., Baltimore, D., 1977. Covalent linkage of a protein to a defined nucleotide sequence at the 5'-terminus of virion and replicative intermediate RNAs of poliovirus. Proc Natl Acad Sci U S A 74, 961-965.

Flint, S.J., American Society for Microbiology., 2009. Principles of virology, 3rd ed. ASM Press, Washington, DC.

Flynt, A., Liu, N., Martin, R., Lai, E.C., 2009. Dicing of viral replication intermediates during silencing of latent Drosophila viruses. Proc Natl Acad Sci U S A 106, 5270-5275.

Fraenkel-Conrat, H., Singer, B., 1999. Virus reconstitution and the proof of the existence of genomic RNA. Philos T Roy Soc B 354, 583-586.

Freed, E.O., 2002. Viral late domains. J Virol 76, 4679-4687.

Fricks, C.E., Hogle, J.M., 1990. Cell-induced conformational change in poliovirus: externalization of the amino terminus of VP1 is responsible for liposome binding. J Virol 64, 1934-1945.

Fujii, K., Hurley, J.H., Freed, E.O., 2007. Beyond Tsg101: the role of Alix in 'ESCRTing' HIV-1. Nature reviews. Microbiology 5, 912-916.

Gallai, N., Salles, J.-M., Settele, J., Vaissière, B.E., 2009. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. Ecological Economics 68, 810-821.

Garrey, J.L., Lee, Y.Y., Au, H.H., Bushell, M., Jan, E., 2010. Host and viral translational mechanisms during cricket paralysis virus infection. J Virol 84, 1124-1138.

Gillis, J., Mistry, M., Pavlidis, P., 2010. Gene function analysis in complex data sets using ErmineJ. Nature protocols 5, 1148-1159.

Graff, J., Torian, U., Nguyen, H., Emerson, S.U., 2006. A bicistronic subgenomic mRNA encodes both the ORF2 and ORF3 proteins of hepatitis E virus. J Virol 80, 5919-5926.

Groppelli, E., Belsham, G.J., Roberts, L.O., 2007. Identification of minimal sequences of the Rhopalosiphum padi virus 5' untranslated region required for internal initiation of protein synthesis in mammalian, plant and insect translation systems. J Gen Virol 88, 1583-1588.

Gross, L., Vicens, Q., Einhorn, E., Noireterre, A., Schaeffer, L., Kuhn, L., Imler, J.L., Eriani, G., Meignin, C., Martin, F., 2017. The IRES5'UTR of the dicistrovirus cricket paralysis virus is a type III IRES containing an essential pseudoknot structure. Nucleic Acids Res 45, 8993-9004.

Gunji, T., N. Kato, M. Hijikata, K. Hayashi, S. Saitoh, and K. Shimotohno, 1993. Specific detection of positive and negative stranded hepatitis C viral RNA using chemical RNA modification. Archives of Virology 134, 293-302.

Guo, L., Allen, E.M., Miller, W.A., 2001. Base-pairing between untranslated regions facilitates translation of uncapped, nonpolyadenylated viral RNA. Mol Cell 7, 1103-1109.

Gustin, K.E., Sarnow, P., 2001. Effects of poliovirus infection on nucleo-cytoplasmic trafficking and nuclear pore complex composition. EMBO J 20, 240-249.

Hanecak, R., Semler, B.L., Anderson, C.W., Wimmer, E., 1982. Proteolytic processing of poliovirus polypeptides: antibodies to polypeptide P3-7c inhibit cleavage at glutamine-glycine pairs. Proc Natl Acad Sci U S A 79, 3973-3977.

Hannun, Y.A., Obeid, L.M., 1995. Ceramide: an intracellular signal for apoptosis. Trends Biochem Sci 20, 73-77.

Harutyunyan, S., Kumar, M., Sedivy, A., Subirats, X., Kowalski, H., Kohler, G., Blaas, D., 2013. Viral uncoating is directional: exit of the genomic RNA in a common cold virus starts with the poly-(A) tail at the 3'-end. PLoS Pathog 9, e1003270.

Hashimoto, Y., Valles, S.M., 2007. Solenopsis invicta virus-1 tissue tropism and intra-colony infection rate in the red imported fire ant: a quantitative PCR-based study. J Invertebr Pathol 96, 156-161.

Hatakeyama, Y., Shibuya, N., Nishiyama, T., Nakashima, N., 2004. Structural variant of the intergenic internal ribosome entry site elements in dicistroviruses and computational search for their counterparts. RNA 10, 779-786.

Henne, W.M., Buchkovich, N.J., Emr, S.D., 2011. The ESCRT pathway. Dev Cell 21, 77-91.

Herr, A.J., Atkins, J.F., Gesteland, R.F., 1999. Mutations which alter the elbow region of tRNA2Gly reduce T4 gene 60 translational bypassing efficiency. EMBO J 18, 2886-2896.

Herr, A.J., Gesteland, R.F., Atkins, J.F., 2000. One protein from two open reading frames: mechanism of a 50 nt translational bypass. EMBO J 19, 2671-2680.

Hertz, M.I., Thompson, S.R., 2011a. In vivo functional analysis of the Dicistroviridae intergenic region internal ribosome entry sites. Nucleic Acids Res 39, 7276-7288.

Hertz, M.I., Thompson, S.R., 2011b. In vivo functional analysis of the Dicistroviridae intergenic region internal ribosome entry sites. Nucleic Acids Res.

Herzog, E., Guilley, H., Fritsch, C., 1995. Translation of the second gene of peanut clump virus RNA 2 occurs by leaky scanning in vitro. Virology 208, 215-225.

Hinnebusch, A.G., Ivanov, I.P., Sonenberg, N., 2016. Translational control by 5'-untranslated regions of eukaryotic mRNAs. Science 352, 1413-1416.

Hinnebusch, A.G., Lorsch, J.R., 2012. The Mechanism of Eukaryotic Translation Initiation: New Insights and Challenges, in: Hershey, J.W.B., Sonenberg, N., Mathews, M.B. (Eds.), Protein

Synthesis and Translational Control. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 29-54.

Hsu, N.Y., Ilnytska, O., Belov, G., Santiana, M., Chen, Y.H., Takvorian, P.M., Pau, C., van der Schaar, H., Kaushik-Basu, N., Balla, T., Cameron, C.E., Ehrenfeld, E., van Kuppeveld, F.J., Altan-Bonnet, N., 2010. Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. Cell 141, 799-811.

Huang, W.M., Ao, S.Z., Casjens, S., Orlandi, R., Zeikus, R., Weiss, R., Winge, D., Fang, M., 1988. A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. Science 239, 1005-1012.

Hurley, J.H., Hanson, P.I., 2010. Membrane budding and scission by the ESCRT machinery: it's all in the neck. Nature reviews. Molecular cell biology 11, 556-566.

Huszar, T., Imler, J.L., 2008. Drosophila viruses and the study of antiviral host-defense. Advances in virus research 72, 227-265.

Ingolia, N.T., Lareau, L.F., Weissman, J.S., 2011. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell 147, 789-802.

Isken, O., Kim, Y.K., Hosoda, N., Mayeur, G.L., Hershey, J.W., Maquat, L.E., 2008. Upf1 phosphorylation triggers translational repression during nonsense-mediated mRNA decay. Cell 133, 314-327.

Jack, K., Bellodi, C., Landry, D.M., Niederer, R.O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, A.M., Thompson, S.R., Ruggero, D., Dinman, J.D., 2011. rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. Mol Cell 44, 660-666.

Jacks, T., Madhani, H.D., Masiarz, F.R., Varmus, H.E., 1988a. Signals for ribosomal frameshifting in the Rous sarcoma virus gag-pol region. Cell 55, 447-458.

Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J., Varmus, H.E., 1988b. Characterization of ribosomal frameshifting in HIV-1 gag-pol expression. Nature 331, 280-283.

Jacks, T., Varmus, H.E., 1985. Expression of the Rous sarcoma virus pol gene by ribosomal frameshifting. Science 230, 1237-1242.

Jackson, R.J., Hellen, C.U., Pestova, T.V., 2010. The mechanism of eukaryotic translation initiation and principles of its regulation. Nature reviews. Molecular cell biology 11, 113-127.

Jackson, R.J., Hellen, C.U., Pestova, T.V., 2012. Termination and post-termination events in eukaryotic translation. Advances in protein chemistry and structural biology 86, 45-93.

Jackson, W.T., 2014. Poliovirus-induced changes in cellular membranes throughout infection. Curr Opin Virol 9, 67-73.

Jackson, W.T., Giddings, T.H., Jr., Taylor, M.P., Mulinyawe, S., Rabinovitch, M., Kopito, R.R., Kirkegaard, K., 2005. Subversion of cellular autophagosomal machinery by RNA viruses. PLoS Biol 3, e156.

Jan, E., 2006. Divergent IRES elements in invertebrates. Virus Res 119
, 16-28.

Jan, E., Kinzy, T.G., Sarnow, P., 2003. Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis. Proc Natl Acad Sci U S A 100, 15410-15415.

Jan, E., Mohr, I., Walsh, D., 2016. A Cap-to-Tail Guide to mRNA Translation Strategies in Virus-Infected Cells. Annual review of virology 3, 283-307.

Jan, E., Sarnow, P., 2002. Factorless ribosome assembly on the internal ribosome entry site of cricket paralysis virus. J Mol Biol 324, 889-902.

Jang, C.J., Jan, E., 2010. Modular domains of the Dicistroviridae intergenic internal ribosome entry site. RNA 16, 1182-1195.

Jang, C.J., Lo, M.C., Jan, E., 2009. Conserved element of the dicistrovirus IGR IRES that mimics an E-site tRNA/ribosome interaction mediates multiple functions. J Mol Biol 387, 42-58.

Jang, S.K., Krausslich, H.G., Nicklin, M.J., Duke, G.M., Palmenberg, A.C., Wimmer, E., 1988. A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation. J Virol 62, 2636-2643.

Johannes, G., Carter, M.S., Eisen, M.B., Brown, P.O., Sarnow, P., 1999. Identification of eukaryotic mRNAs that are translated at reduced cap binding complex eIF4F concentrations using a cDNA microarray. Proc Natl Acad Sci U S A 96, 13118-13123.

Kamoshita, N., Nomoto, A., RajBhandary, U.L., 2009. Translation initiation from the ribosomal A site or the P site, dependent on the conformation of RNA pseudoknot I in dicistrovirus RNAs. Mol Cell 35, 181-190.

Kanamori, Y., Nakashima, N., 2001. A tertiary structure model of the internal ribosome entry site (IRES) for methionine-independent initiation of translation. RNA 7, 266-274.

Kapp, L.D., Lorsch, J.R., 2004. GTP-dependent recognition of the methionine moiety on initiator tRNA by translation factor eIF2. J Mol Biol 335, 923-936.

Kemp, C., Mueller, S., Goto, A., Barbier, V., Paro, S., Bonnay, F., Dostert, C., Troxler, L., Hetru, C., Meignin, C., Pfeffer, S., Hoffmann, J.A., Imler, J.L., 2013. Broad RNA interference-mediated antiviral immunity and virus-specific inducible responses in Drosophila. J Immunol 190, 650-658.

Kerr, C.H., Ma, Z.W., Jang, C.J., Thompson, S.R., Jan, E., 2016. Molecular analysis of the factorless internal ribosome entry site in Cricket Paralysis virus infection. Scientific reports 6, 37319.

Kerr, C.H., Wang, Q.S., Keatings, K., Khong, A., Allan, D., Yip, C.K., Foster, L.J., Jan, E., 2015. The 5' untranslated region of a novel infectious molecular clone of the dicistrovirus cricket paralysis virus modulates infection. J Virol 89, 5919-5934.

Khong, A., Bonderoff, J.M., Spriggs, R.V., Tammpere, E., Kerr, C.H., Jackson, T.J., Willis, A.E., Jan, E., 2016. Temporal Regulation of Distinct Internal Ribosome Entry Sites of the Dicistroviridae Cricket Paralysis Virus. Viruses 8.

Khong, A., Jan, E., 2011. Modulation of stress granules and P bodies during dicistrovirus infection. J Virol 85, 1439-1451.

Khong, A., Kerr, C.H., Yeung, C.H., Keatings, K., Nayak, A., Allan, D.W., Jan, E., 2017. Disruption of Stress Granule Formation by the Multifunctional Cricket Paralysis Virus 1A Protein. J Virol 91.

Khromykh, A.A., Westaway, E.G., 1994. Completion of Kunjin virus RNA sequence and recovery of an infectious RNA transcribed from stably cloned full-length cDNA. J Virol 68, 4580-4588.

Kieft, J.S., 2008. Viral IRES RNA structures and ribosome interactions. Trends Biochem Sci 33, 274-283.

Kim, Y.G., Su, L., Maas, S., O'Neill, A., Rich, A., 1999. Specific mutations in a viral RNA pseudoknot drastically change ribosomal frameshifting efficiency. Proc Natl Acad Sci U S A 96, 14234-14239.

Klock, H.E., Lesley, S.A., 2009. The Polymerase Incomplete Primer Extension (PIPE) method applied to high-throughput cloning and site-directed mutagenesis. Methods Mol Biol 498, 91-103.

Koh, C.S., Brilot, A.F., Grigorieff, N., Korostelev, A.A., 2014. Taura syndrome virus IRES initiates translation by binding its tRNA-mRNA-like structural element in the ribosomal decoding center. Proc Natl Acad Sci U S A 111, 9139-9144.

Konan, K.V., Sanchez-Felipe, L., 2014. Lipids and RNA virus replication. Curr Opin Virol 9, 45-52.

Konarska, M.M., Padgett, R.A., Sharp, P.A., 1984. Recognition of cap structure in splicing in vitro of mRNA precursors. Cell 38, 731-736.

Koppen, T., Weckmann, A., Muller, S., Staubach, S., Bloch, W., Dohmen, R.J., Schwientek, T., 2011. Proteomics analyses of microvesicles released by Drosophila Kc167 and S2 cells. Proteomics 11, 4397-4410.

Kozak, M., 1986. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell 44, 283-292.

Kozak, M., 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J Biol Chem 266, 19867-19870.

Kozak, M., 2002. Pushing the limits of the scanning mechanism for initiation of translation. Gene 299, 1-34.

Kozak, M., 2003. Alternative ways to think about mRNA sequences and proteins that appear to promote internal initiation of translation. Gene 318, 1-23.

Kozak, M., 2005. A second look at cellular mRNA sequences said to function as internal ribosome entry sites. Nucleic Acids Res 33, 6593-6602.

Kozak, M., Shatkin, A.J., 1978. Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. J Biol Chem 253, 6568-6577.

Krishna, N.K., Marshall, D., Schneemann, A., 2003. Analysis of RNA packaging in wild-type and mosaic protein capsids of flock house virus using recombinant baculovirus vectors. Virology 305, 10-24.

Lacerda, R., Menezes, J., Romao, L., 2017. More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. Cellular and molecular life sciences : CMLS 74, 1659-1680.

Lai, C.J., Zhao, B.T., Hori, H., Bray, M., 1991. Infectious RNA transcribed from stably cloned full-length cDNA of dengue type 4 virus. Proc Natl Acad Sci U S A 88, 5139-5143.

Landry, D.M., Hertz, M.I., Thompson, S.R., 2009. RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs. Genes Dev 23, 2753-2764.

Lang, B.F., Jakubkova, M., Hegedusova, E., Daoud, R., Forget, L., Brejova, B., Vinar, T., Kosa, P., Fricova, D., Nebohacova, M., Griac, P., Tomaska, L., Burger, G., Nosek, J., 2014. Massive programmed translational jumping in mitochondria. Proc Natl Acad Sci U S A 111, 5926-5931.

Lauring, A.S., Andino, R., 2010. Quasispecies theory and the behavior of RNA viruses. PLoS Pathog 6, e1001005.

Leat, N., Ball, B., Govan, V., Davison, S., 2000. Analysis of the complete genome sequence of black queen-cell virus, a picorna-like virus of honey bees. J Gen Virol 81, 2111-2119.

Lee, C.K., Wimmer, E., 1988. Proteolytic processing of poliovirus polyprotein: elimination of 2Apro-mediated, alternative cleavage of polypeptide 3CD by in vitro mutagenesis. Virology 166, 405-414.

Lee, J.C., Wu, T.Y., Huang, C.F., Yang, F.M., Shih, S.R., Hsu, J.T., 2005. High-efficiency protein expression mediated by enterovirus 71 internal ribosome entry site. Biotechnol Bioeng 90, 656-662.

Lee, Y.F., Nomoto, A., Detjen, B.M., Wimmer, E., 1977. A protein covalently linked to poliovirus genome RNA. Proc Natl Acad Sci U S A 74, 59-63.

Lemon, S.M., 1985. Type A viral hepatitis. New developments in an old disease. N Engl J Med 313, 1059-1067.

Lemon, S.M., Jansen, R.W., Newbold, J.E., 1985. Infectious hepatitis A virus particles produced in cell culture consist of three distinct types with different buoyant densities in CsCl. J Virol 54, 78-85.

Leveque, N., Semler, B.L., 2015. A 21st century perspective of poliovirus replication. PLoS Pathog 11, e1004825.

Lewis, J.D., Izaurralde, E., 1997. The role of the cap structure in RNA processing and nuclear export. Eur J Biochem 247, 461-469.

Li, D., Aaskov, J., Lott, W.B., 2011. Identification of a cryptic prokaryotic promoter within the cDNA encoding the 5' end of dengue virus RNA genome. PLoS ONE 6, e18197.

Li, X., Song, J., Yi, C., 2014a. Genome-wide mapping of cellular protein-RNA interactions enabled by chemical crosslinking. Genomics Proteomics Bioinformatics 12, 72-78.

Li, Y., Treffers, E.E., Napthine, S., Tas, A., Zhu, L., Sun, Z., Bell, S., Mark, B.L., van Veelen, P.A., van Hemert, M.J., Firth, A.E., Brierley, I., Snijder, E.J., Fang, Y., 2014b. Transactivation of programmed ribosomal frameshifting by a viral protein. Proc Natl Acad Sci U S A 111, E2172-2181.

Liang, H., Yao, N., Song, J.T., Luo, S., Lu, H., Greenberg, J.T., 2003. Ceramides modulate programmed cell death in plants. Genes Dev 17, 2636-2641.

Liebig, H.D., Ziegler, E., Yan, R., Hartmuth, K., Klump, H., Kowalski, H., Blaas, D., Sommergruber, W., Frasel, L., Lamphear, B., et al., 1993. Purification of two picornaviral 2A proteinases: interaction with eIF-4 gamma and influence on in vitro translation. Biochemistry 32, 7581-7588.

Lightner, D.V., 1996. Epizootiology, distribution and the impact on international trade of two penaeid shrimp viruses in the Americas. Rev Sci Tech 15, 579-601.

Lightner, D.V., Redman, R.M., 1998. Strategies for the control of viral diseases of shrimp in the Americas. Fish Pathology, 165-180.

Liu, Y., Wimmer, E., Paul, A.V., 2009. Cis-acting RNA elements in human and animal plus-strand RNA viruses. Biochim Biophys Acta 1789, 495-517.

Llacer, J.L., Hussain, T., Marler, L., Aitken, C.E., Thakur, A., Lorsch, J.R., Hinnebusch, A.G., Ramakrishnan, V., 2015. Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex. Mol Cell 59, 399-412.

Lloyd, R.E., Bovee, M., 1993. Persistent infection of human erythroblastoid cells by poliovirus. Virology 194, 200-209.

Ma, X.M., Blenis, J., 2009. Molecular mechanisms of mTOR-mediated translational control. Nature reviews. Molecular cell biology 10, 307-318.

Magny, E.G., Pueyo, J.I., Pearl, F.M., Cespedes, M.A., Niven, J.E., Bishop, S.A., Couso, J.P., 2013. Conserved regulation of cardiac calcium uptake by peptides encoded in small open reading frames. Science 341, 1116-1120.

Majzoub, K., Hafirassou, M.L., Meignin, C., Goto, A., Marzi, S., Fedorova, A., Verdier, Y., Vinh, J., Hoffmann, J.A., Martin, F., Baumert, T.F., Schuster, C., Imler, J.L., 2014. RACK1 controls IRES-mediated translation of viruses. Cell 159, 1086-1095.

Maldonado, R., Herr, A.J., 1998. Efficiency of T4 gene 60 translational bypassing. J Bacteriol 180, 1822-1830.

Manrubia, S.C., Escarmís, C., Domingo, E., Lázaro, E., 2005. High mutation rates, bottlenecks, and robustness of RNA viral quasispecies. Gene 347, 273-282.

Maquat, L.E., Tarn, W.Y., Isken, O., 2010. The pioneer round of translation: features and functions. Cell 142, 368-374.

Mari, J., Poulos, B.T., Lightner, D.V., Bonami, J.R., 2002. Shrimp Taura syndrome virus: genomic characterization and similarity with members of the genus Cricket paralysis-like viruses. J Gen Virol 83, 915-926.

Masoumi, A., Hanzlik, T.N., Christian, P.D., 2003. Functionality of the 5′- and intergenic IRES elements of cricket paralysis virus in a range of insect cell lines, and its relationship with viral activities. Virus Research 94, 113-120.

Matsuda, D., Dreher, T.W., 2006. Close spacing of AUG initiation codons confers dicistronic character on a eukaryotic mRNA. RNA 12, 1338-1349.

Meckes, D.G., Jr., Raab-Traub, N., 2011. Microvesicles and viral infection. J Virol 85, 12844-12854.

Mendelsohn, C.L., Wimmer, E., Racaniello, V.R., 1989. Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell 56, 855-865.

Miras, M., Miller, W.A., Truniger, V., Aranda, M.A., 2017. Non-canonical Translation in Plant RNA Viruses. Front Plant Sci 8, 494.

Molla, A., Paul, A.V., Wimmer, E., 1991. Cell-free, de novo synthesis of poliovirus. Science 254, 1647-1651.

Moon, J.S., Domier, L.L., McCoppin, N.K., D'Arcy, C.J., Jin, H., 1998. Nucleotide sequence analysis shows that Rhopalosiphum padi virus is a member of a novel group of insect-infecting RNA viruses. Virology 243, 54-65.

Moore, N.F., Kearns, A., Pullin, J.S.K., 1980. Characterization of Cricket Paralysis Virus-Induced Polypeptides in Drosophila cells. J Virology 33, 1-9.

Morris, D.R., Geballe, A.P., 2000. Upstream open reading frames as regulators of mRNA translation. Mol Cell Biol 20, 8635-8642.

Muhs, M., Hilal, T., Mielke, T., Skabkin, M.A., Sanbonmatsu, K.Y., Pestova, T.V., Spahn, C.M., 2015. Cryo-EM of ribosomal 80S complexes with termination factors reveals the translocated cricket paralysis virus IRES. Mol Cell 57, 422-432.

Muhs, M., Yamamoto, H., Ismer, J., Takaku, H., Nashimoto, M., Uchiumi, T., Nakashima, N., Mielke, T., Hildebrand, P.W., Nierhaus, K.H., Spahn, C.M., 2011. Structural basis for the binding of IRES RNAs to the head of the ribosomal 40S subunit. Nucleic Acids Res 39, 5264-5275.

Murray, J., Savva, C.G., Shin, B.S., Dever, T.E., Ramakrishnan, V., Fernandez, I.S., 2016. Structural characterization of ribosome recruitment and translocation by type IV IRES. eLife 5.

Murray, N.E., Quam, M.B., Wilder-Smith, A., 2013. Epidemiology of dengue: past, present and future prospects. Clin Epidemiol 5, 299-309.

Muscio, O.A., LaTorre, J.L., Scodeller, E.A., 1987. Small nonoccluded viruses from triatomine bug Triatoma infestans (Hemiptera: Reduviidae). J Invertebr Pathol 49, 218-220.

Nagashima, S., Takahashi, M., Jirintai, Tanaka, T., Yamada, K., Nishizawa, T., Okamoto, H., 2011a. A PSAP motif in the ORF3 protein of hepatitis E virus is necessary for virion release from infected cells. J Gen Virol 92, 269-278.

Nagashima, S., Takahashi, M., Jirintai, S., Tanaka, T., Nishizawa, T., Yasuda, J., Okamoto, H., 2011b. Tumour susceptibility gene 101 and the vacuolar protein sorting pathway are required for the release of hepatitis E virions. J Gen Virol 92, 2838-2848.

Nakashima, N., Shibuya, N., 2006. Multiple coding sequences for the genome-linked virus protein (VPg) in dicistroviruses. J Invertebr Pathol 92, 100-104.

Nakashima, N., Uchiumi, T., 2009. Functional analysis of structural motifs in dicistroviruses. Virus Res 139, 137-147.

Napthine, S., Ling, R., Finch, L.K., Jones, J.D., Bell, S., Brierley, I., Firth, A.E., 2017. Protein-directed ribosomal frameshifting temporally regulates gene expression. Nature communications 8, 15582.

Napthine, S., Treffers, E.E., Bell, S., Goodfellow, I., Fang, Y., Firth, A.E., Snijder, E.J., Brierley, I., 2016. A novel role for poly(C) binding proteins in programmed ribosomal frameshifting. Nucleic Acids Res 44, 5491-5503.

Nathanson, N., Kew, O.M., 2010. From emergence to eradication: the epidemiology of poliomyelitis deconstructed. Am J Epidemiol 172, 1213-1229.

Nayak, A., Berry, B., Tassetto, M., Kunitomi, M., Acevedo, A., Deng, C., Krutchinsky, A., Gross, J., Antoniewski, C., Andino, R., 2010a. Cricket paralysis virus antagonizes Argonaute 2 to modulate antiviral defense in Drosophila. Nature Structural & Molecular Biology 17, 547-554.

Nayak, A., Berry, B., Tassetto, M., Kunitomi, M., Acevedo, A., Deng, C., Krutchinsky, A., Gross, J., Antoniewski, C., Andino, R., 2010b. Cricket paralysis virus antagonizes Argonaute 2 to modulate antiviral defense in Drosophila. Nat Struct Mol Biol 17, 547-554.

Nayak, A., Goodfellow, I.G., Belsham, G.J., 2005. Factors required for the Uridylylation of the foot-and-mouth disease virus 3B1, 3B2, and 3B3 peptides by the RNA-dependent RNA polymerase (3Dpol) in vitro. J Virol 79, 7698-7706.

Nieva, J.L., Madan, V., Carrasco, L., 2012. Viroporins: structure and biological functions. Nature reviews. Microbiology 10, 563-574.

Nishiyama, T., Yamamoto, H., Uchiumi, T., Nakashima, N., 2007. Eukaryotic ribosomal protein RPS25 interacts with the conserved loop region in a dicistroviral intergenic internal ribosome entry site. Nucleic Acids Res 35, 1514-1521.

Nomura, M., Gourse, R., Baughman, G., 1984. Regulation of the synthesis of ribosomes and ribosomal components. Annu Rev Biochem 53, 75-117.

Pacheco, J.M., Henry, T.M., O'Donnell, V.K., Gregory, J.B., Mason, P.W., 2003. Role of nonstructural proteins 3A and 3B in host range and pathogenicity of foot-and-mouth disease virus. J Virol 77, 13017-13027.

Pacheco, J.M., Piccone, M.E., Rieder, E., Pauszek, S.J., Borca, M.V., Rodriguez, L.L., 2010. Domain disruptions of individual 3B proteins of foot-and-mouth disease virus do not alter growth in cell culture or virulence in cattle. Virology 405, 149-156.

Paek, K.Y., Hong, K.Y., Ryu, I., Park, S.M., Keum, S.J., Kwon, O.S., Jang, S.K., 2015. Translation initiation mediated by RNA looping. Proc Natl Acad Sci U S A 112, 1041-1046.

Pal, N., Boyapalle, S., Beckett, R.J., Miller, W.A., Bonning, B.C., 2014. A baculovirus-expressed dicistrovirus that is infectious to aphids. J Virol 88, 3610.

Paloheimo, O., Ihalainen, T.O., Tauriainen, S., Valilehto, O., Kirjavainen, S., Niskanen, E.A., Laakkonen, J.P., Hyoty, H., Vihinen-Ranta, M., 2011. Coxsackievirus B3-induced cellular protrusions: structural characteristics and functional competence. J Virol 85, 6714-6724.

Passmore, L.A., Schmeing, T.M., Maag, D., Applefield, D.J., Acker, M.G., Algire, M.A., Lorsch, J.R., Ramakrishnan, V., 2007. The eukaryotic translation initiation factors eIF1 and eIF1A induce an open conformation of the 40S ribosome. Mol Cell 26, 41-50.

Patel, N., Dykeman, E.C., Coutts, R.H., Lomonossoff, G.P., Rowlands, D.J., Phillips, S.E., Ranson, N., Twarock, R., Tuma, R., Stockley, P.G., 2015. Revealing the density of encoded functions in a viral RNA. Proc Natl Acad Sci U S A.

Paul, A.V., Rieder, E., Kim, D.W., van Boom, J.H., Wimmer, E., 2000. Identification of an RNA hairpin in poliovirus RNA that serves as the primary template in the in vitro uridylylation of VPg. J Virol 74, 10359-10370.

Paulin, F.E., Campbell, L.E., O'Brien, K., Loughlin, J., Proud, C.G., 2001. Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein. Curr Biol 11, 55-59.

Pelletier, J., Sonenberg, N., 1988. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. Nature 334, 320-325.

Pestova, T.V., Hellen, C.U., 2003. Translation elongation after assembly of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA. Genes Dev 17, 181-186.

Pestova, T.V., Lomakin, I.B., Lee, J.H., Choi, S.K., Dever, T.E., Hellen, C.U., 2000. The joining of ribosomal subunits in eukaryotes requires eIF5B. Nature 403, 332-335.

Pestova, T.V., Shatsky, I.N., Fletcher, S.P., Jackson, R.J., Hellen, C.U., 1998. A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initiation of hepatitis C and classical swine fever virus RNAs. Genes Dev 12, 67-83.

Petersen, C.P., Bordeleau, M.E., Pelletier, J., Sharp, P.A., 2006. Short RNAs repress translation after initiation in mammalian cells. Mol Cell 21, 533-542.

Petrov, A., Grosely, R., Chen, J., O'Leary, S.E., Puglisi, J.D., 2016. Multiple Parallel Pathways of Translation Initiation on the CrPV IRES. Mol Cell 62, 92-103.

Peyrefitte, C.N., Pastorino, B., Bessaud, M., Tolou, H.J., Couissinier-Paris, P., 2003. Evidence for in vitro falsely-primed cDNAs that prevent specific detection of virus negative strand RNAs in dengue-infected cells: improvement by tagged RT-PCR. Journal of Virological Methods 113, 19-28.

Pfingsten, J.S., Costantino, D.A., Kieft, J.S., 2006. Structural basis for ribosome recruitment and manipulation by a viral IRES RNA. Science 314, 1450-1454.

Pfister, T., Wimmer, E., 1999. Characterization of the nucleoside triphosphatase activity of poliovirus protein 2C reveals a mechanism by which guanidine inhibits poliovirus replication. J Biol Chem 274, 6992-7001.

Pickl-Herk, A., Luque, D., Vives-Adrian, L., Querol-Audi, J., Garriga, D., Trus, B.L., Verdaguer, N., Blaas, D., Caston, J.R., 2013. Uncoating of common cold virus is preceded by RNA switching as determined by X-ray and cryo-EM analyses of the subviral A-particle. Proc Natl Acad Sci U S A 110, 20063-20068.

Plank, T.D., Kieft, J.S., 2012. The structures of nonprotein-coding RNAs that drive internal ribosome entry site function. Wiley interdisciplinary reviews. RNA 3, 195-212.

Plaskon, N.E., Adelman, Z.N., Myles, K.M., 2009. Accurate strand-specific quantification of viral RNA. PLoS ONE 4, e7468.

Poyry, T.A., Kaminski, A., Connell, E.J., Fraser, C.S., Jackson, R.J., 2007. The mechanism of an exceptional case of reinitiation after translation of a long ORF reveals why such events do not generally occur in mammalian mRNA translation. Genes Dev 21, 3149-3162.

Poyry, T.A., Kaminski, A., Jackson, R.J., 2004. What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? Genes Dev 18, 62-75.

Raab-Traub, N., Dittmer, D.P., 2017. Viral effects on the content and function of extracellular vesicles. Nature reviews. Microbiology 15, 559-572.

Racaniello, V.R., Baltimore, D., 1981. Cloned poliovirus complementary DNA is infectious in mammalian cells. Science 214, 916-919.

Raposo, G., Stoorvogel, W., 2013. Extracellular vesicles: exosomes, microvesicles, and friends. J Cell Biol 200, 373-383.

Reddy, K.E., Noh, J.H., Kim, Y.H., Yoo, M.S., Doan, H.T., Ramya, M., Jung, S.C., Quyen, D.V., Kang, S.W., 2013. Analysis of the nonstructural and structural polyprotein regions, and complete genome sequences of Israel acute paralysis viruses identified from honeybees (Apis mellifera) in Korea. Virology 444, 211-217.

Reddy, K.E., Yoo, M.-S., Kim, Y.-H., Kim, N.-H., Jung, H.-N., Thao, L.T.B., Ramya, M., Doan, H.T.T., Nguyen, L.T.K., Jung, S.-C., Kang, S.-W., 2014. Analysis of the RdRp, intergenic and structural polyprotein regions, and the complete genome sequence of Kashmir bee virus from infected honeybees (Apis mellifera) in Korea. Virus Genes 49, 137-144.

Reinganum, C., 1975. The isolation of cricket paralysis virus from the emperor gum moth, Antheraea eucalypti Scott, and its infectivity towards a range of insect species. Intervirology 5, 97-102.

Reinganum, C., G.T. O'Loughlin, and T.W. Hogan, 1970. A nonoccluded virus of the field crickets Teleogryllus aceanicus and T. commodus (Orthoptera: Gryllidae). Journal of Invertebrate Pathology 16, 214-220.

Ren, Q., Au, H.H., Wang, Q.S., Lee, S., Jan, E., 2014. Structural determinants of an internal ribosome entry site that direct translational reading frame selection. Nucleic Acids Res 42, 9366-9382.

Ren, Q., Wang, Q.S., Firth, A.E., Chan, M.M., Gouw, J.W., Guarna, M.M., Foster, L.J., Atkins, J.F., Jan, E., 2012. Alternative reading frame selection mediated by a tRNA-like domain of an internal ribosome entry site. Proc Natl Acad Sci U S A 109, E630-639.

Richter, J.D., Sonenberg, N., 2005. Regulation of cap-dependent translation by eIF4E inhibitory proteins. Nature 433, 477-480.

Riley, A., Jordan, L.E., Holcik, M., 2010. Distinct 5' UTRs regulate XIAP expression under normal growth conditions and during cellular stress. Nucleic Acids Res 38, 4665-4674.

Robbins, P.D., Morelli, A.E., 2014. Regulation of immune responses by extracellular vesicles. Nature reviews. Immunology 14, 195-208.

Roberts, L.O., Groppelli, E., 2009. An atypical IRES within the 5' UTR of a dicistrovirus genome. Virus Res 139, 157-165.

Robinson, S.M., Tsueng, G., Sin, J., Mangale, V., Rahawi, S., McIntyre, L.L., Williams, W., Kha, N., Cruz, C., Hancock, B.M., Nguyen, D.P., Sayen, M.R., Hilton, B.J., Doran, K.S., Segall, A.M., Wolkowicz, R., Cornell, C.T., Whitton, J.L., Gottlieb, R.A., Feuer, R., 2014. Coxsackievirus B exits the host cell in shed microvesicles displaying autophagosomal markers. PLoS Pathog 10, e1004045.

Romero-Brey, I., Merz, A., Chiramel, A., Lee, J.Y., Chlanda, P., Haselman, U., Santarella-Mellwig, R., Habermann, A., Hoppe, S., Kallis, S., Walther, P., Antony, C., Krijnse-Locker, J., Bartenschlager, R., 2012. Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. PLoS Pathog 8, e1003056.

Rossman, J.S., Lamb, R.A., 2011. Influenza virus assembly and budding. Virology 411, 229-236.

Roussarie, J.P., Ruffie, C., Edgar, J.M., Griffiths, I., Brahic, M., 2007. Axon myelin transfer of a non-enveloped virus. PLoS One 2, e1331.

Roux, K.J., Kim, D.I., Burke, B., 2013. BioID: a screen for protein-protein interactions. Curr Protoc Protein Sci 74, Unit 19 23.

Rowlands, A.G., Panniers, R., Henshaw, E.C., 1988. The catalytic mechanism of guanine nucleotide exchange factor action and competitive inhibition by phosphorylated eukaryotic initiation factor 2. J Biol Chem 263, 5526-5533.

Roy, G., Miron, M., Khaleghpour, K., Lasko, P., Sonenberg, N., 2004. The Drosophila poly(A) binding protein-interacting protein, dPaip2, is a novel effector of cell growth. Mol Cell Biol 24, 1143-1154.

Royall, E., Woolaway, K.E., Schacherl, J., Kubick, S., Belsham, G.J., Roberts, L.O., 2004. The Rhopalosiphum padi virus 5' internal ribosome entry site is functional in Spodoptera frugiperda 21 cells and in their cell-free lysates: implications for the baculovirus expression system. J Gen Virol 85, 1565-1569.

Ruehle, M.D., Zhang, H., Sheridan, R.M., Mitra, S., Chen, Y., Gonzalez, R.L., Cooperman, B.S., Kieft, J.S., 2015. A dynamic RNA loop in an IRES affects multiple steps of elongation factor-mediated translation initiation. eLife 4.

Ryan, R.M., Morris, R.E., Rice, W.R., Ciraolo, G., Whitsett, J.A., 1989. Binding and uptake of pulmonary surfactant protein (SP-A) by pulmonary type II epithelial cells. J Histochem Cytochem 37, 429-440.

Sabath, N., Price, N., Graur, D., 2009. A potentially novel overlapping gene in the genomes of Israeli acute paralysis virus and its relatives. Virol J 6, 144.

Sabin, L.R., Hanna, S.L., Cherry, S., 2010. Innate antiviral immunity in Drosophila. Curr Opin Immunol 22, 4-9.

Sahu, R., Kaushik, S., Clement, C.C., Cannizzo, E.S., Scharf, B., Follenzi, A., Potolicchio, I., Nieves, E., Cuervo, A.M., Santambrogio, L., 2011. Microautophagy of cytosolic proteins by late endosomes. Dev Cell 20, 131-139.

Samatova, E., Konevega, A.L., Wills, N.M., Atkins, J.F., Rodnina, M.V., 2014. High-efficiency translational bypassing of non-coding nucleotides specified by mRNA structure and nascent peptide. Nature communications 5, 4459.

Sanjuan, R., Nebot, M.R., Chirico, N., Mansky, L.M., Belshaw, R., 2010. Viral mutation rates. J Virol 84, 9733-9748.

Sarabhai, A., Brenner, S., 1967. A mutant which reinitiates the polypeptide chain after chain termination. J Mol Biol 27, 145-162.

Sarnow, P., 1989. Role of 3'-end sequences in infectivity of poliovirus transcripts made in vitro. J Virol 63, 467-470.

Sasaki, J., Nakashima, N., 1999. Translation initiation at the CUU codon is mediated by the internal ribosome entry site of an insect picorna-like virus in vitro. J Virol 73, 1219-1226.

Sasaki, J., Nakashima, N., 2000. Methionine-independent initiation of translation in the capsid protein of an insect RNA virus. Proc Natl Acad Sci U S A 97, 1512-1515.

Sasaki, J., Nakashima, N., Saito, H., Noda, H., 1998. An insect picorna-like virus, Plautia stali intestine virus, has genes of capsid proteins in the 3' part of the genome. Virology 244, 50-58.

Schuler, M., Connell, S.R., Lescoute, A., Giesebrecht, J., Dabrowski, M., Schroeer, B., Mielke, T., Penczek, P.A., Westhof, E., Spahn, C.M., 2006. Structure of the ribosome-bound cricket paralysis virus IRES RNA. Nat Struct Mol Biol 13, 1092-1096.

Schwer, B., Mao, X., Shuman, S., 1998. Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. Nucleic Acids Res 26, 2050-2057.

Sharma, S.D., Kraft, J.J., Miller, W.A., Goss, D.J., 2015. Recruitment of the 40S ribosome subunit to the 3'-untranslated region (UTR) of a viral mRNA, via the eIF4 complex, facilitates cap-independent translation. J Biol Chem 290, 11268-11281.

Shehu-Xhilaga, M., Crowe, S.M., Mak, J., 2001. Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity. J Virol 75, 1834-1841.

Shi, M., Lin, X.D., Tian, J.H., Chen, L.J., Chen, X., Li, C.X., Qin, X.C., Li, J., Cao, J.P., Eden, J.S., Buchmann, J., Wang, W., Xu, J., Holmes, E.C., Zhang, Y.Z., 2016. Redefining the invertebrate RNA virosphere. Nature.

Shibuya, N., Nakashima, N., 2006. Characterization of the 5' internal ribosome entry site of Plautia stali intestine virus. J Gen Virol 87, 3679-3686.

Shuman, S., 2002. What messenger RNA capping tells us about eukaryotic evolution. Nature reviews. Molecular cell biology 3, 619-625.

Simon, A.E., Miller, W.A., 2013. 3' cap-independent translation enhancers of plant viruses. Annu Rev Microbiol 67, 21-42.

Simon-Loriere, E., Holmes, E.C., 2011. Why do RNA viruses recombine? Nature Reviews Microbiology 9, 617-626.

Skabkin, M.A., Skabkina, O.V., Hellen, C.U., Pestova, T.V., 2013. Reinitiation and other unconventional posttermination events during eukaryotic translation. Mol Cell 51, 249-264.

Smith, M.C., Hendrix, R.W., Dedrick, R., Mitchell, K., Ko, C.C., Russell, D., Bell, E., Gregory, M., Bibb, M.J., Pethick, F., Jacobs-Sera, D., Herron, P., Buttner, M.J., Hatfull, G.F., 2013. Evolutionary relationships among actinophages and a putative adaptation for growth in Streptomyces spp. J Bacteriol 195, 4924-4935.

Smith, W., 1939. The action of bile salts on viruses. J Pathol Bacteriol 48, 557-571.

Sonenberg, N., Hinnebusch, A.G., 2009. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell 136, 731-745.

Spahn, C.M., Jan, E., Mulder, A., Grassucci, R.A., Sarnow, P., Frank, J., 2004. Cryo-EM visualization of a viral internal ribosome entry site bound to human ribosomes: the IRES functions as an RNA-based translation factor. Cell 118, 465-475.

Spector, D.H., Baltimore, D., 1974. Requirement of 3'-terminal poly(adenylic acid) for the infectivity of poliovirus RNA. Proc Natl Acad Sci U S A 71, 2983-2987.

Stuart, L.M., Ezekowitz, R.A., 2008. Phagocytosis and comparative innate immunity: learning on the fly. Nature reviews. Immunology 8, 131-141.

Stuffers, S., Sem Wegner, C., Stenmark, H., Brech, A., 2009. Multivesicular endosome biogenesis in the absence of ESCRTs. Traffic 10, 925-937.

Sumiyoshi, H., Hoke, C.H., Trent, D.W., 1992. Infectious Japanese encephalitis virus RNA can be synthesized from in vitro-ligated cDNA templates. J Virol 66, 5425-5431.

Suttle, C.A., 2007. Marine viruses--major players in the global ecosystem. Nature reviews. Microbiology 5, 801-812.

Svitkin, Y.V., Imataka, H., Khaleghpour, K., Kahvejian, A., Liebig, H.D., Sonenberg, N., 2001. Poly(A)-binding protein interaction with elF4G stimulates picornavirus IRES-dependent translation. RNA 7, 1743-1752.

Sweeney, T.R., Abaeva, I.S., Pestova, T.V., Hellen, C.U., 2014. The mechanism of translation initiation on Type 1 picornavirus IRESs. EMBO J 33, 76-92.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., Yamanaka, S., 2007. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861-872.

Takahashi, M., Tanaka, T., Takahashi, H., Hoshino, Y., Nagashima, S., Jirintai, Mizuo, H., Yazaki, Y., Takagi, T., Azuma, M., Kusano, E., Isoda, N., Sugano, K., Okamoto, H., 2010. Hepatitis E Virus (HEV) strains in serum samples can replicate efficiently in cultured cells despite the coexistence of HEV antibodies: characterization of HEV virions in blood circulation. J Clin Microbiol 48, 1112-1125.

Tassetto, M., Kunitomi, M., Andino, R., 2017. Circulating Immune Cells Mediate a Systemic RNAi-Based Adaptive Antiviral Response in Drosophila. Cell 169, 314-325 e313.

Tate, J., Liljas, L., Scotti, P., Christian, P., Lin, T., Johnson, J.E., 1999. The crystal structure of cricket paralysis virus: the first view of a new virus family. Nat Struct Biol 6, 765-774.

Taylor, M.P., Burgon, T.B., Kirkegaard, K., Jackson, W.T., 2009. Role of microtubules in extracellular release of poliovirus. J Virol 83, 6599-6609.

Telpalo-Carpio, S.A., Diaz-Mitoma, F., Moreno-Cuevas, J.E., Aguilar-Yanez, J.M., 2015. Internal ribosome entry site (IRES) from Encephalomyocarditis virus (EMCV) as a tool for shuttle expression plasmids. Biochem Biophys Res Commun 468, 548-553.

Terenin, I.M., Dmitriev, S.E., Andreev, D.E., Royall, E., Belsham, G.J., Roberts, L.O., Shatsky, I.N., 2005. A cross-kingdom internal ribosome entry site reveals a simplified mode of internal ribosome entry. Mol Cell Biol 25, 7879-7888.

Thery, C., Zitvogel, L., Amigorena, S., 2002. Exosomes: composition, biogenesis and function. Nature reviews. Immunology 2, 569-579.

Thompson, S.R., 2012. So you want to know if your message has an IRES? Wiley interdisciplinary reviews. RNA 3, 697-705.

Thompson, S.R., Gulyas, K.D., Sarnow, P., 2001. Internal initiation in Saccharomyces cerevisiae mediated by an initiator tRNA/eIF2-independent internal ribosome entry site element. Proc Natl Acad Sci U S A 98, 12972-12977.

Toyoda, H., Koide, N., Kamiyama, M., Tobita, K., Mizumoto, K., Imura, N., 1994. Host factors required for internal initiation of translation on poliovirus RNA. Arch Virol 138, 1-15.

Trajkovic, K., Hsu, C., Chiantia, S., Rajendran, L., Wenzel, D., Wieland, F., Schwille, P., Brugger, B., Simons, M., 2008. Ceramide triggers budding of exosome vesicles into multivesicular endosomes. Science 319, 1244-1247.

Tucker, S.P., Thornton, C.L., Wimmer, E., Compans, R.W., 1993. Vectorial release of poliovirus from polarized human intestinal epithelial cells. J Virol 67, 4274-4282.

Turina, M., Desvoyes, B., Scholthof, K.B., 2000. A gene cluster encoded by panicum mosaic virus is associated with virus movement. Virology 266, 120-128.

Vaheri, A., Strandin, T., Hepojoki, J., Sironen, T., Henttonen, H., Makela, S., Mustonen, J., 2013. Uncovering the mysteries of hantavirus infections. Nature reviews. Microbiology 11, 539-550.

Valles, S.M., Strong, C.A., Dang, P.M., Hunter, W.B., Pereira, R.M., Oi, D.H., Shapiro, A.M., Williams, D.F., 2004. A picorna-like virus from the red imported fire ant, Solenopsis invicta: initial discovery, genome sequence, and characterization. Virology 328, 151-157.

Valles, S.M., Strong, C.A., Hunter, W.B., Dang, P.M., Pereira, R.M., Oi, D.H., Williams, D.F., 2008. Expressed sequence tags from the red imported fire ant, Solenopsis invicta: annotation and utilization for discovery of viruses. J Invertebr Pathol 99, 74-81.

van Kuppeveld, F.J., Melchers, W.J., Willems, P.H., Gadella, T.W., Jr., 2002. Homomultimerization of the coxsackievirus 2B protein in living cells visualized by fluorescence resonance energy transfer microscopy. J Virol 76, 9446-9456.

van Niel, G., Charrin, S., Simoes, S., Romao, M., Rochin, L., Saftig, P., Marks, M.S., Rubinstein, E., Raposo, G., 2011. The tetraspanin CD63 regulates ESCRT-independent and -dependent endosomal sorting during melanogenesis. Dev Cell 21, 708-721.

van Rij, R.P., Saleh, M.C., Berry, B., Foo, C., Houk, A., Antoniewski, C., Andino, R., 2006. The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. Genes Dev 20, 2985-2995.

Vardi, A., Van Mooy, B.A., Fredricks, H.F., Popendorf, K.J., Ossolinski, J.E., Haramaty, L., Bidle, K.D., 2009. Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton. Science 326, 861-865.

Vassilaki, N., Mavromara, P., 2009. The HCV ARFP/F/core+1 protein: production and functional analysis of an unconventional viral product. IUBMB Life 61, 739-752.

Victoria, J.G., Kapoor, A., Li, L., Blinkova, O., Slikas, B., Wang, C., Naeem, A., Zaidi, S., Delwart, E., 2009. Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. J Virol 83, 4642-4651.

Wang, Q.S., Au, H.H., Jan, E., 2013. Methods for studying IRES-mediated translation of positive-strand RNA viruses. Methods 59, 167-179.

Wang, S., Miller, W.A., 1995. A sequence located 4.5 to 5 kilobases from the 5' end of the barley yellow dwarf virus (PAV) genome strongly stimulates translation of uncapped mRNA. J Biol Chem 270, 13446-13452.

Wang, X.H., 2006. RNA Interference Directs Innate Immunity Against Viruses in Adult Drosophila. Science 312, 452-454.

Watson, F.L., Puttmann-Holgado, R., Thomas, F., Lamar, D.L., Hughes, M., Kondo, M., Rebel, V.I., Schmucker, D., 2005. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science 309, 1874-1878.

Williams, M.A., Lamb, R.A., 1989. Effect of mutations and deletions in a bicistronic mRNA on the synthesis of influenza B virus NB and NA glycoproteins. J Virol 63, 28-35.

Wills, N.M., O'Connor, M., Nelson, C.C., Rettberg, C.C., Huang, W.M., Gesteland, R.F., Atkins, J.F., 2008. Translational bypassing without peptidyl-tRNA anticodon scanning of coding gap mRNA. EMBO J 27, 2533-2544.

Wilson, J.E., Pestova, T.V., Hellen, C.U., Sarnow, P., 2000a. Initiation of protein synthesis from the A site of the ribosome. Cell 102, 511-520.

Wilson, J.E., Powell, M.J., Hoover, S.E., Sarnow, P., 2000b. Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites. Mol Cell Biol 20, 4990-4999.

Woolaway, K.E., Lazaridis, K., Belsham, G.J., Carter, M.J., Roberts, L.O., 2001. The 5' untranslated region of Rhopalosiphum padi virus contains an internal ribosome entry site which functions efficiently in mammalian, plant, and insect translation systems. J Virol 75, 10244-10249.

Wu, Q., Luo, Y., Lu, R., Lau, N., Lai, E.C., Li, W.X., Ding, S.W., 2010. Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs. Proc Natl Acad Sci U S A 107, 1606-1611.

Xiang, W., Harris, K.S., Alexander, L., Wimmer, E., 1995. Interaction between the 5'-terminal cloverleaf and 3AB/3CDpro of poliovirus is essential for RNA replication. J Virol 69, 3658-3667.

Yamamoto, H., Unbehaun, A., Spahn, C.M.T., 2017. Ribosomal Chamber Music: Toward an Understanding of IRES Mechanisms. Trends Biochem Sci 42, 655-668.

Yamamoto, H., Wittek, D., Gupta, R., Qin, B., Ueda, T., Krause, R., Yamamoto, K., Albrecht, R., Pech, M., Nierhaus, K.H., 2016. 70S-scanning initiation is a novel and frequent initiation mode of ribosomal translation in bacteria. Proc Natl Acad Sci U S A 113, E1180-1189.

Yamasaki, K., Weihl, C.C., Roos, R.P., 1999. Alternative translation initiation of Theiler's murine encephalomyelitis virus. J Virol 73, 8519-8526.

Yu, C.H., Noteborn, M.H., Pleij, C.W., Olsthoorn, R.C., 2011. Stem-loop structures can effectively substitute for an RNA pseudoknot in -1 ribosomal frameshifting. Nucleic Acids Res 39, 8952-8959.

Zhang, H., Ng, M.Y., Chen, Y., Cooperman, B.S., 2016. Kinetics of initiating polypeptide elongation in an IRES-dependent system. eLife 5.

Zhu, J., Korostelev, A., Costantino, D.A., Donohue, J.P., Noller, H.F., Kieft, J.S., 2011. Crystal structures of complexes containing domains from two viral internal ribosome entry site (IRES) RNAs bound to the 70S ribosome. Proc Natl Acad Sci U S A 108, 1839-1844.

Appendices

Appendix A : CHAPTER 4 SUPPLEMENTARY DATA

A

Dicistrovirus	IGR IRES Type	Potential +1 basepair	Predicted +1 ORFx Length (aa)	First +1 frame amino acid
Solenopsis invicta virus-1 (SINV-1)	II	C-G	125	L
Israeli acute paralysis virus (IAPV)	II	U-G	94	A
Kashmir bee virus (KBV)	II	C-G	93	L
Acute bee paralysis virus (ABPV)	II	U-G	92	P
Cricket paralysis virus (CrPV)	I	U-G	53	L
Drosophila C virus (DCV)	I	U-G	51	L
Plautia stali intestine virus (PSIV)	I	U C (?)	43	K
Big Sioux River virus (BSRV)	I	U-G	40	L
Black queen cell virus (BQCV)	I	U-G	26	L
Homalodisca coagulata virus- 1	I	U-G	17	Q
Taura Syndrome virus (TSV)	II	U-G	16	L
Himetobi P virus (HiPV)	I	U-G	14	Q
Bat guano dicistrovirus	I	U-G	14	Q
Mud crab dicistrovirus (MCDV)	II	U-G	9	L
Aphid lethal paralysis virus (ALPV)	I	U-G	7	L
ALPV like Brookings virus	I	U-G	7	Q
Rhopalosiphum padi virus (RhPV)	I	U-G	2	Q
Triatoma virus (TrV)	II	C-G	1	L

B

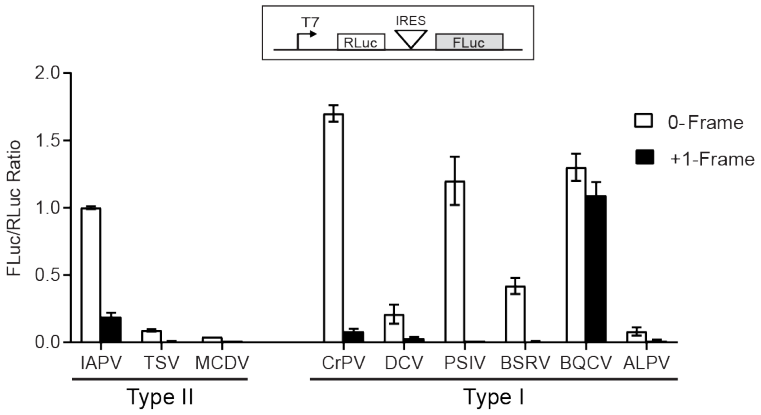


Figure A.1. Biochemical analysis of potential +1-frame translation by various Dicistrovirus IGR IRESs. A) *In silico* analysis for potential +1-frame translation mediated by IGR IRESs. List of type I or type II IGR IRES based on their secondary structures. According to NCBI sequences, the potential +1-frame base pair which locates immediate downstream of their IGR IRES PKI region as in IAPV IGR IRES, together with the predicted length of ORFxs and the first amino acids, are shown. **B)** IGR IRES mediated +1-frame translation using an *in vitro* translation assay. The translation of firefly luciferase (FLuc), which is fused in the 0-frame or +1-frame, is driven by the individual IGR IRES within the bicistronic reporter construct. Linearized reporter constructs are incubated in Sf21 *in vitro* translation insect cell extract at 30 °C for 2 hours in the presence of [^{35}S] methionine/cysteine.

Translation of FLuc and RLuc was monitored by phosphorimager analysis after resolving on a 16% SDS PAGE. On the bottom panel, the ratio of FLuc/RLuc are quantified and normalized to the IAPV 0-frame translation. Shown are averages from at least three independent biological experiments (± SD).

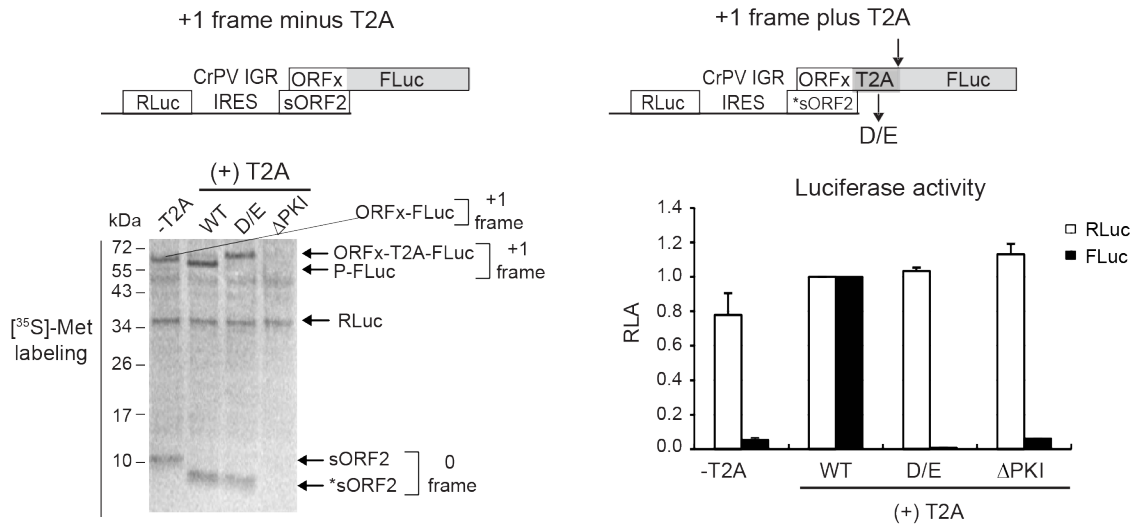


Figure A.2. Construction of a T2A construct +1-frame reporter construct that recover FLuc enzyme activity. **A)** The schemes of bicistronic constructs in which firefly luciferase gene (FLuc) is in frame with the CrPV ORFx are shown. The T2A sequence (dark grey) is inserted between the CrPV ORFx and the FLuc gene. The arrow indicates the 'self-cleavage' or 'stop-go' site. A mutation within the T2A peptide (D to E) inactivates T2A 'self-cleavage' activity. T2A-minus (-T2A) and T2A-containing ((+)T2A) +1-frame bicistronic constructs were incubated in Sf21 extracts for 2 h in the presence of [^{35}S] methionine and cysteine. The left panel is a representative SDS-PAGE monitored by autoradiography, whereas the right panel is the quantitation of relative luciferase activity (RLA) measured by Dual-Luciferase reporter assay (Promega). RLA are normalized to that observed with the +1-frame T2A-containing reporter construct. Shown are averages from at least three independent biological experiments (± SD).

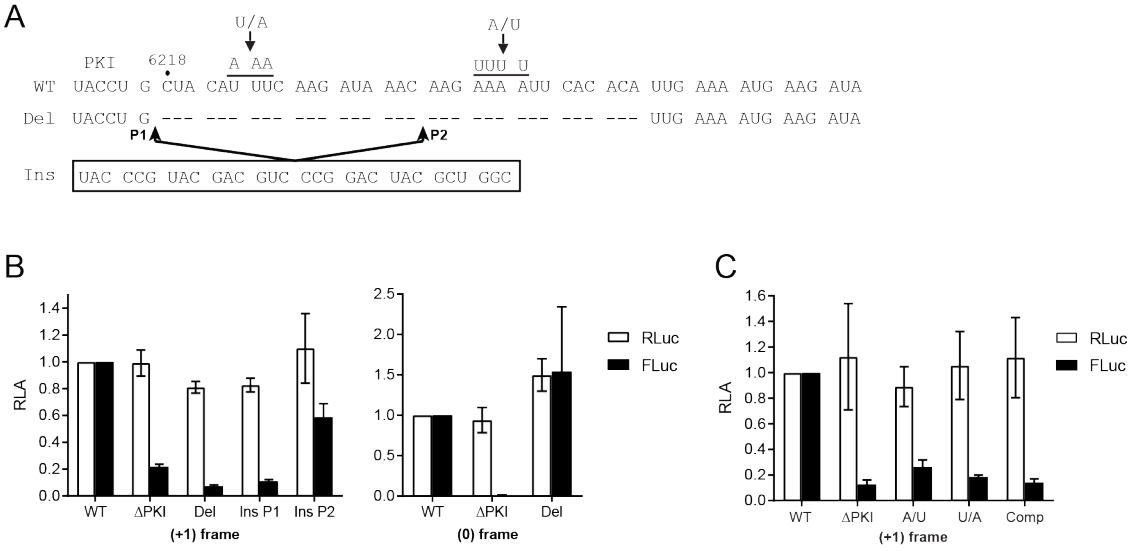


Figure A.3. Effects of deletion and insertion mutations in the CrPV spacer region on +1-frame translation. Schematic of the spacer region between the PKI domain and the +1 frame start site is shown above. *In vitro* translation assay of the CrPV IRES bicistronic construct with the indicated mutations or an insertion mutation (Ins) of 30 random nucleotides between G_{6217} and C_{6218} or between C_{6235} and A_{6236} are monitored by luciferase activities. Both 0- and +1-frame activities are normalized to that of WT constructs. Shown are averages from at least three independent biological experiments (± SD). RLA = relative luciferase activity.

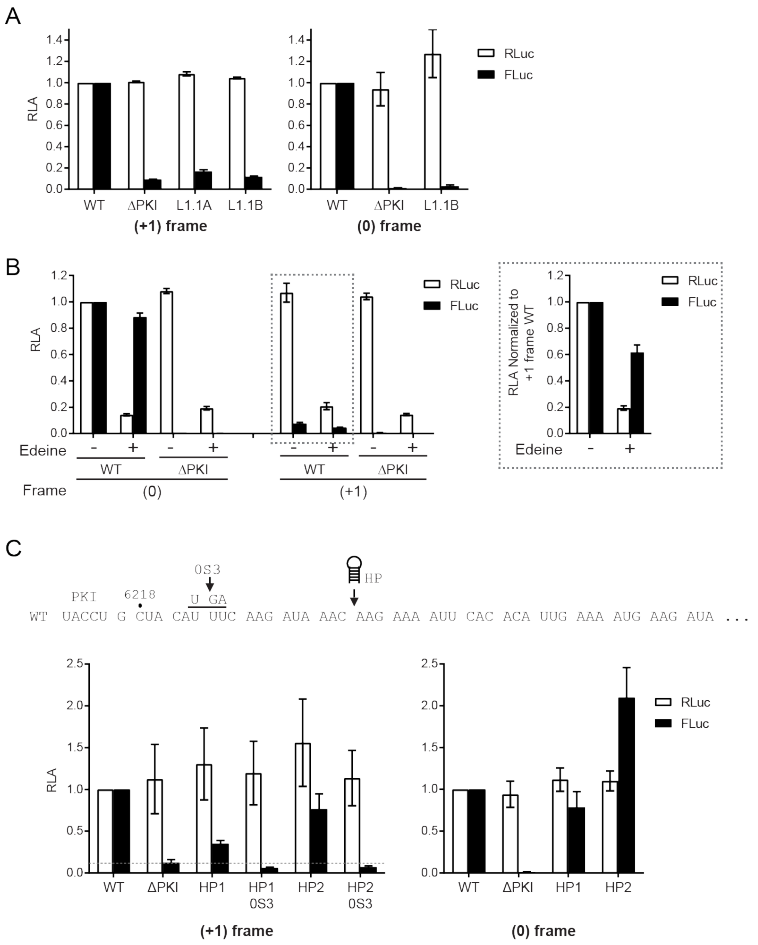


Figure A.4. Translation of CrPV ORFx is dependent on 80S binding but not on scanning-dependent. A) Effect of L1.1 mutation on +1 frame translation**.** Comparison of luciferase activities between WT IGR IRES and L1.1 mutants (L1.1A and L1.1B) that do not bind the 60S subunit. The value of RLuc and FLuc are normalized to WT. **B)** +1-frame translation is insensitive to edeine**.** CrPV IRES-mediated 0- and +1-frame translation is monitored in the absence or presence of 2 μM edeine. The ΔPKI mutants in both 0- and +1-frame are used as controls for IGR IRES dependent translation. The luciferase activities of RLuc and FLuc on the left panel are normalized to the WT of the 0-frame, whereas the luciferase activities of RLuc and FLuc on the right panel are normalized to the WT of the +1-frame. **C)** Effect of RNA hairpin structures on +1 frame translation**.** Translation of 0- and +1-frame translation of bicistronic reporter constructs that contain RNA hairpins structures (HP1 ΔG = -17.00 kcal/mole and HP2 ΔG = -32.70 kcal/mole) between the nucleotides 6135 and 6136. To block the ribosome scanning the downstream sequence driven by the 0-frame initiation, we introduced a stop codon at the 3rd codon on the 0-frame on both HP1 and HP2 constructs (0S3). Shown are averages from at least three independent biological experiments (± SD).

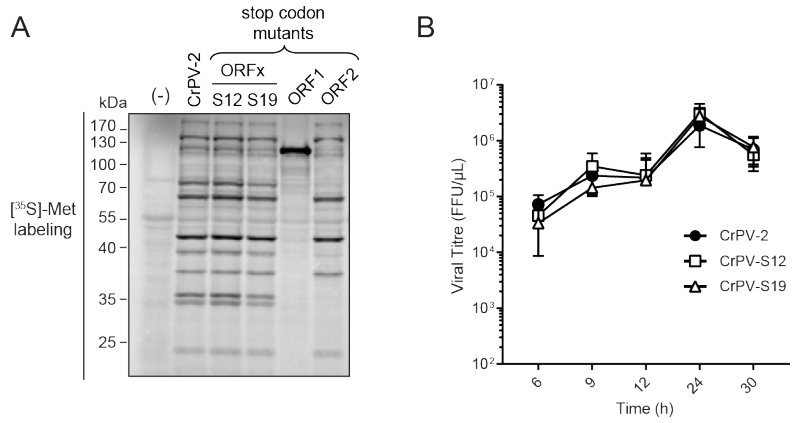


Figure A.5. CrPV ORFx has no effect on virus infection in S2 cells. A) CrPV ORFx has no effect on viral protein synthesis *in vitro*. The *in vitro* transcribed RNA from CrPV-2 or the cognate mutants was incubated in Sf21 cell extract at 30°C for 2 hours in the present of [^{35}S]-methionine/cysteine. Translation of viral proteins were monitored by autoradiography using 12% SDS PAGE. A representative gel of at least three independent biological experiments is shown. **B)** CrPV ORFx has no effect on viral replication in S2 cells. S2 cells were infected with CrPV-2, CrPV-S12, or CrPV-S19 viruses at MOI 1.0. Viral titres were measured at 6, 9, 12, 24 and 30 h.p.i for MOI 1. Shown are average from at least three independent biological experiments (±SD).

Appendix B : CHAPTER 5 SUPPLEMENTARY DATA

Table B.1. Significantly enriched GO terms found in the exosome-like vesicle pellet of from CrPV-infected cells.

Name	GO ID	# Genes	Corrected p value*
aerobic respiration	GO:0009060	13	1.00E-12
ATP biosynthetic process	GO:0006754	7	1.00E-04
ATP metabolic process	GO:0046034	29	1.00E-12
ATP synthesis coupled electron transport	GO:0042773	11	1.00E-12
ATP synthesis coupled proton transport	GO:0015986	7	1.00E-04
carbohydrate derivative metabolic process	GO:1901135	49	4.36E-08
carboxylic acid metabolic process	GO:0019752	52	7.31E-11
cellular lipid catabolic process	GO:0044242	6	1.70E-03
cellular respiration	GO:0045333	25	1.00E-12
coenzyme metabolic process	GO:0006732	14	1.50E-03
cofactor metabolic process	GO:0051186	16	1.00E-04
electron transport chain	GO:0022900	15	1.00E-12
energy derivation by oxidation of organic compounds	GO:0015980	26	1.00E-12
fatty acid beta-oxidation	GO:0006635	5	9.00E-04
fatty acid catabolic process	GO:0009062	6	1.70E-03
fatty acid metabolic process	GO:0006631	8	2.30E-03
fatty acid oxidation	GO:0019395	6	1.70E-03
generation of precursor metabolites and energy	GO:0006091	35	1.00E-12
hydrogen ion transmembrane transport	GO:1902600	16	8.00E-04
hydrogen transport	GO:0006818	17	7.00E-04
ion transport	GO:0006811	37	8.00E-04
lipid catabolic process	GO:0016042	6	1.70E-03
lipid modification	GO:0030258	10	7.00E-04
lipid oxidation	GO:0034440	6	1.70E-03
mitochondrial ATP synthesis coupled electron transport	GO:0042775	11	1.00E-12
mitochondrial electron transport, ubiquinol to cytochrome c	GO:0006122	6	1.00E-04
mitochondrial transport	GO:0006839	3	1.90E-03
mitochondrion organization	GO:0007005	8	1.90E-03
monocarboxylic acid catabolic process	GO:0072329	6	1.70E-03
monocarboxylic acid metabolic process	GO:0032787	19	2.00E-04
monovalent inorganic cation transport	GO:0015672	19	1.00E-03
nucleobase-containing small molecule metabolic process	GO:0055086	41	3.74E-10
nucleoside monophosphate biosynthetic process	GO:0009124	8	1.00E-12
nucleoside monophosphate metabolic process	GO:0009123	30	1.00E-12
nucleoside phosphate metabolic process	GO:0006753	38	1.00E-12
nucleoside triphosphate biosynthetic process	GO:0009142	8	1.00E-12

Table B.1. Continued.

nucleoside triphosphate biosynthetic process	GO:0009142	8	1.00E-12
nucleoside triphosphate metabolic process	GO:0009141	30	1.00E-12
nucleotide metabolic process	GO:0009117	38	1.00E-12
organic acid catabolic process	GO:0016054	9	6.00E-04
organic acid metabolic process	GO:0006082	53	2.20E-10
organophosphate metabolic process	GO:0019637	45	3.12E-10
oxidation-reduction process	GO:0055114	60	1.00E-12
oxidative phosphorylation	GO:0006119	12	1.00E-12
phosphate-containing compound metabolic process	GO:0006796	73	1.59E-06
phosphorus metabolic process	GO:0006793	76	8.35E-08
phosphorylation	GO:0016310	42	2.51E-04
proton transport	GO:0015992	17	7.00E-04
purine nucleoside monophosphate biosynthetic process	GO:0009127	8	1.00E-12
purine nucleoside triphosphate biosynthetic process	GO:0009145	8	1.00E-12
purine nucleoside triphosphate metabolic process	GO:0009144	30	1.00E-12
purine nucleotide biosynthetic process	GO:0006164	11	4.00E-04
purine nucleotide metabolic process	GO:0006163	33	6.12E-10
purine ribonucleotide biosynthetic process	GO:0009152	10	1.00E-04
purine ribonucleotide metabolic process	GO:0009150	32	9.45E-11
purine-containing compound biosynthetic process	GO:0072522	11	4.00E-04
purine-containing compound metabolic process	GO:0072521	35	6.65E-09
respiratory electron transport chain	GO:0022904	13	1.00E-12
ribonucleoside monophosphate biosynthetic process	GO:0009156	8	1.00E-12
ribonucleoside monophosphate metabolic process	GO:0009161	30	1.00E-12
ribonucleoside triphosphate biosynthetic process	GO:0009201	8	1.00E-12
ribonucleoside triphosphate metabolic process	GO:0009199	30	1.00E-12
ribonucleotide biosynthetic process	GO:0009260	10	1.00E-04
ribonucleotide metabolic process	GO:0009259	32	9.45E-11
ribose phosphate metabolic process	GO:0019693	34	7.70E-10
small molecule catabolic process	GO:0044282	9	6.00E-04
small molecule metabolic process	GO:0044281	86	1.00E-12
tricarboxylic acid cycle	GO:0006099	13	1.00E-12
tricarboxylic acid metabolic process	GO:0072350	13	1.00E-12