

**EXPLORING THE EVOLUTION OF A VIRAL INTERNAL RIBOSOME
ENTRY SITE**

by

Xinying Wang

B.Sc., The University of British Columbia, 2018

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Biochemistry and Molecular Biology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2021

© Xinying Wang, 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Exploring the evolution of a viral internal ribosome entry site

submitted by Xinying Wang in partial fulfillment of the requirements for

the degree of Master of Science

in Biochemistry and Molecular Biology

Examining Committee:

Dr. Eric Jan, Department of Biochemistry and Molecular Biology, UBC

Supervisor

Dr. Vivien Measday, Department of Biochemistry and Molecular Biology, UBC

Supervisory Committee Member

Dr. Ivan Sadowski, Department of Biochemistry and Molecular Biology, UBC

Supervisory Committee Member

Dr. Ethan Greenblatt, Department of Biochemistry and Molecular Biology, UBC

Additional Examiner

Abstract

The dicistrovirus intergenic region internal ribosome entry site (IGR IRES) adopts a triple-pseudoknot (PK) structure to directly bind to the conserved core of the ribosome and drive translation from a non-AUG codon. The origin of this IRES mechanism is not known. In this thesis, I describe two studies that attempt to examine how the IGR IRES may have come about. In the first study, I characterized an IGR IRES from a 700-year-old dicistrovirus, named ancient Northwest territories criparavirus (aNCV). From structural prediction of the aNCV IGR sequence and filter binding assays, we showed that the aNCV IGR secondary structure is similar to contemporary IGR IRES structures and could tightly bind to purified human ribosomes. However, there are differences including 105 nucleotides upstream of the IRES of unknown function. We also demonstrated that the aNCV IGR IRES can direct internal ribosome entry in vitro. Lastly, we generated a chimeric virus clone by swapping the aNCV IRES into the cricket paralysis virus (CrPV) infectious clone. The chimeric infectious clone with an aNCV IGR IRES supported translation and virus infection. The characterization and resurrection of a functional IGR IRES from a divergent 700-year-old virus provides a historical framework of the importance of this viral translational mechanism.

In the second study, I have examined candidate RNAs that may have IGR IRES-like properties from the *Drosophila* genome. Previously, we adapted a selective evolution approach to identify RNA elements in the *Drosophila* genome which have IGR IRES-like properties. From the potential candidate RNAs, RNA3, RNA5 and RNA7 showed tight binding to purified human ribosomes. However, in a competition assay only RNA5 could compete with excess wild-type but not mutant CrPV IGR IRESs for ribosomes. However, we demonstrated that RNA5 did not bind

to ribosomal core, as it was accessible by RNase I. Structural predictions were used to identify stemloop (SL) structures of RNA5. Mutations at SL2 altered RNA5 binding affinity, suggesting a potential interaction region. Finally, incubation of RNA5 in rabbit reticulocyte lysate (RRL) did not affect translation *in vitro*.

Lay Summary

Translation is a key step of gene expression in all organisms. Viruses have evolved mechanisms to translate viral proteins. One such mechanism is through viral structural element to bypass the translation regulation. By studying the mechanisms in which different viruses translate, the knowledge we gain can be applied to boost the translation of proteins of interest under stress regulations. The translation mechanism from dicistroviruses, which have two coding regions in their genomes, is well-studied as it requires the most simplified mechanism. However, the origin of this mechanism is unclear. This thesis explored the evolution of this mechanism by characterizing the viral structural element of a virus from the same family which is at least 700-year-old and comparing with contemporary models. We also searched for similar element from the host genome, which would solidify host-virus coevolution theory.

Preface

A version of Chapter 2 has been published: Wang, X.; Vlok, M.; Flibotte, S.; Jan, E. Resurrection of a Viral Internal Ribosome Entry Site from a 700 Year Old Ancient Northwest Territories Cripavirus. *Viruses* **2021**, *13* (3), 493. I conducted all the experiments and data analysis presented in this chapter. The manuscript was written with guidance from Dr. E. Jan.

Chapter 3 is based on the original screen conducted by Dr. Q.S. Wang. Experiments for Figures 3.1 were performed by Dr. Q.S. Wang, and I conducted and analyzed the rest of the experiments.

Table of Contents

Abstract.....	iii
Lay Summary	v
Preface.....	vi
Table of Contents	vii
List of Tables	xi
List of Figures.....	xii
List of Abbreviations	xiv
Acknowledgements	xvi
Dedication	xviii
Chapter 1: Introduction	1
1.1 Overview of positive-sense single-stranded RNA viruses.....	1
1.1.1 Positive-sense single-stranded RNA viruses	1
1.1.2 Viral life cycle.....	2
1.2 Eukaryotic cap-dependent translation initiation	6
1.3 Non-canonical translation initiation.....	10
1.3.1 Viral internal ribosome entry site	10
1.3.2 Cellular internal ribosome entry site.....	15
1.4 <i>Dicistroviridae</i> family of viruses	18

1.4.1 Genome characteristics and organization	19
1.4.2 Viral-host interaction	23
1.4.3 Intergenic region internal ribosome entry site	23
1.4.3.1 Mechanism of IGR IRES-mediated translation initiation.....	26
1.4.3.2 Significance of IGR IRES.....	30
1.5 Thesis investigation	31
Chapter 2: Resurrection of an IGR IRES from ancient Northwest Territories cripavirus	33
2.1 Introduction.....	33
2.2 Materials and methods	34
2.2.1 Cell culture.....	34
2.2.2 Virus infection	34
2.2.3 DNA constructs.....	34
2.2.4 CrPV/aNCV chimeric infectious clone constructs	35
2.2.5 In vitro transcription and translation.....	35
2.2.6 Purification of the 40S and 60S subunits.....	36
2.2.7 Filter-binding assays	36
2.2.8 Ribosome protection assay	37
2.2.9 Toeprinting/Primer extension analysis	38
2.2.10 RNA transfection	39
2.2.11 RT-PCR and sequence confirmation	39
2.2.12 Western blotting.....	39
2.3 Results.....	40

2.3.1 aNCV IGR IRES adopts a triple pseudoknot structure.....	40
2.3.2 aNCV IGR IRES binds tightly to human 80S ribosomes	43
2.3.3 RNase protection analysis of aNCV IGR-ribosome complexes	46
2.3.4 Determination of the aNCV IGR IRES initiation site	48
2.3.5 aNCV IGR IRES directs translation in vitro.....	51
2.3.6 Chimeric CrPV clone containing the aNCV IGR is infectious.....	53
2.4 Discussion	56
Chapter 3: The search for Drosophila RNAs that have viral IGR IRES-like properties	59
3.1 Introduction.....	59
3.2 Materials and methods	61
3.2.1 Plasmids and constructs	61
3.2.2 <i>In vitro</i> transcription	61
3.2.3 Purification of the 40S and 60S subunits.....	62
3.2.4 Filter-binding assays	62
3.2.5 Ribosome protection assay	63
3.2.6 Translation competition in RRL	64
3.3 Results.....	64
3.3.1 SELEX screen identifies potential RNA candidates that can bind directly to ribosomes	64
3.3.2 Ribosome binding of candidate RNAs	70
3.3.3 Ribosome competition assays	74
3.3.4 Mutations in RNA5 alter its binding affinity to ribosomes	74

3.3.5 None of the candidates show protection by ribosomes from RNase I.....	78
3.3.6 RNA5 does not affect translation level in vitro	80
3.4 Discussion	82
Chapter 4: Conclusions and Future Directions	86
References	92

List of Tables

Table 3.1. Genomic information of candidates being tested from ribosome binding assay. 72

List of Figures

Figure 1.1. Overview of the poliovirus replication cycle.	5
Figure 1.2. Pathway of eukaryotic cap-dependent translation initiation.	9
Figure 1.3. Classification of viral internal ribosome entry sites.	14
Figure 1.4. Genome organization of dicistroviruses.	20
Figure 1.5. Morphology and virion structure of dicistroviruses.	22
Figure 1.6. Dicistroviruses intergenic region internal ribosome entry site.	25
Figure 1.7. A pathway model of IGR IRES-mediated translation initiation.....	29
Figure 2.1. Structure information of aNCV IGR IRES.	42
Figure 2.2. Affinity of 80S-aNCV IGR IRES complexes.....	45
Figure 2.3. aNCV IGR IRES bound to 80S is RNase I-resistant.....	47
Figure 2.4. Toeprinting analysis of 80S-aNCV IRES complexes.....	50
Figure 2.5. In vitro translation assays in RRL.	52
Figure 2.6. Chimeric CrPV-aNCV clone is infectious in <i>Drosophila</i> S2 cells.....	55
Figure 3.1. Selection IGR IRES-like RNA elements in genomic RNA library.....	67
Figure 3.2. Summary of SELEX screen.....	69
Figure 3.3. Affinity of 80S-RNA candidate complexes.	73
Figure 3.4. Alignment of RNA5 within <i>Drosophila</i> and other species.	76
Figure 3.5. Mutations at certain stem and loop regions of RNA5 affect its ribosome binding affinity.....	77
Figure 3.6. RNA candidates bound to the 80S subunit are not resistant to RNase I.....	79

Figure 3.7. Incubation of RNA5 in an in-vitro translation reaction of bicistronic reporter has no impact on reporter RNA translation in RRL. 81

List of Abbreviations

aNCV	ancient Northwest Territories Cripavirus
ASL	anticodon stem loop
CBC	cap binding complex
CRE	cis-acting replication element
CrPV	Cricket paralysis virus
DCV	Drosophila C virus
EMCV	Encephalomyocarditis virus
ER	endoplasmic reticulum
FLuc	Firefly luciferase
HaIV	Halastavi árva virus
HCV	hepatitis C virus
HDV	hepatitis D virus
h.p.i.	hours post-infection
h.p.t.	hours post-transfection
IAPV	Israeli acute paralysis virus
IGR	intergenic region
IRES	internal ribosome entry site
ITAF	IRES <i>trans-acting</i> factor
kb	kilobase
ncRNA	non-protein-coding RNA
NMD	nonsense mediated decay

ORF	open reading frame
PABP	poly(A) binding protein
PIC	preinitiation complex
PK	pseudoknot
PV	Poliovirus
RdRp	RNA dependent RNA polymerase
RhPV	Rhopalosiphum padi virus
RLuc	Renilla luciferase
rp	ribosomal protein
RRL	rabbit reticulocyte lysate
rRNA	ribosomal RNA
sf-21	<i>Spodoptera frugiperda</i>
SL	stem loop
+ssRNA	single-stranded positive-sense RNA
UTR	untranslated region
VP	virion protein
VPg	viral protein genome-linked
VRL	variable loop
WT	wild-type

Acknowledgements

First and foremost, I would like to thank my supervisor Dr. Eric Jan for providing me the opportunity to work in his laboratory and conduct this research. I am extremely grateful for his patience, guidance and mentorship throughout my entire graduate life. I will never forget that when I just started my graduate research and had trouble with an experiment for the entire term, I started to question and doubt myself whether I had the ability to do the research especially I came with a bachelor's degree of different background. It is Eric who keeps encouraging me and providing mental support which make me confident and eventually establish myself as an independent scientist.

To the Jan lab members. I am proud that I am part of this big family, and definitely the best lab ever! Thank you all for your guidance and support from the past three years. I will always remember the “happy time” every Friday and those lovely birthday celebrations. Jibin Sadasivan, you are my colleague, my friend, but also my mentor in this lab. You can always provide useful suggestions about my experiments and train me on so many experiments. I know I never show my appreciation in front of you, but here I really want to give a big thank to you. Without you, I will not be able to grow so fast during my graduate life, and all of the knowledge I learned from you are precious. Reid Warsaba, your attitude to life is always positive which definitely changes my attitude when facing troubles. Thanks for always bringing laughs to the lab. Ciara Stevenson, I am going to miss the discussions about dressing and other adorable things with you and hope you can graduate very soon as well. Dr. Marli Vlok, thanks for the past help from you and those kind reminders, preventing me from getting into trouble. Daniel Andrews, you always give the most detailed training and provide the most comprehensive plan. Yihang Chen, I love those desserts

from you. I believe you will become experienced and also be a confident girl. To the new members of the lab – Brenna Hay, Amanda Rahardjo, Dr. Katherine Badior and Dr. Corina Stewart – all of you are incredible and your enthusiasm for science encourages me.

Lastly, I would like to thank my parents for unconditional love and support throughout my life, without whom I would never have enjoyed so many opportunities. Thank you for giving me strength to chase my dreams and do whatever I want. I love you both.

For my Mother and Father. Enjoy.

Chapter 1: Introduction

1.1 Overview of positive-sense single-stranded RNA viruses

1.1.1 Positive-sense single-stranded RNA viruses

Viruses are obligate pathogens that replicate inside living cells, from mammals, plants to microorganisms. Viruses are classified into seven groups based on their types of genetic information (RNA or DNA), strandedness, and the method of replication (Baltimore 1971). Single-stranded positive-sense RNA (+ssRNA) viruses comprise one of the largest class of viruses having RNA genome sizes ranging from 2 kilobases (kb) genome (porcine circovirus) (Dhindwal et al. 2019) to 32 kb genome (coronavirus) (Snijder et al. 2003; D. Kim et al. 2020).

The single-stranded viral RNA genomes function as both the replication template and as a template for translation to produce viral proteins. +ssRNA viruses replicate through RNA dependent RNA polymerase (RdRp) that is virally encoded. Because RdRp lacks proof-reading ability (Venkataraman, Prasad, and Selvarajan 2018), viruses replicate with high mutation rates and exhibit genetic diversity. Coronaviruses are the exception, and the increased fidelity of their replication is mediated by the exoribonuclease encoded from nonstructural protein 14 (Minskaia et al. 2006; Gribble et al. 2021). For +ssRNA viruses, the mutation rate was examined to as high as 10^{-4} mutation per nucleotide in a single infection (Sanjuán et al. 2010). Thus, these viruses adapt rapidly when environmental dynamics change and evolve resistance to selective pressures such as vaccines and antiviral drugs, therefore it is a constant challenge to develop new antiviral therapeutic strategies (Lauring and Andino 2010). Despite the high mutation rate, understanding

the fundamental mechanisms of the viral life cycle will provide potential targets for developing antiviral drugs.

Due to the limited compact genome size, +ssRNA viruses contain single or multiple open reading frames (ORFs) that are translated to polyproteins. The polyproteins are then processed by viral or cellular proteases or via noncanonical translation mechanisms such as the 2A peptide "stop-go" mechanism (Luke et al. 2008). The ORFs encode viral proteins required for different steps of the viral life cycle. Viral non-structural proteins include the viral protease that cleaves both viral and host substrates, viral protein genome-linked (VPg) that participates in viral replication and protects viral mRNA from degradation (Goodfellow 2011), RdRp, and other essential proteins involved in both translation and replication. Viral structural proteins include capsid and envelope proteins that package and protect the viral genome, transport and release their genomes inside other cells (Roos et al. 2007). Depending on characteristics such as genome composition, host range, sequence similarity and virion structures, viruses with similar properties are grouped following a hierarchical relationship which includes order, family, genus and species. For the purpose of this study, I will be focusing on the member of *Dicistroviridae*, which is classified with the family of *Picornaviridae*. Its genome composition as well as protein functions will be discussed more in **section 1.4.1**.

1.1.2 Viral life cycle

Most +ssRNA viruses share a similar viral life cycle, including binding to cell surface receptors for viral entry, followed by uncoating and release of the viral RNA genome into the cytoplasm. Subsequently, the viral RNA is translated to produce viral proteins, which in turn initiate the replication step to produce the -strands and +strands. Viral RNAs are then assembled with structural capsid proteins and egress out of the cell either through lysis or budding (reviewed

in (De Jesus 2007)). Here, the poliovirus (PV) life cycle will be used as the model to be reviewed because it is one of the most well-studied viruses in this group, and dicistroviruses are part of *Picornavirales* which are likely to have a similar life cycle.

As a member of *Picornaviridae*, PV is a non-enveloped virus composed of an RNA genome and capsid proteins. The receptor for PV entry is the cell surface receptor CD155, which is an immunoglobulin-like receptor (Mendelsohn, Wimmer, and Racaniello 1989). Upon binding to CD155, the viral particle of PV undergoes a conformational change which is necessary for viral entry through receptor mediated endocytosis (de Sena and Mandel 1977). This conformational change externalizes the capsid protein VP4 and the N terminus of VP1 (M. Chow et al. 1987; Fricks and Hogle 1990) and the externalized peptides insert into membranes to anchor to the cell membrane in a receptor-independent manner (Tosteson and Chow 1997; Brandenburg et al. 2007). As a result, pores are formed on the cellular membrane and the viral RNA genome is released into the cytoplasm.

PV possesses an RNA genome with a single ORF which is ~7.5 kb in length. After its genomic RNA is released into the cytoplasm, the covalently linked VPg is removed from the 5' end of PV RNA by a cellular unlinkase (Nomoto, Fon Lee, and Wimmer 1976; Lee et al. 1977; Virgen-Slane et al. 2012). Instead of having a 5' m⁷G-cap structure to recruit the host translational machinery, PV possesses an internal ribosome entry site (IRES) within the 5' untranslated region (5'UTR) to initiate translation. More detailed mechanisms of cap-dependent translation and IRES-dependent translation will be discussed in **section 1.2** and **1.3**. After translation, the nascent polypeptide is produced and then cleaved into about 10 individual proteins by virally encoded proteases, 2A and 3C/3CD. The 2A and 3C/CD also target host proteins to modulate cellular processes that promote infection. Notably, translation factors eIF4G and poly(A) binding protein

(PABP) are cleaved thereby resulting in inhibition of cellular translation (Gradi et al. 1998; Kuyumcu-Martinez et al. 2004). After viral proteins are synthesized, translation shifts to replication to produce more RNAs. During PV replication, the intracellular membrane landscape rearranges massively to provide vesicles for RNA replication (Belov et al. 2012), induced by 2C and 2BC which tightly associate with PV membranous replication complexes (Cho et al. 1994). PV replication is thought to happen at endoplasmic reticulum (ER) because viral vesicles were found to colocalize with ER marker proteins (Bienz, Egger, and Pasamontes 1987; Suhy, Giddings, and Kirkegaard 2000), and the PV-induced vesicles most resemble the ER (Suhy, Giddings, and Kirkegaard 2000). PV is a +ssRNA virus, so (-)RNA needs to be replicated to synthesize more (+)RNA. VPg is covalently linked to (+)RNA which functions as a primer for both negative- and positive-strand synthesis (Vogt and Andino 2010). To generate (-)RNA, VPg is uridylylated by addition of two uracil nucleotides on its third tyrosine residue by an cis-acting replication element (CRE). CRE contains a highly conserved secondary structure bedded within the 2C-coding region and together with uridylylated VPg to synthesize (-)RNA (Vogt and Andino 2010). After (-)RNA is synthesized, it serves as the template to produce (+)RNA which serves as both the translating template and genomic RNA. At the last step, structural proteins (capsid VP1, VP2, VP3, VP4) produced by viral translation encapsulate viral RNA genomes and generate progeny virions. Membrane lysis is believed to be the main method to release virions from the cell, though a different mechanism involving extracellular vesicles has been suggested to release virions (Bird and Kirkegaard 2015; Y. H. Chen et al. 2015). (**Figure 1.1**)

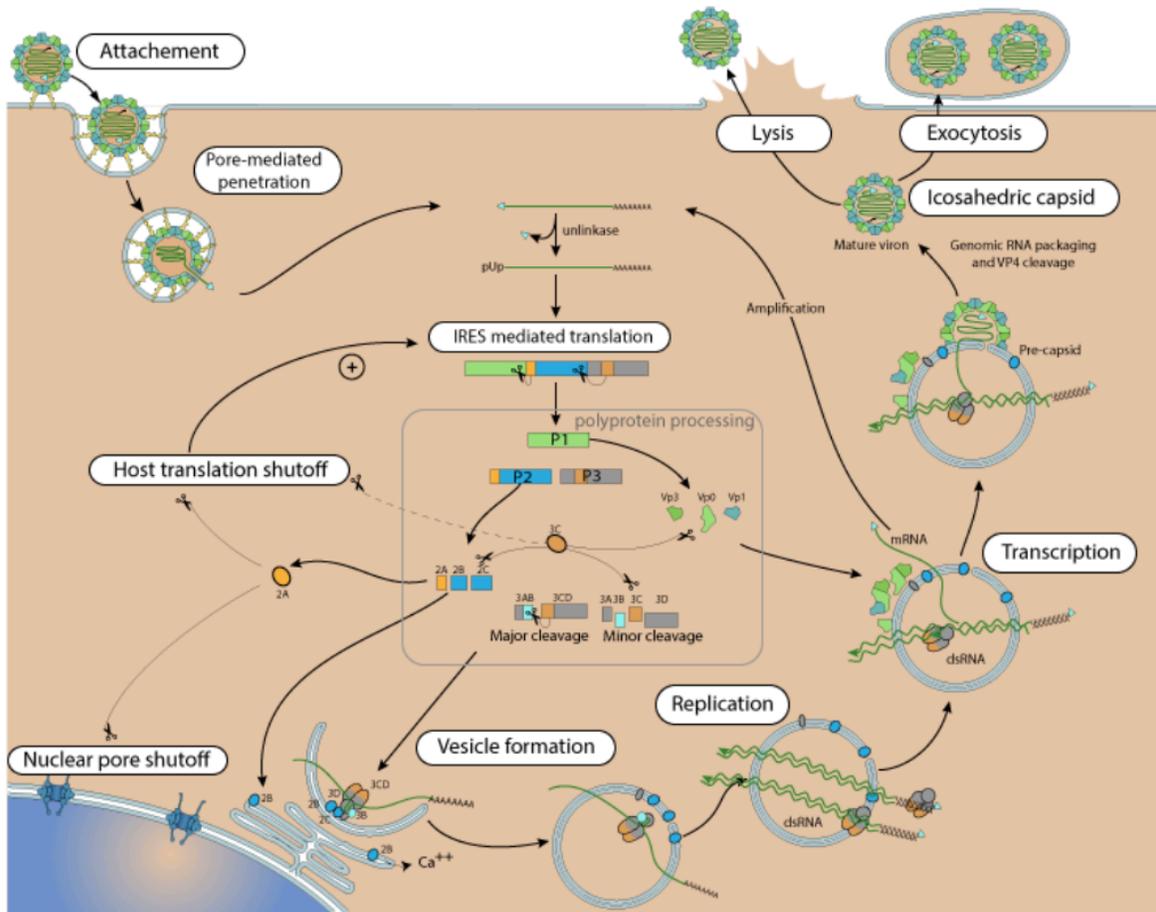


Figure 1.1. Overview of the poliovirus replication cycle.

Poliovirus attaches to the cell membrane through binding to CD155 receptor. After receptor-mediated endocytosis, viral genomes are released into the cytoplasm. The released genome is translated into a polyprotein, processed into individual proteins. Viral RNA replicates at membrane vesicles together with replication complexes. Encapsidation happens closely coupled to replication to package newly synthesized genomic RNA into capsids. Finally viruses are released via lysis or exocytosis. Adapted from ViralZone (<https://viralzone.expasy.org>; Swiss Institute of Bioinformatics).

1.2 Eukaryotic cap-dependent translation initiation

Translation is a key step of gene expression in all organisms and can be divided into four stages: initiation, elongation, termination and ribosome recycling. Initiation is typically the rate limiting step of translation and requires complex step-by-step actions to start translation. The majority of eukaryotic mRNAs use a scanning cap-dependent mechanism that requires upwards of 12 core translation initiation factors to initiate translation. Two hallmarks of eukaryotic mRNAs are the 5' cap and 3' poly(A) tail that both play an important role during translation. The 5' cap has three main functions: 1) prevents mRNAs from degradation by exoribonucleases. 2) promotes pre-mRNA splicing (Edery and Sonenberg 1985; Konarska, Padgett, and Sharp 1984; Fresco and Buratowski 1996). 3) regulates nuclear export of mRNAs (Lewis and Izaurralde 1997). The 5' cap, consisting of a guanine nucleotide, is connected to the first nucleotide of mRNA by an inverted 5'-5' triphosphate linkage (reviewed in (Shuman 2002)). The cap is added co-transcriptionally when the nascent mRNA is generated and then guanosine is methylated at position 7, abbreviated as m⁷G. By blocking the 5' end, the 5' cap can stabilize mRNAs from exoribonucleases degradation. Cap binding complex (CBC) regulates RNA nuclear export by recognizing and binding to the 5' cap. Then CBC is recognized by the nuclear pore complex and exported to support a pioneer round of mRNA translation essential for mRNA quality control (reviewed in (Maquat, Tarn, and Isken 2010)). Subsequently, CBC is replaced by the translation initiation factor, eIF4E, and the new complex is recognized by other translation initiation machinery to start the translation (Maquat, Tarn, and Isken 2010).

The translation initiation starts by recruiting the cap-binding complex, eIF4F and the formation of the 43S preinitiation complex (PIC) (reviewed in (Jackson, Hellen, and Pestova 2010;

Hinnebusch and Lorsch 2012) (**Figure 1.2**). eIF4F comprises the cap-binding protein eIF4E which binds to the 5' cap, the DEAD-box RNA helicase eIF4A and eIF4G which is a scaffold protein mediating the circularization of mRNA by interacting with poly(A) binding protein (PABP) bound to the 3' poly(A) tail and eIF4E. The 43S PIC consists of 40S ribosomal subunit, the ternary complex eIF2-Met-tRNA_i-GTP, eIF1, eIF1A, eIF3 and eIF5. The 43S PIC is recruited to the 5' end through interaction between eIF4G and eIF3 and then scans along the mRNA until it encounters the first AUG codon. Scanning is assisted by the helicase eIF4A which unwinds the RNA secondary structure in the 5' UTR, together with eIF1 and eIF1A which are essential for recognizing the AUG start codon by inducing an “open” conformation of the 40S at the binding channel (Passmore et al. 2007). To ensure the fidelity of initiation, scanning complexes must have discrimination ability to screen and bypass partial base-pairing of initiation codon triplets and recognize the correct site (Yu et al. 2009; Simonetti et al. 2020). An optimal initiation site is within the “Kozak” sequence, which has an optimum context GCC(A/G)CCAAUGG, with a purine in -3 and a guanine in +4 positions (relative to the A of the AUG codon, designated as +1) (Kozak 1986; 1991). Once the start codon is recognized and Met-tRNA_i establishes the anticodon-codon base-pairing with AUG start codon at the ribosomal P site, eIF1 is displaced followed by the ribosome complex adopting a scanning-incompetent “closed” conformation change (Maag et al. 2005; Passmore et al. 2007). The arrested ribosomal complex commits to that initiation codon, and this step is mediated by eIF5, a GTPase-activating protein, which specifically stimulates the hydrolysis of eIF2-bound GTP and therefore decreases the binding affinity of eIF2-GDP and results in its dissociation from the ribosomal complex (Paulin et al. 2001; Kapp and Lorsch 2004). The dissociation of eIF1, eIF1A, eIF3 are mediated by eIF5B, a ribosome-dependent GTPase, which is proposed to promote 60S subunit joining (Pestova et al. 2000). The release of eIF5B is achieved

by hydrolyzing eIF5B-bound GTP. At last, an 80S ribosome is formed at the translational start site, with the Met-tRNA_i positioned in the P site and an adjacent vacant A site ready to accommodate the next aminoacyl-tRNA, ready for the elongation.

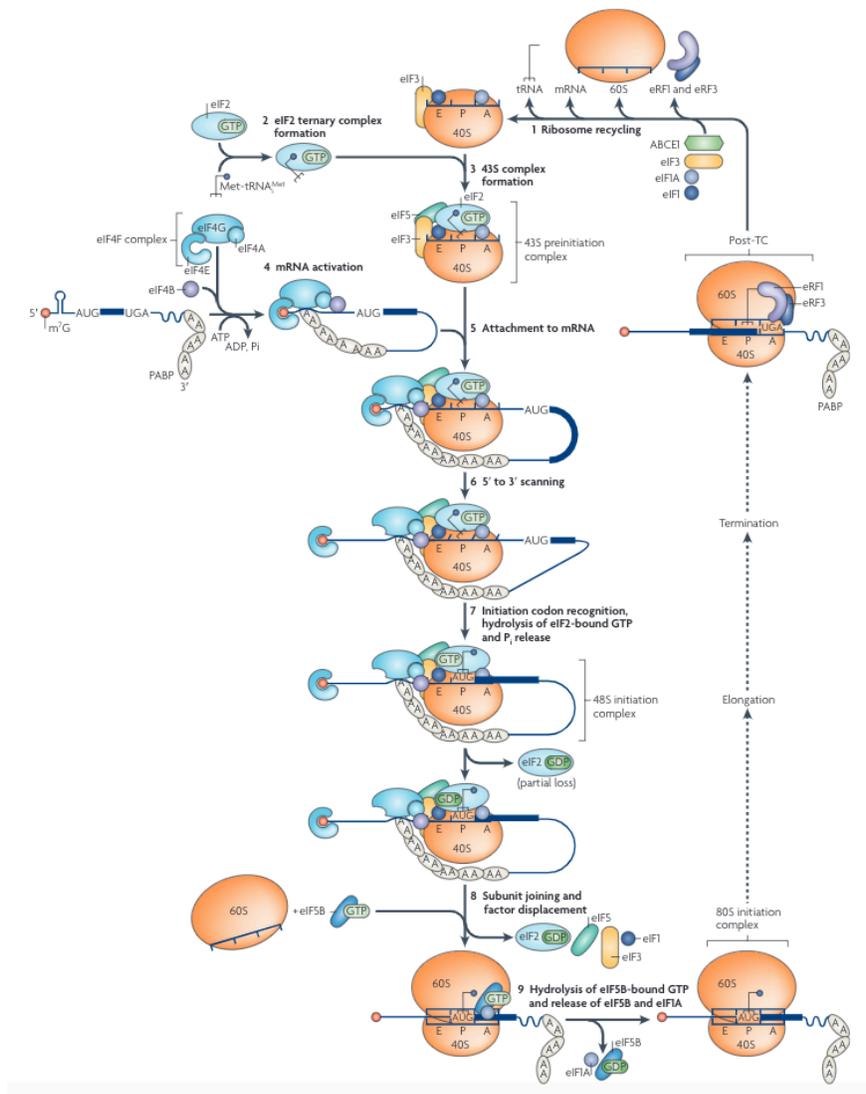


Figure 1.2. Pathway of eukaryotic cap-dependent translation initiation.

The eukaryotic translation initiation pathway is divided into eight stages (2-9) followed as: (2) eIF2-Met-tRNA_i-GTP ternary complex formation; (3) 43S preinitiation complex formation; (4) mRNA activation by binding of eIF4F; (5) attachment of 43S complex to the mRNA cap region; (6) 5' to 3' scanning of the 5' UTR by 43S complex; (7) initiation codon recognition followed by 48S initiation complex formation, switching the scanning complex to a “closed” conformation; (8) 60S subunits joining to 48S complex and initiation factors displacement; (9) eIF5B-bound GTP hydrolysis and release of eIF1A and eIF5B to form elongation-competent 80S ribosomes. Adapted with permission from (Jackson, Hellen, and Pestova 2010).

1.3 Non-canonical translation initiation

Viruses have evolved noncanonical mechanisms for viral protein synthesis including but not restricted to leaky scanning, which allows for expression of multiple isoforms of a protein; ribosome shunting, which allows ribosomes to access downstream ORFs in a scanning-independent pattern; reinitiation, which encodes multiple ORFs without the dissociation of 40S subunit; internal ribosome entry site (IRES), which allows internal entry of ribosomes on a mRNA in a 5' cap-independent manner (reviewed in (Firth and Brierley 2012)). These mechanisms allow for the virus to co-opt the ribosomes and may allow preferential viral protein synthesis under virus infection. For the purpose of this study, IRES as one of the viral strategies to initiate translation will be discussed in detail.

1.3.1 Viral internal ribosome entry site

Most of the IRESs discovered are from viruses. IRESs allow viruses to subvert host translation machinery to synthesize viral proteins. Based on the sequences, secondary structures and the mode of action for translation initiation, IRESs are classified to four different classes (**Figure 1.3**). Class 1 and 2 IRESs generally have longer sequences with RNA structures consisting of basic short and long hairpins and tertiary structures such as pseudoknots. These IRESs require all initiation factors except eIF4E as well as protein cofactors. Class 1 IRESs still require a scanning process whereas Class 2 IRESs do not. Class 3 IRESs bind directly to the ribosome but still require a few eIFs. Class 4 IRESs have the most streamlined mechanism which does not require any eIFs to bind to ribosomes and initiate at a non-AUG codon (reviewed in (Jaafar and Kieft 2018; Mailliot and Martin 2018)).

Class 1 and 2 IRESs: Typical examples of these two types of IRESs are from *Picornaviridae* and are similar to one another in terms of the requirement to initiate translation. The first studied models poliovirus (PV) and encephalomyocarditis virus (EMCV) that containing IRESs are from Class 1 and 2 respectively (S. K. Jang et al. 1988; Pelletier and Sonenberg 1988). Both viruses lack a conventional 5' m⁷G cap and instead contain highly structured 5' UTRs. Their 5' UTRs internal initiation ability were determined by inserting the 5' UTR between two reporter luciferases within a bicistronic reporter construct. Indeed, the translation directed by viral 5' UTR occurred, demonstrating that internal translation initiation is possible. This discovery was further proved by artificially synthesizing a circularized RNA harboring the EMCV which can initiate translation (C. Y. Chen and Sarnow 1995). Later, IRESs were found in other picornaviruses including foot-and-mouth disease virus (Belsham and Brangwyn 1990; Kühn, Luz, and Beck 1990), human rhinoviruses (Borman and Jackson 1992), and hepatitis A virus (Glass, Jia, and Summers 1993). Both classes of IRESs require the entire set of canonical eIFs excluding eIF4E and also require IRES *trans-acting* factors (ITAFs). They are unable to recruit 40S subunit directly. Although most of the protein complements associated with IRES are not clear, ITAFs are thought to be important for maintaining an active IRES by promoting IRES remodeling in an active structure state (reviewed in (Plank and Kieft 2012)). Class 1 members need to undergo a classical 5'-3' scanning to find the start codon, whereas Class 2 members can tether translation directly to the start codon without any scanning. Both classes of viruses are partially refractory to eIF2 phosphorylation (Meurs et al. 1992) but can operate in an eIF2-independent mode by using eIF5B as a substitute (White, Reineke, and Lloyd 2011). Taking IRES-driven translation as an advantage, viruses containing Class 1 or 2 IRESs depress cap-dependent translation by targeting and cleaving eIF4G and PABP by the viral 2A^{pro} and 3C^{pro} proteases (Gradi et al. 1998; Etchison et al. 1982;

Joachims, Van Breugel, and Lloyd 1999; Kräusslich et al. 1987). For example, PV cleaves eIF4G using the 2A^{pro} protease and abolishes *de novo* cellular translation initiation (Gradi et al. 1998). PV can bypass this abolishment by recruiting the carboxy-terminal fragment of eIF4G that interacts with eIF3 and retain translation activity (Ohlmann et al. 1996; Sweeney et al. 2014).

Class 3 IRESs: This type of IRESs is found in *Flaviviridae* family and classical swine fever virus. Most of the current knowledge relies on hepatitis C virus (HCV) from *Flaviviridae* family. The IRES from HCV can directly bind to the 40S subunit with only a few subsets of initiation factors including eIF3 and eIF2 (Pestova et al. 1998). The IRES binding to the 40S subunit positions the P site AUG start codon within the decoding region without any scanning process (Pestova et al. 1998). The position of eIF3 in the IRES-40S-eIF3 complex is different from its normal binding condition (Hashem et al. 2013), which is hypothesized to help the IRES access the 40S subunit or participate in the remodeling step (Jaafar et al. 2016). Class 3 IRES initiation is proposed to have two ways of delivering Met-tRNA_i. In most cases, if eIF2 is available, it delivers Met-tRNA_i. When eIF2 is not available such as phosphorylation of eIF2 α , it is proposed that Met-tRNA_i can be delivered by alternative factors including eIF5B and eIF2A (Terenin et al. 2008). eIF5B, the homolog of IF2 which delivers initiator tRNA in prokaryotes, can recognize the methionylated acceptor stem of initiator tRNA and delivers it to the ribosome (Kuhle and Ficner 2014). eIF2A mediates viral translation via directly interacting with HCV IRES followed by Met-tRNA_i loading, and knocking down of eIF2A reduced viral infectivity (J. H. Kim et al. 2011). Once Met-tRNA_i is delivered, 60S subunit is directly recruited to form 80S ribosome and start translation.

Class 4 IRESs: IRESs from this class have been found exclusively in the *Dicistroviridae* family, of which the genomic RNA contains two ORFs separated by an intergenic region (IGR).

The IGR possesses the most streamlined IRES mechanism to date and initiates translation of the downstream ORF. IGR IRESs dispense the need for all canonical translation factors and bind directly to the 40S subunit followed by 60S subunit joining (Jan and Sarnow 2002; J E Wilson, Pestova, et al. 2000). Remarkably, Class 4 IRESs mediate translation initiation at a non-AUG codon (J Sasaki and Nakashima 1999; Jun Sasaki and Nakashima 2000; Jan and Sarnow 2002; J E Wilson, Pestova, et al. 2000). The detailed mechanism of how IGR IRESs recruit and direct translation will be discussed in **section 1.4.3**. All IGR IRESs are ~150-200 nucleotides in length and are able to direct translation in a variety of cell types including yeast, plant, insect and mammalian systems (Colussi et al. 2015; Thompson, Gulyas, and Sarnow 2001; J E Wilson, Powell, et al. 2000), showing that their translation mechanism is very robust.

In general, as one of the strategies to bypass the cellular translation regulation, IRESs allow viruses to synthesize their own proteins while global host translation is repressed, following the increasing accessibility of free cellular ribosomes and translation factors for viral translation.

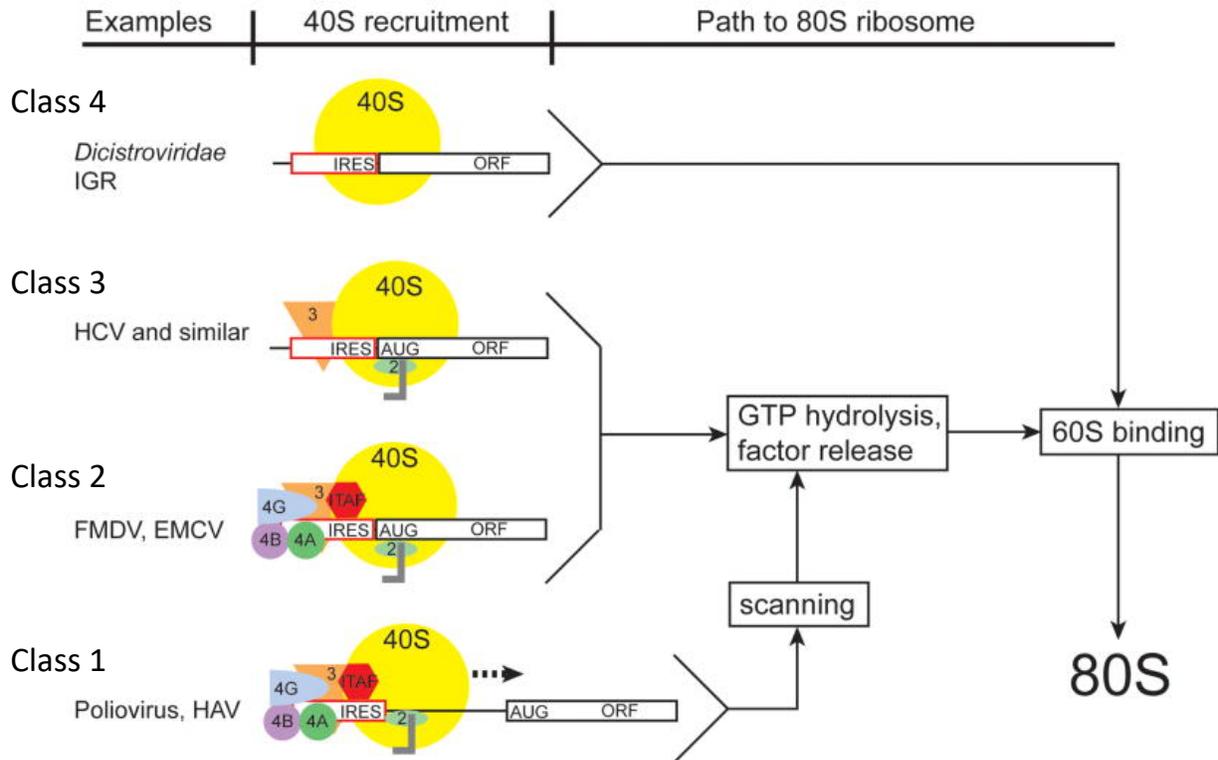


Figure 1.3. Classification of viral internal ribosome entry sites.

Viral IRESs are classified to four classes. Class 1 IRESs require all canonical initiation factors except eIF4E and *trans*-acting factors to recruit the ribosome. Ribosomes must undergo scanning to initiate at AUG start codon site. Class 2 IRESs require all factors as those of Class I IRESs but can tether translation directly to the start codon without any scanning. Class 3 IRESs require only eIF2 and eIF3 and can directly bind the 40S ribosomal subunit and initiate at AUG codon site. Class 4 IRESs bind directly to the 40S ribosomal subunit without any factors and initiate at non-AUG codon site. Adapted with permission from (Plank and Kieft 2012).

1.3.2 Cellular internal ribosome entry site

Although studies on viral IRESs are extensive, IRES-mediated mechanisms in cellular transcripts are less understood. The first cellular IRES was found within the mRNA encoding the immunoglobulin heavy chain binding protein under PV infection (Sarnow 1989; Johannes and Sarnow 1998). eIF4GI, an isoform of eIF4G, is another example of being identified to contain an IRES element during PV infection (Johannes and Sarnow 1998). Identifying bona fide cellular IRESs has been challenging. One of the reasons is that cellular IRES-mediated translation is usually less efficient compared to viral IRES-mediated translation. For example, during mitosis, when the global mRNA translation is reduced, though the nucleophosmin mRNA contains IRES element, its translation level is also reduced, but only half of global translation reduced level (Qin and Sarnow 2004). Similarly, c-myc is also known as to contain an IRES in its 5' UTR. During apoptosis, the translation level of c-myc decreases as well, but the decrease is delayed by 2 h compared to the global translation repression (Bushell et al. 2006). Another key factor is that cellular IRES-mediated translation does not happen all the time. Instead, it is activated under certain conditions. Increasing evidence suggests that these cellular IRESs have two main physiological functions: 1) They only support a low level of translation initiation when cap-dependent translation is active. 2) Cellular IRES-mediated translation is robust when the cap-dependent translation is depressed under a variety of stress conditions such as mitosis, virus infection (reviewed in (Komar and Hatzoglou 2011)). It has also been shown that both cap-dependent and IRES-mediated translation can operate on the same mRNA. For example, the mRNA for synthesizing neurogranin, a neuronal calmodulin-binding protein, can be translated through both 5' cap-dependent and internal initiation mechanisms (Pinkstaff et al. 2001).

Unlike viral IRESs which share structure similarity within each class, cellular IRESs are much more diverse and do not contain a shared structure nor a similar sequence. Chemical and enzymatic probing on a subset of cellular IRESs revealed complex structures including stem loops and pseudoknots (Le Quesne et al. 2001; Bonnal et al. 2003; Mitchell et al. 2003; Yaman et al. 2003; Jopling et al. 2004; Martineau et al. 2004). However, the detailed mechanism of promoting internal entry of ribosomes by these cellular IRESs is still unclear. The majority of cellular IRESs are located in the 5' UTR immediately upstream of the initiation codon, but cases in which IRESs are located in the coding region also exist (Komar and Hatzoglou 2011). This results in a truncated version of protein with alternative functions (Komar et al. 2003; Grover et al. 2009).

There have not been extensive systematic studies on cellular IRES-mediated translation mechanism in terms of the ability to recruit the ribosome complex. However, given that these cellular IRESs are highly structured at the 5' UTR, it is unlikely to have the conventional scanning mechanism from the 5' end (Hellen and Sarnow 2001; Stoneley and Willis 2004). Alternatively, IRESs from c-myc, L-myc, and N-myc mRNAs were suggested to utilize the “land” and the “scan” mechanism which is typical in picornavirus IRES-mediated translation (Spriggs et al. 2009). Noticeably, eIF4E and eIF4G, which function as cap-binding protein and the scaffolding protein, are not required in some cellular IRES-mediated translation (Spriggs et al. 2009). Another translation regulation target, eIF2, also does not affect certain cellular IRES-mediated translation including c-myc, platelet-derived growth factor-2 and vascular endothelial growth factor when eIF2 α is phosphorylated (Gerlitz, Jagus, and Elroy-Stein 2002; Subkhankulova, Mitchell, and Willis 2001). Interestingly, the IRES from cationic amino acid transporter-1 is activated during eIF2 α phosphorylation (Fernandez et al. 2002). Using chemical inhibitors that target the helicase eIF4A revealed that c-myc and N-myc do require eIF4A and eIF3, which function similarly as

Class 1 or 2 IRESs (Spriggs et al. 2009). ITAFs have also been reported to modulate the efficiency of cellular IRES-mediated translation (Stoneley and Willis 2004).

After identifying potential cellular IRESs, each IRES has to be tested on a case-by-case basis. The bicistronic construct is the standard method for testing cellular mRNA IRESs (reviewed in (Jackson 2013)). DNA-based bicistronic constructs are the classical tools for investigating IRES activity, which was originally used to investigate PV IRES (Pelletier and Sonenberg 1988). In general, a typical bicistronic reporter will be a dual luciferase reporter containing Renilla luciferase (RLuc) as the upstream cistron and Firefly luciferase (FLuc) as the downstream cistron. The RNA of interest that is predicted to contain an IRES is inserted in the intergenic region between RLuc and FLuc. IRES activity is indicated by a significant increase in the FLuc/RLuc expression ratio (RLuc serves as a normalizing control) as compared to the control bicistronic construct with an empty intergenic region. The translation ability can be tested both *in vitro* by coupling with lysate extracts and *in vivo* by transfecting the reporter construct or RNA into a cell line. However, the use of this reporter construct can generate false-positive or negative results. When using a DNA-based reporter, the cryptic promoter activity or cryptic splice acceptor site within the 5' UTR mRNA (reviewed in (Kozak 2003; 2005)) will generate a shorter or separate monocistronic RNA including the second intron that is being expressed through a cap-dependent mechanism (Riley, Lindsay, and Holcik 2010). Therefore, to avoid verifying a false-positive IRES element, comprehensive analysis must be taken. For example, Northern Blotting or quantitative real-time PCR are required to confirm the integrity of the bicistronic reporter, Alternatively, siRNAs targeting the upstream reporter gene can be co-transfected with the bicistronic reporter (Jacobs and Dinman 2004). A similar reduction in both luciferase protein should be observed if the reporter

RNA remains intact. Finally, an RNA-based reporter has been recently used to avoid the artifacts made from DNA-based reporter (reviewed in (Thompson 2012))(Q. S. Wang, Au, and Jan 2013).

1.4 *Dicistroviridae* family of viruses

Dicistroviridae belongs to the order *Picornavirales*. At the time of writing, the *Dicistroviridae* contain three genera including *Cripavirus*, *Aparavirus*, and *Triatovirus* that in total include 15 viruses (<https://talk.ictvonline.org/taxonomy/>). However, this classification will likely evolve as there have been hundreds of new dicistrovirus-like genomes discovered via metagenomic studies (Shi et al. 2016; Wolfe, Dunavan, and Diamond 2007). Classification of each genera is based on the phylogenetic analysis of IGR IRES region (Warsaba, Sadasivan, and Jan 2020). All members of the dicistrovirus family infect arthropods and are distributed widely in nature. Marine environments and invertebrates have all been identified to have dicistroviruses or dicistro-like viruses (Culley, Lang, and Suttle 2003; Shi et al. 2016; Suttle 2007). CrPV and *Drosophila C virus* (DCV) are two well-studied members from this family, both of which mainly infect *Drosophila* though CrPV was initially discovered in crickets. Both viruses are studied extensively in the laboratory serving as a model of virus-host interactions (Kerr et al. 2015; Khong et al. 2016; Cherry and Perrimon 2004). Other members from *Dicistroviridae* can have devastating economic consequences as well. For example, the Taura syndrome virus infects penaeid shrimp species and has caused significant losses in Latin America and the United states economies (Bonami et al. 1997). Four dicistroviruses are found to infect honeybees, including Acute bee paralysis virus, Kashmir bee virus, Black queen cell virus, and Israeli acute paralysis virus (IAPV) (de Miranda, Cordon, and Budge 2010). Honeybees are the world's most important pollinator of food crops and their health defects are detrimental to the agriculture industry and economy. IAPV

infections correlated with honeybee Colony Collapse Disorder (Cox-Foster et al. 2007), which can lead up to 50% loss of worker bees (vanEngelsdorp et al. 2008). On the other hand, dicistroviruses can also be utilized as biopesticides. (reviewed in(Warsaba, Sadasivan, and Jan 2020)). Considering the huge impact of dicistroviruses on agriculture economy, understanding their biology and interactions with the host is the key.

1.4.1 Genome characteristics and organization

Dicistroviruses are small non-enveloped viruses that contain a monopartite positive-sense, single-stranded RNA genome ranging from 8 to 10 kb in size (reviewed in(Warsaba, Sadasivan, and Jan 2020)). The family derives its name from the two ORFs encoded in the genome (**Figure 1.4**). The first ORF encodes nonstructural proteins, including a 1A suppressor of RNA-mediated silencing, 2C RNA helicase, 3C chymotrypsin-like cysteine protease, VPg and RdRp, essential for viral replication. The second ORF encodes structural proteins, including virion proteins (VP1-4), that participate in viral assembly. The genome of dicistroviruses is covalently linked to VPg at the 5' end and includes a poly(A) tail at the 3' end. The ORFs are translated by two distinct IRESs (Joan E. Wilson, Powell, et al. 2000), and subsequently processed by the viral-encoded protease to generate mature viral proteins. ORF1 is regulated by an IRES located at the 5' UTR, and ORF2 is regulated by IGR IRES, which is classified to Class 4 IRES. This bicistronic organization enables viruses to independently control the expression of structural and non-structural proteins and shift from translation to replication (Khong et al. 2016). Dicistroviruses produce an excess of structural proteins over non-structural proteins during translation (Garrey et al. 2010; Moore, Kearns, and Pullin 1980; Joan E. Wilson, Powell, et al. 2000).

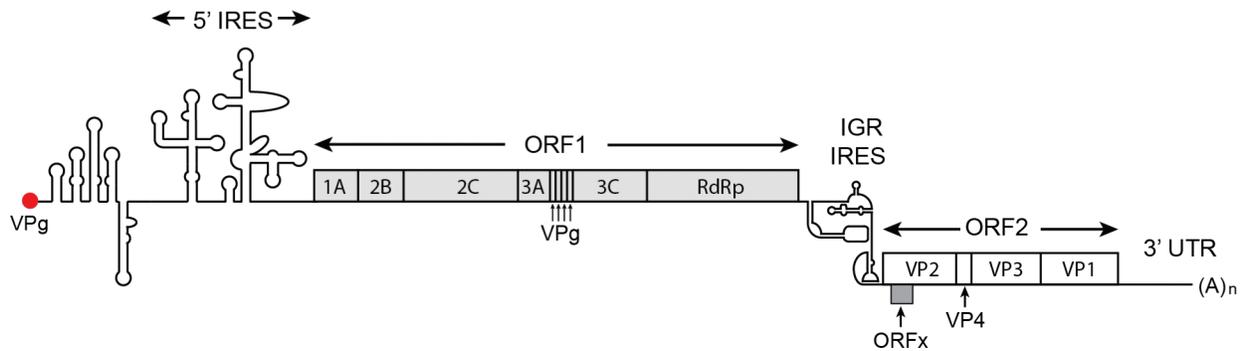


Figure 1.4. Genome organization of dicistroviruses.

Displayed is the model of CrPV genome organization. The genome contains a positive-sense single-stranded RNA which has VPg covalently linked to 5' end and poly(A) tail at 3' end. The approximately 9 kb genome contains two open reading frames (ORFs) encoding viral non-structural proteins including the suppressor of RNAi (1A), helicase (2C), VPg, protease (3C), and RNA-dependent RNA polymerase (RdRp), and structural proteins, including VP1 to VP4. Both of them are driven by distinct IRESs. Translation of the first ORF is directed by an IRES located at the 5' UTR, and the downstream ORF translation is controlled by the IGR IRES.

The dicistroviruses virion structure resembles picornaviruses (Tate et al. 1999). It consists of a 30 nm diameter icosahedron, including 60 copies of each of virion proteins (VP1, VP2, VP3). VP4 is packaged inside of the virion as a fusion protein and is attached to the inner surface, possibly contacting the viral genome. Cleavage of the precursor protein VP0 into VP3 and VP4 happens after the viral genome is encapsidated, proposed to mature the virion (Tate et al. 1999) **(Figure 1.5)**.

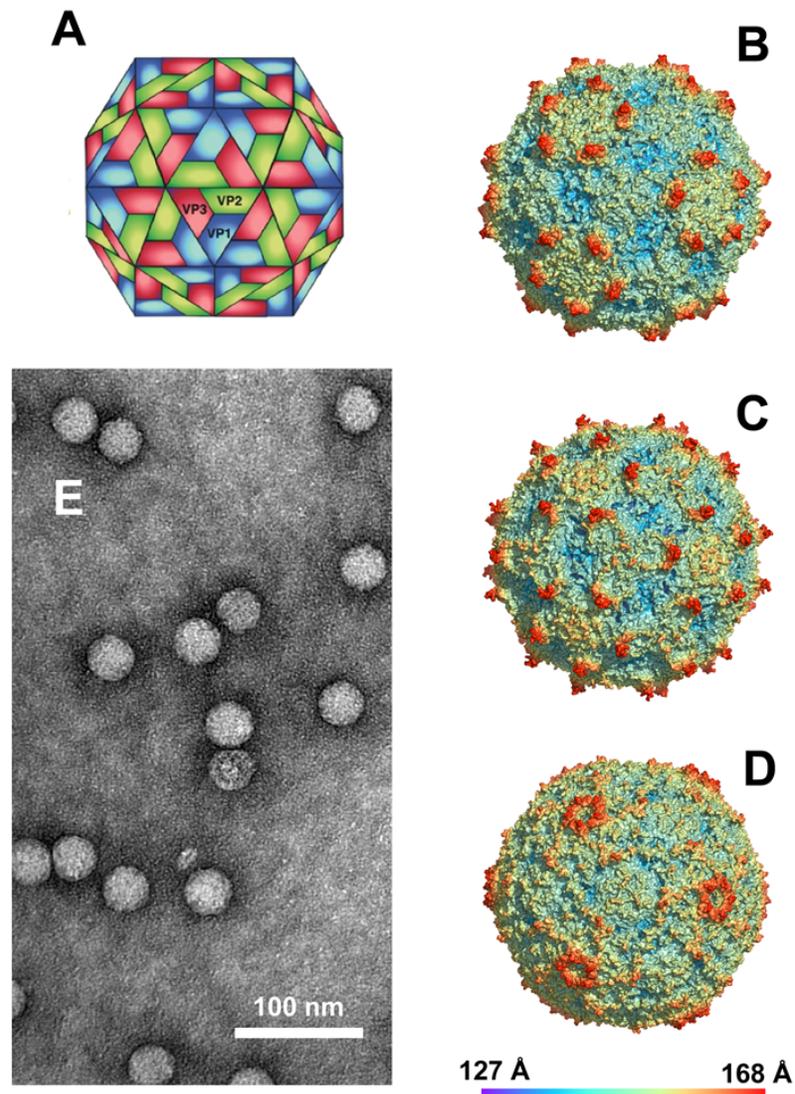


Figure 1.5. Morphology and virion structure of dicistroviruses.

(A) Diagram illustrating the packing surface of viral proteins VP1, VP2 and VP3. VP4 is located inside of the virion. VP1 subunits are shown in blue, VP2 in green and VP3 in red. Rendered X-ray crystal structures of (B) triatoma virus; (C) Israeli acute paralysis virus; (D) Cricket paralysis virus (Surface color indicates the distance from the virion center). (E) Negative-contrast electron micrograph of purified triatoma virus. Adapted with permission from (Valles et al. 2017).

1.4.2 Viral-host interaction

Dicistroviruses serve as a model to study fundamental viral-host interactions and have led to studies of conserved responses in vertebrates including translation control and innate immune pathways. Dicistroviruses infection leads to a rapid shut-off of host protein synthesis, and shifts to viral IRES-dependent translation with increasing production of viral proteins (Joan E. Wilson, Powell, et al. 2000; Garrey et al. 2010). When CrPV infects *Drosophila* S2 tissue culture cells, the host translational shutoff occurs two to three hours post-infection (Garrey et al. 2010; Khong et al. 2016). During CrPV infection, eIF2 α is phosphorylated and cap binding protein eIF4E is prohibited from binding to the scaffold protein eIF4G (Garrey et al. 2010). Although blocking eIF2 phosphorylation itself in CrPV infected cells does not inhibit host translation (Garrey et al. 2010; Khong et al. 2016), blocking the interaction between eIF4E and eIF4G is correlated to host translational shutoff (Garrey et al. 2010). This may facilitate viral protein synthesis with an increased pool of free ribosomes. Dicistroviruses can bypass the translational block through their IRES-dependent mechanism. Radioactive pulse-labelling and ribosome profiling of CrPV post-infection suggest that translation of first ORF directed by the 5' UTR IRES occurs early in infection, and that translation of the second ORF directed by the IGR IRES is delayed until three to four hours post-infection (Khong et al. 2016; Garrey et al. 2010). The temporal regulation of 5'UTR and IGR IRES-dependent translation may serve as an optimized mechanism to firstly express optimal levels of non-structural proteins for viral translation and replication, and then shift to synthesis of structural proteins for viral packing and assembly (Khong et al. 2016).

1.4.3 Intergenic region internal ribosome entry site

The IGR IRES utilizes the simplest mechanism to initiate translation. IGR IRESs have been identified exclusively in dicistroviruses and are about 150-200 nucleotides in length. Its compact

structure allows it to bind directly to ribosomes without any eIFs. Based on biochemical and structural probing, it has been shown that IGR IRES is composed of three pseudoknots (PKI, PKII, PKIII) in a two-domain architecture (Pfungsten, Costantino, and Kieft 2006; Schüler et al. 2006; Fernández et al. 2014) (**Figure 1.6A**). PKII and PKIII form a core domain that is responsible for recruiting 40S and 60S ribosomal subunits (**Figure 1.6B**), while PKI contains a tRNA-like anticodon:codon structure that occupies the ribosomal A-site (**Figure 1.6B, C**) (Mailliot and Martin 2018; Kerr et al. 2016; Pisareva, Pisarev, and Fernández 2018). Structural and biochemical analyses have revealed key contacts between the IGR IRES and the ribosome that drive factorless translation. In the initial assembly of ribosomes on the IGR IRES, stem-loops SLIV and SLV interact with uS7 and eS25 of the 40S subunit, and loop L1.1 interacts with the L1 stalk of the 60S subunit (Kerr et al. 2016). Mutations of these terminal loop regions disrupt 40S and/or 60S recruitment. Because PKI controls the position of ribosomes, another remarkable aspect of the IGR IRES is that it can initiate at non-AUG codon site (Pestova and Hellen 2003; Joan E. Wilson, Pestova, et al. 2000; Jan and Sarnow 2002). Based on the structures identified, IGR IRESs are classified into two subgroups, Type I and Type II. (**Figure 1.6A**). CrPV and IAPV are representatives of Type I and Type II respectively. Type II differs from Type I by an extra stem-loop (SLIII) in PKI domain, a larger L1.1 loop and a lack of a “shoulder” element next to SLIV. Although the sequences of IGR IRESs are variable, their overall two-domain conformations are conserved. There are also conserved sequences at certain loop regions which make key contacts with the ribosome (Kerr et al. 2016). It has been shown that although Type I and Type II have subtle difference in the tRNA-like domain, their PKI regions can be interchangeable to generate functional chimeric IGR IRESs that can still lead translation (Hertz and Thompson 2011; Au and Jan 2012; C. J. Jang and Jan 2010).

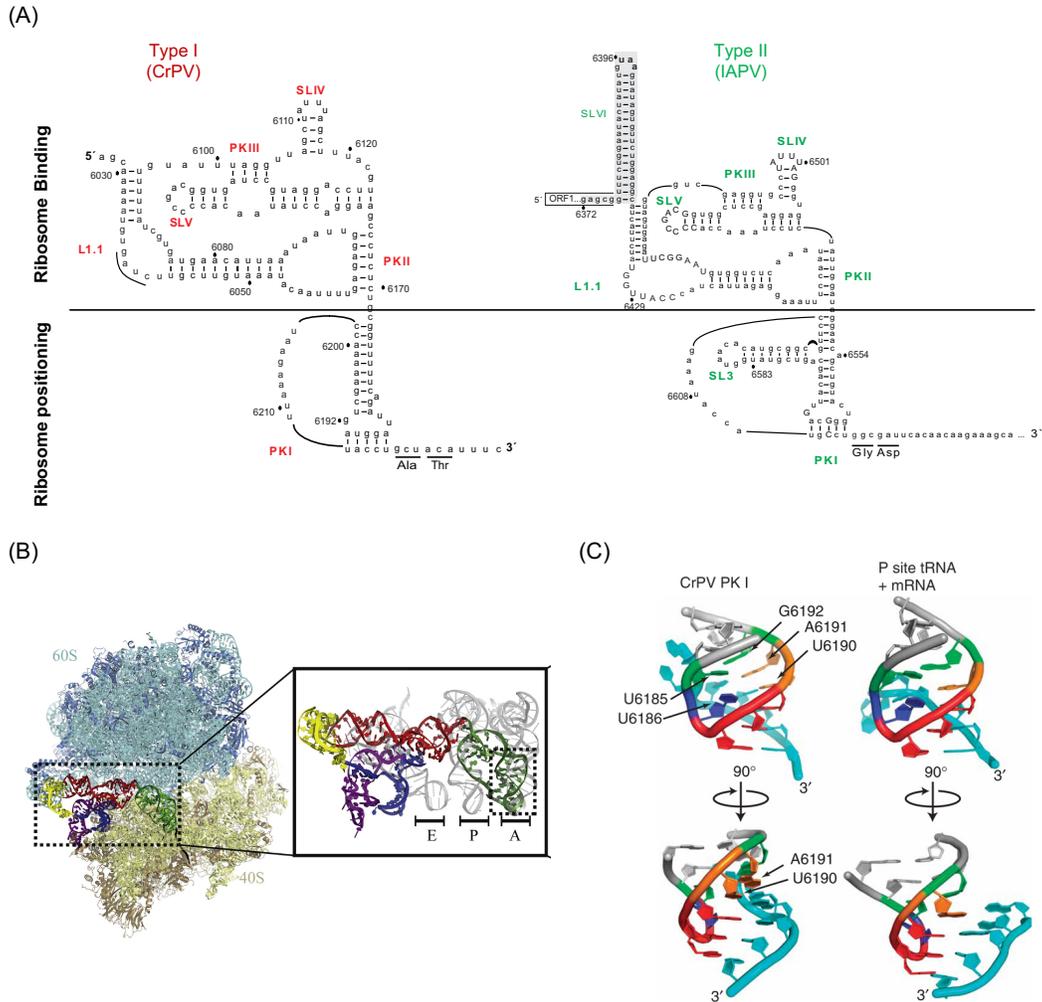


Figure 1.6. Dicistroviruses intergenic region internal ribosome entry site.

(A) Secondary structure of CrPV (left) and IAPV (right) IGR IRES as representatives of Type I and Type II IGR IRESs. SLIV and V interacts with 40S subunit and L1.1 interacts with 60S subunit forming ribosome binding domain. The PKI domain mimics a tRNA-like anticodon:codon structure to position the IRES at the A site. (B) Cryo-EM structure of the CrPV IGR IRES bound to the yeast *Kluyveromyces lactis* 80S ribosome solved at 3.7 Å. Inset: Superposition of the CrPV IGR IRES with A, P, and E site of tRNAs, while PKI (green) is in the decoding region where an A site tRNA normally occupies. Reproduced with permission from (Fernández et al. 2014). (C) Comparison of the CrPV PKI anticodon:codon interaction and a P-site tRNA-mRNA anticodon:codon interaction. Analogous bases are highlighted with the same color. Reproduced with permission from (Costantino et al. 2008).

1.4.3.1 Mechanism of IGR IRES-mediated translation initiation

The ribosome binding domain consisting of PKII and PKIII interacts with the 40S and 60S ribosome subunits respectively. SLIV and SLV establish critical contacts with ribosomal protein (rp) S7 and S25 of 40S subunit (Nishiyama et al. 2007; Pfingsten, Costantino, and Kieft 2006; Schüler et al. 2006) (**Figure 1.6B**). It was shown that ribosomes deficient in rpS25 exhibited significant reduced binding of CrPV and *Plautia stali* intestine virus (PSIV) IRESs (Landry, Hertz, and Thompson 2009; Muhs et al. 2011). Mutational analysis revealed that the sequences identity at apical loops of SLIV and SLV are important when interacting with 40S subunit (Kerr et al. 2016). L1.1 region is also conserved in each subgroup of IGR IRESs, and mainly interacts with the L1 stalk of the 60S subunit to facilitate 80S complex formation (Pfingsten, Costantino, and Kieft 2006; Spahn et al. 2004) (**Figure 1.6B**). The tRNA-like domain consisting of PKI establishes the translation reading frame by positioning at A-site of ribosomes through its anticodon-codon basepairing derived from canonical tRNA:mRNA anticodon:codon basepairing (**Figure 1.6C**). Mutating PKI does not affect IGR IRES recruiting ribosomes, but ribosomes cannot position properly and therefore block the initiation (Jan and Sarnow 2002; C. J. Jang, Lo, and Jan 2008). The flexibility of variable loop (VRL) is also critical for PKI positioning in the ribosome (Ruehle et al. 2015; Au and Jan 2012).

Structures at low-resolution by cryo-electron microscopy firstly showed the 40S-IRES complex where IGR IRES adopts an elongated shape with PKI extruding from the E to the P site (Spahn et al. 2004; Schüler et al. 2006). When the IRES is in the complex with the 80S ribosome, the IRES binds in the intersubunit space between the 40S and 60S subunits (**Figure 1.7**). The head of the 40S subunit adopts two states with canonical and rotated states, which facilitate the opening of the mRNA channel and the lock respectively. Later, a high-resolution structural models were

acquired that brought indepth mechanistic insight of IGR IRES-ribosome interaction (Fernández et al. 2014; Muhs et al. 2011; Koh et al. 2014). The structural model showed that the PKI domain is positioned in the A site of the ribosome (Fernández et al. 2014; Koh et al. 2014), unlike the canonical initiator tRNA which is delivered to the P site of the ribosome. Upon positioning, the IRES undergoes pseudotranslocation by moving the PKI domain from A site to P site allowing the delivery of the first aminoacyl-tRNA to the A site and starting peptidyl synthesis (**Figure 1.7**). This step differs from the canonical translocation step by not having a peptide bond formation. It has also been proposed that without the delivery of an aminoacyl-tRNA or in the absence of eEF1A, IRES can undergo back-translocation from the P site to the A site (Kerr and Jan 2016). By utilizing this predominant streamlined mechanism, the IGR IRESs from dicistroviruses can bypass the translational regulation and usurp host ribosomes for protein synthesis. The IGR IRES is functional in other systems including mammal, yeast and also bacteria, suggesting a universal mechanism across kingdoms (Colussi et al. 2015; Thompson, Gulyas, and Sarnow 2001; J E Wilson, Powell, et al. 2000).

The dynamics of the 40S head is important for Met-tRNA_i positioning in the P-site during canonical eukaryotic translation initiation (Llácer et al. 2015). In IGR IRES-mediated translation, the flexible 40S head is restricted by protruding SLIV and SLV into the cleft formed between the head and the body of the 40S, which facilitate PKI binding to the decoding region (Murray et al. 2016). The PKI domain shows marked similarity with the canonical tRNA decoding.

Once the 60S subunit is recruited, the binary IRES:80S complex oscillates between rotated state and canonical state which are similar to ribosome complexes with tRNA prior to translocation (Muhs et al. 2011; Koh et al. 2014; Murray et al. 2016). In the rotated state, the 40S subunit rotates counterclockwise relative to the 60S together with 40S head swiveling. The L1 stalk from 60S

subunit is dynamic in the absence of interaction. Once the IRES is bound, its orientation is fixed to an outward conformation by the interaction with the L1.1 loop region of IRES (Murray et al. 2016). Because the IRES is in contact with these dynamic ribosomal subunits, it also adopts conformational changes based on the rotational state of the ribosome to maintain the interaction (Murray et al. 2016; Fernández et al. 2014). When the 40S subunit is in the rotated state, it is induced a further ~ 3 -degree rotation by the delivery of eEF2-GTP (Fernández et al. 2014; Muhs et al. 2011; Koh et al. 2014) (**Figure 1.7**). eEF2 stabilizes PKI in an intermediate A/P site hybrid state by its domain IV interaction (Murray et al. 2016; Abeyrathne et al. 2016). Once the eEF2 bound GTP hydrolyzes, eEF2 undergoes conformational change followed by translocation of PKI to P site. Meanwhile the original contact by SLIV and SLV with the 40S subunit is broken but the L1 stalk interaction with L1.1 loop region of IRES is maintained (Abeyrathne et al. 2016). This conformational change cycle continues until the PKI completely translocates to E site and aminoacyl-tRNA is placed in the P site, which triggers the disassembly of anticodon stem loop (ASL) from PKI and mimicking the acceptor stem of a canonical E site tRNA (Pisareva, Pisarev, and Fernández 2018).

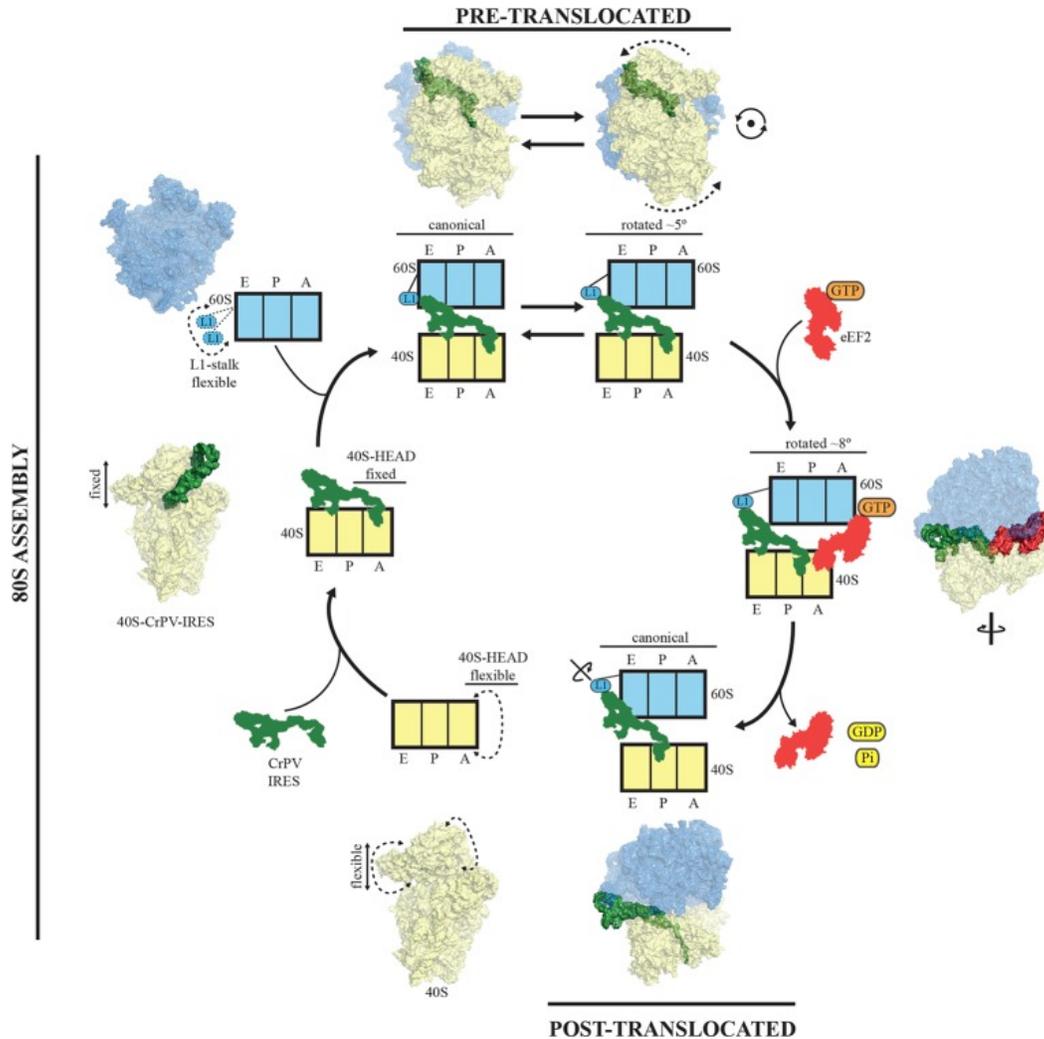


Figure 1.7. A pathway model of IGR IRES-mediated translation initiation.

Through the interaction between SLIV and V with the 40S subunit, the 40S head flexibility is restricted, allowing 60S subunit to join. Upon 60S subunit recruited, the binary IRES/80S complex fluctuates between canonical and rotated states with PKI positioning at A site ('pre-translocation states'). After eEF2-GTP binds to the rotated state of 40S subunit, it induces an extra ~ 3 degree of rotation of the 40S. The contact between Domain IV of eEF2 and PKI stabilizes IRES in an intermediate ap/P-like state. During translocation, eEF2 bound GTP hydrolyzes and contacts of SL IV and V with the 40S are disrupted, which promote the translocation of PKI to the P site ('post-translocation state'). Reproduced with permission from (Murray et al. 2016).

1.4.3.2 Significance of IGR IRES

The IGR IRES serves as a powerful model to understand the mechanism of IRES-dependent translation initiation. With the most simplified mechanism, the information obtained has been applied to other more complex IRESs such as the HCV IRES (Landry, Hertz, and Thompson 2009). Indeed similarities were found between IGR IRES and HCV IRES, for example, which both require Rps25 for translation activity (Landry, Hertz, and Thompson 2009). Both also utilize a similar mechanism to facilitate the delivery of the first aminoacyl-tRNA to the P site of the ribosome by repositioning PKI in IGR IRES or domain II in HCV IRES to resemble endogenous tRNA states (Pisareva, Pisarev, and Fernández 2018; Yamamoto et al. 2015). IGR IRES can also serve as an important tool to reconstitute a minimal translation system thus allowing studies on translation elongation and termination by bypassing the canonical translation initiation requirement (Pestova and Hellen 2005; Jan, Kinzy, and Sarnow 2003). IGR IRES can determine whether a cellular process is initiation factor dependent. By utilizing IGR IRES-containing reporter mRNAs, microRNA (miRNA)-mediated regulation has been determined to occur post initiation, based on the observation that miRNA translation inhibition is independent of cap-dependent mechanism (Petersen et al. 2006). IGR IRES can also be employed to determine the eIFs dependency of nonsense mediated decay (NMD). NMD degrades mRNAs containing premature stop codons. However, introducing a premature stop codon into a reporter containing IGR IRES does not induce NMD and suggests that NMD requires at least some eIFs (Isken et al. 2008). eIF3 has been suggested to be required for NMD recognition (Chiu et al. 2004). Indeed, a reporter including EMCV IRES which requires eIF3 for initiation is recognized and degraded by NMD (Isken et al. 2008).

The IGR IRES can be used as a translation promoter in order to express proteins of interest especially under the stressed conditions. Under different stress conditions, translation regulation is active to limit the amounts of proteins synthesized. Because the IGR IRES can bypass the requirement of eIFs, even when ternary complex availability is limited, the IGR IRES can still function. Therefore, by having an IGR IRES in expression vectors, the protein of interest can be efficiently translated and studied even under the stress conditions. Taken together, IGR IRES can both improve the understanding of IRES-mediated translation initiation by serving as the fundamental model and serve as a tool in different biological investigation.

1.5 Thesis investigation

The IGR IRESs from dicistroviruses utilize an unprecedented streamlined mechanism to recruit ribosomes independently of initiation factors and start translation. Extensive analysis including biomolecular and structural studies revealed how the IGR IRES functions and recruits ribosomes explicitly. However, the origin of this IGR IRES and how it evolved are still unclear. Here, I hypothesize that IGR IRES mediated mechanism is essential for viral translation in a historical framework and the structure is required through virus-host co-evolution. In my proposal, I will be focusing on two different objectives to address this question. One objective is to address the viral evolution by investigating the IGR IRES property of an ancient dicistrovirus recovered from 700-year-old caribou feces. Another objective is to address the virus-host co-evolution by searching for structured RNA elements in the host genome that can bind directly to ribosomes. It is apparent that viruses and hosts have co-evolved leading to exchange of genetic material. For example, Hepatitis D virus (HDV) ribozymes, self-cleaving catalytic RNAs, have a conserved sequence in the human *CPEB3* gene (Author et al. 2006). In *Drosophilla*, fragments of diverse

non-retroviral RNA are embedded in the *Drosophila* genome, making viral infection persistent (Goic et al. 2013). In Chapter 2, I characterized an IGR IRES from a divergent dicistrovirus RNA genome extracted from 700-year-old caribou feces trapped in a subarctic ice patch. Using several biomolecular assays, including the ribosome binding assay, the ribosome competition assay, the primer extension assay, the *in vitro* translation assay, I showed that this “ancient” IGR IRES functions as a modern IGR IRES. By cloning it into the CrPV infectious clone, the new chimeric virus clone was still infectious and could produce viruses. In Chapter 3, I validated and analyzed candidates generated from the screen which tried to identify IGR IRES-like elements from *Drosophila*. Their ability to bind to ribosomes independently of initiation factors and modulation effect on translation were tested. By doing ribosome binding assays, I showed that three candidates can bind to ribosomes tightly. Among these three candidates, one of them can also compete with CrPV IGR IRES for ribosome binding. However, from the *in vitro* translation assay, it did not affect translation level.

Chapter 2: Resurrection of an IGR IRES from ancient Northwest

Territories cripavirus

2.1 Introduction

The origin of the IGR IRES mechanism is not known. Currently, the IGR IRESs studied to date can be divided into two main types; Type I and II IGR IRESs are exemplified by the CrPV and IAPV IGR IRES, respectively, with the main differences in the L1.1 nucleotides, and that the Type II IGR IRESs have an extra stem-loop within the PKI domain (Nakashima and Uchiumi 2009). Previous studies have shown that the PKI domains of Type I and II can be functionally interchanged and that the IRESs are modular, thus suggesting that the IRESs may have evolved through recombination of modular domains (C. J. Jang and Jan 2010; Hertz and Thompson 2011). Recent metagenomic approaches have identified an increasing diversity of dicistro-like viral genomes and IGR IRESs that may provide hints into the origins of the IRES (Shi et al. 2016; Wolfe, Dunavan, and Diamond 2007). Alternatively, identifying ancient RNA viral genomes may provide historical context for the evolution of viral strategies; however, this is challenging given the relatively labile nature of RNA. However, some preserved RNA viral genomes have been discovered. The pioneering benchmark in identifying ancient RNA viruses is the recovery of the influenza virus genome from 1918–1919 (Taubenberger et al. 1997; 2005). Furthermore, a complete genome of ancient Barley Stripe Mosaic Virus was identified from barley grain dated (~700 years) and a 1000 year old RNA virus related to plant chryso-viruses was isolated from old maize samples with a nearly complete genome (Smith et al. 2014; Peyambari et al. 2019). Recently, Ng et al. recovered two novel viruses from 700-year-old caribou feces trapped in a subarctic ice

patch (Ng et al. 2014), one of which is a fragment of a divergent dicistrovirus RNA genome, named ancient Northwest territories criparvirus (aNCV). The partial aNCV sequence that was recovered included the IGR domain, thus providing an opportunity to characterize and compare the aNCV IGR to contemporary dicistrovirus IGR IRESs. In this study, we examine the molecular and biochemical properties of the aNCV IGR and demonstrate that the aNCV IGR directly assembles ribosomes and can direct internal ribosome entry both in vitro and in vivo. We show that the intact aNCV IGR including the IRES region and 105 nucleotides upstream of the IRES can support viral translation and infection in a heterologous dicistrovirus clone, thus highlighting the significance of this translational mechanism in an ancient virus.

2.2 Materials and methods

2.2.1 Cell culture

Drosophila Schneider line 2 (S2) cells were maintained and passaged at 25 °C in Shields and Sang M3 insect medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum.

2.2.2 Virus infection

S2 cells were infected at the desired multiplicity of infection in minimal phosphate-buffered saline (PBS) at 25 °C. After 30 min absorption, complete medium was added and harvested at the desired time point. Virus titres were monitored as described (Au, Elspass, and Jan 2017) using immunofluorescence (anti-VP2).

2.2.3 DNA constructs

The aNCV IGR sequence (Accession KJ938718.1) was synthesized (Twist Biosciences) and cloned into pEJ4 containing EcoRI and NcoI sites upstream of the firefly luciferase (FLuc) open reading frame. For the bicistronic reporter construct, PCR-amplified full-length aNCV IGR

or truncated aNCV IGR (Δ 1-99) was ligated into the standard bicistronic reporter construct (pEJ253) using EcoRI and NdeI sites.

2.2.4 CrPV/aNCV chimeric infectious clone constructs

The chimeric CrPV-aNCV infectious clone was derived from the full-length CrPV infectious clone (pCrPV-3; Accession KP974707) (Kerr et al. 2015) using Gibson assembly (New England BioLabs), per the manufacturer's instructions. All constructs were verified by sequencing.

2.2.5 In vitro transcription and translation

Mono-cistronic and bicistronic reporters were linearized with NcoI and BamHI, respectively. pCrPV-3 and chimeric CrPV-aNCV clones were linearized with Eco53kI. RNAs were in vitro transcribed using a bacteriophage T7 RNA polymerase reaction and RNA was purified using a RNeasy Kit (Qiagen). Radiolabeled RNAs were bulk labelled by incorporating α -[32 P] UTP (3000 Ci/mmol). The integrity and purity of RNAs were confirmed by agarose gel analysis. Uncapped bicistronic RNAs were first pre-folded by heating at 65 °C for 3 min, followed by the addition of 1xbuffer E (final concentration: 20 mM Tris pH 7.5, 100 mM KCl, 2.5 mM MgOAc, 0.25 mM Spermidine, and 2 mM DTT) and slowly cooled at room temperature for 10 min. The pre-folded RNAs (20–40 ng/ μ L) were incubated in RRL containing 8 U Ribolock inhibitor (Thermo Fisher Scientific), 20 μ M amino acid mix minus methionine, 0.3 μ L [35 S]-methionine/cysteine (PerkinElmer, >1000 Ci/mmol), and 75 mM KOAc pH 7.5 at 30 °C for 1 hr. For the infectious clones, 2 μ g RNA was incubated at 30 °C for 2 hr in *Spodoptera frugiperda* (Sf21) extract (Promega) in the presence of [35 S]-methionine/cysteine (PerkinElmer) and an additional 40 mM KOAc and 0.5 mM MgCl₂. The translated proteins were resolved using SDS-PAGE and analyzed by phosphor-imager analysis (Typhoon, GE life sciences).

2.2.6 Purification of the 40S and 60S subunits

Ribosomal subunits were purified from HeLa cell pellets (Cell Culture Company) as described [31]. In brief, HeLa cells were lysed in a lysis buffer (15 mM Tris-HCl (pH 7.5), 300 mM NaCl, 6 mM MgCl₂, 1% (v/v) Triton X-100, 1 mg/mL heparin). Debris was removed by centrifuging at 23,000 × g and the supernatant was layered on a 30% (w/w) cushion of sucrose in 0.5 M KCl and centrifuged at 100,000 × g to pellet crude ribosomes. Ribosomes were gently resuspended in buffer B (20 mM Tris-HCl (pH 7.5), 6 mM magnesium acetate, 150 mM KCl, 6.8% (w/v) sucrose, 1 mM DTT) at 4 °C, treated with puromycin (final 2.3 mM) to release ribosomes from mRNA, and KCl (final 500 mM) was added to wash and separate 80S ribosomes into 40S and 60S. The dissociated ribosomes were then separated on a 10%–30% (w/w) sucrose gradient. The 40S and 60S peaks were detected by measuring the absorbance at 260 nm. Corresponding fractions were pooled and concentrated using Amicon Ultra spin concentrators (Millipore Sigma) in buffer C (20 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, 10 mM KCl, 1 mM MgCl₂, 6.8% sucrose). The concentration of 40S and 60S subunits was determined by spectrophotometry, using the conversions 1 A_{260 nm} = 50 nM for 40S subunits, and 1 A_{260 nm} = 25 nM for 60S subunits.

2.2.7 Filter-binding assays

RNAs (final 0.5 nM) were preheated at 65 °C for 3 min, followed by the addition of 1xbuffer E and slowed cooling in a water bath preheated to 60 °C for 20 min. The pre-folded RNAs, with 50 ng/μL of non-competitor RNA, were incubated with an increasing amount of 40S subunits from 0.1 nM to 100 nM, and a 1.5-fold excess of 60S subunits for 20 min at room temperature. Non-competitor RNAs were in vitro transcribed from the pcDNA3 vector 880–948 nucleotides. Reactions were then loaded onto a Bio-Dot filtration apparatus (Bio-Rad) including a double membrane of nitrocellulose and nylon pre-washed with buffer E. Membranes were then washed

three times with buffer E, dried, and the radioactivity was imaged and quantified by phosphor-imager analysis. The dissociation constant is determined by the formula,

$$\frac{[AB]}{[A]_{total}} = f_{max} \left(\frac{[B]}{[B] + K_D} \right) \quad (1)$$

where [A] is the concentration of RNAs, [B] is the concentration of ribosomes, [AB] is the concentration of RNAs bound to the ribosomes, f_{max} is the saturation point, and K_D is the dissociation constant.

Bulk-labeled RNAs (final 0.5 nM), IRES competitors, and non-competitor RNAs (50 ng/ μ L) were pre-folded in buffer E. Unlabeled IRES competitors were added in increasing concentrations from 2 nM to 250 nM. RNAs were then incubated with 6 nM 40S and 9 nM 60S subunits at room temperature for 20 min. Reactions were then loaded onto the Bio-Dot filtration apparatus, and data were fitted to the Linn-Riggs equation that describes competitive ligand binding to the target:

$$\theta = \frac{[S](1 - \theta)}{K_D \{ (1 + [C])K_C \} + [R](1 - \theta)} \quad (1)$$

where [S] is 80S ribosome concentration, [R] is radiolabelled RNA concentration, θ is fraction of radiolabelled RNAs bound to ribosomes, [C] is competitor RNA concentration, K_D and K_C are dissociation constants of labelled RNAs and competitor RNAs, respectively.

2.2.8 Ribosome protection assay

[³²P]-labeled RNAs (final 0.1 μ M) were pre-folded in buffer E and incubated with 0.6 μ M 40S and 0.9 μ M 60S at room temperature for 20 min. 1 μ L of 1 U/ μ L of RNase I (Thermo Fisher Scientific) was added to the mixture and incubated at 20 °C for 1 hr. RNAs without RNase I

treatment were incubated at 20 °C with the same time length. RNAs from mixtures with or without RNase I treatment were TRIZOL-extracted and loaded onto 6% (w/v) polyacrylamide/8M urea gels to separate them. RNA Ladders (Thermo Fisher Scientific) were synthesized per manufacturer's instruction. Gels were dried and imaged by phosphor-imager analysis.

2.2.9 Toeprinting/Primer extension analysis

Toeprinting analysis of ribosomal complexes in RRL was performed as previously described (J E Wilson, Pestova, et al. 2000). 0.4 μ g of bicistronic WT or mutant CrPV IGR IRES RNAs and WT or mutant aNCV IGR IRES RNAs were annealed to primer 5'-GTAAAAGCAATTGTTCCAGGAACCAG-3' and primer 5'-GTTAGCAGACTTCCTCTGCCCTCTC-3', respectively in 40 mM Tris (pH 7.5) and 0.2 mM EDTA by slow cooling from 65 °C to 30 °C. Annealed RNAs were added to RRL pre-incubated with 0.68 mg/mL cycloheximide and containing 20 μ M amino acid mix, 8 units of Ribolock (Thermo Fisher Scientific) and 154 nM final concentration of potassium acetate (pH 7.5). The reaction was incubated at 30 °C for 20 min. Toeprinting analysis using purified ribosomes was performed as follows: 75 ng of RNAs were annealed to primers in 40 mM Tris·Cl, pH 7.5, and 0.2 mM EDTA by slow cooling from 65 °C to 35 °C. Annealed RNAs were incubated in buffer E (containing 100 mM KCl) containing 40S (final concentration 100 nM); 60S (final concentration 150 nM); ribo-lock (0.02 U/ μ L) at 30 °C for 20 min. Following incubation, ribosome positioning was determined by primer extension/reverse transcription using AMV reverse transcriptase (1 U/ μ L) (Promega) in the presence of 125 μ M of each of dTTP, dGTP, dCTP, 25 μ M dATP, 0.5 μ L of α - [³²P] dATP (3.33 μ M, 3000 Ci/mmol), 8 mM MgOAc, in the final reaction volume. The reverse transcription reaction was incubated at 30 °C for 1 h, after which it was quenched by the

addition of STOP solution (0.45 M NH₄OAc, 0.1% SDS, 1 mM EDTA). Following the reverse transcription reaction, the samples were extracted by phenol/chloroform (twice), chloroform alone (once), and ethanol precipitated. The cDNA was analyzed under denaturing conditions on 6% (w/v) polyacrylamide/8M urea gels, which were dried and subjected to phosphor-imager analysis.

2.2.10 RNA transfection

3 µg of in vitro-transcribed RNA derived from the pCrPV-3 or chimeric clones was transfected into 3 x 10⁶ S2 cells using lipofectamine 2000 reagent (Thermo Fisher Scientific) per the manufacturer's instructions.

2.2.11 RT-PCR and sequence confirmation

Viruses were passaged three times after transfection of viral RNAs in S2 cells. Total RNA was isolated from S2 cells using TRIzol reagent. RT was performed using 1 µg of RNA using LunaScript™ RT SuperMix Kit (New England BioLabs) per manufacturer's instructions. For reverse transcription of the negative-sense CrPV viral RNA to detect replication, tagged primer (5'-CTATGGATCCATGGGAGAAGATCAGCAAAT-3'; tag is underlined) was used. Primer (5'-CTATGGATCCATGGGAGAAG-3') and primer (5'-GTGGCTGAAATACTATCTCTGG-3') were used for PCR amplification of the negative-sense strand of the CrPV genome. Rps6 was amplified using primers (5'-CGATATCCTCGGTGACGAGT-3') and (5'-CCCTTCTTCAAGACGACCAG-3').

2.2.12 Western blotting

Cells were washed once using 1 × PBS and harvested in lysis buffer (20 mM HEPES, 150 mM sodium chloride, 1% Triton X-100, 10% glycerol, 1 mM EDTA, 10 mM tetra-pyrophosphate, 100 mM sodium fluoride, 17.5 mM β-glycerophosphate, protease inhibitor cocktail (Sigma-

Aldrich). Equal amounts of protein lysate were resolved by SDS-PAGE and subsequently transferred onto polyvinylidene difluoride Immobilon-FL membrane (Millipore Sigma). Following transfer, the membrane was blocked for 30 min at room temperature in 5% skim milk in Tris-buffered saline containing 0.1% Tween 20 (TBST) and incubated for 1 h with rabbit polyclonal antibody raised against CrPV 1A (1:1000; Genscript) or CrPV VP2 (1:5000; Genscript). Membranes were washed three times with TBST and subsequently incubated with IRDye 800CW goat anti-rabbit IgG (1:20,000; Li-Cor Biosciences) for 1 h at room temperature.

2.3 Results

2.3.1 aNCV IGR IRES adopts a triple pseudoknot structure

Using computational and compensatory base-pairing analysis, a secondary structure model of the aNCV IGR was predicted (**Figure 2.1A**). The aNCV IGR is predicted to adopt an overall similar structure to modern dicistrovirus IGR IRESs, all possessing the main PKI, PKII and PKIII structures and is classified as a Type I IGR IRES. However, there were notable differences. An alignment analysis of the aNCV IGR with classic Type I and II IGR IRESs showed that the aNCV IGR structure contains a chimera mix of Type I and II features, especially at key domains that direct distinct steps of IGR IRES translation (**Figure 2.1B**). Specifically, the aNCV IGR shares sequence similarities to Type I IGR IRESs within Loop 1.1B, Loop 3, and to Type II at Loop 1.1A. Furthermore, whereas all Type I and II IGR IRESs contain an invariant AUUU within the loop of SLIV, the aNCV contains an AUUA loop sequence. Strikingly, the aNCV IGR IRES also includes an extra 105 nucleotides between the stop codon of ORF1 and the start of the predicted IGR IRES structure. Several smaller stem-loops (SLVI, SLVII and SLVIII) are predicted within this 105-nucleotide region. In summary, the aNCV IGR adopts an overall secondary structure that is similar

to other known dicistrovirus IGR IRESs but is a chimera of Type I and II IGR IRESs at key loop domains and has an extended upstream region.

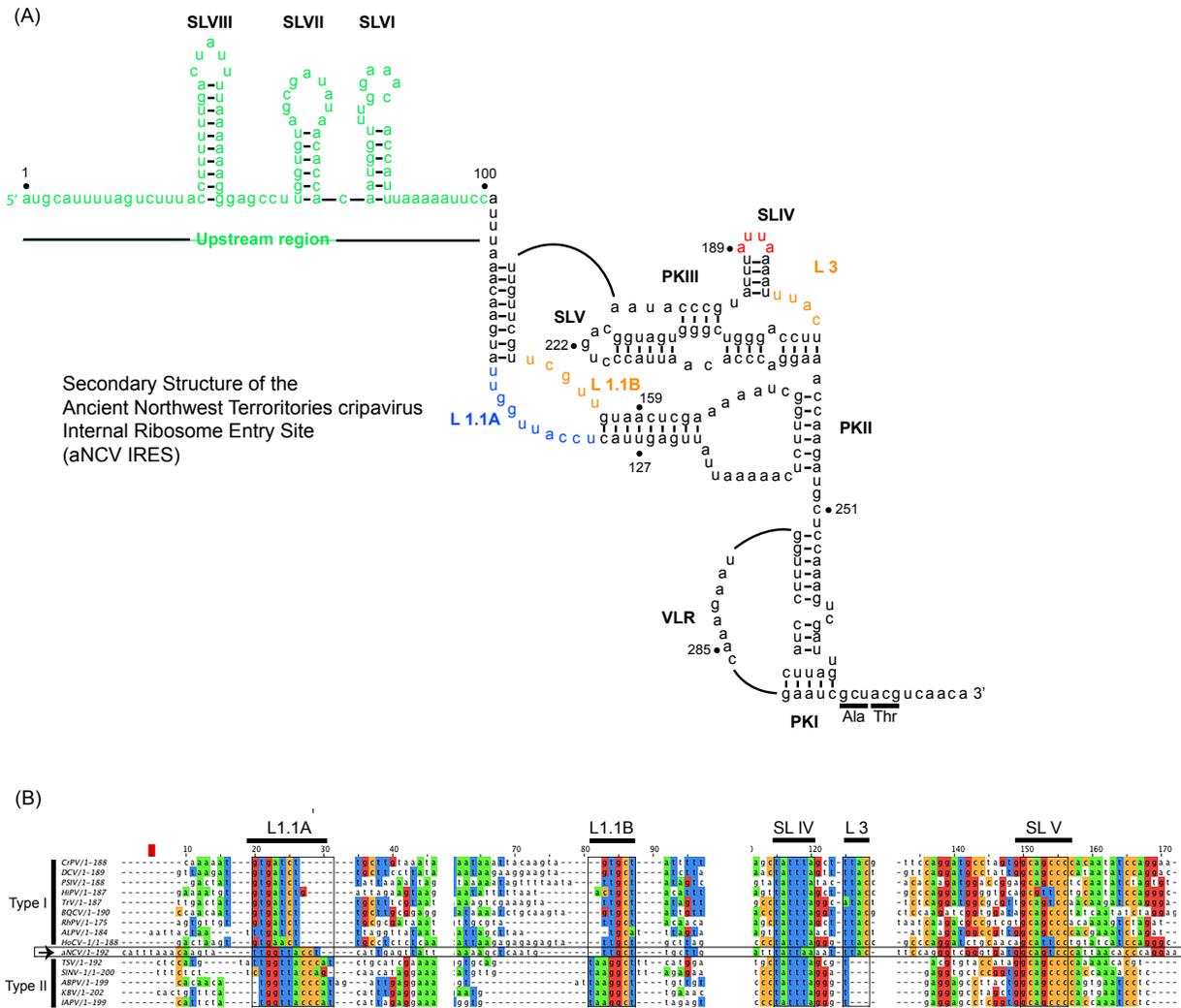


Figure 2.1. Structure information of aNCV IGR IRES.

(A) Secondary structure model of the ancient Northwest Territories Cripavirus (aNCV) intergenic region (IGR) (accession KJ938718.1) Pseudoknots (PK) I, II, III are indicated. Orange-colored nucleotides at Loop (L)1.1B, L3 denote sequences conserved in Type I IGR internal ribosome entry site (IRES), and blue-colored nucleotides at L1.1A indicate sequences conserved in Type II IGR IRES. Red-colored nucleotides at stem loop (SL) IV are unique in sequence in aNCV IGR IRES. Green-colored nucleotides represent IGR sequences outside the core IRES. The predicted start codon is GCU adjacent to the PKI domain. Numbering refers to the nucleotide position in the IGR as the genome of aNCV is not complete. (B) Alignment of Type I and II IGR IRESs with aNCV IGR.

2.3.2 aNCV IGR IRES binds tightly to human 80S ribosomes

Given the chimeric makeup of the aNCV IGR, we next examined whether the aNCV IGR possesses properties similar to dicistrovirus IGR IRESs. The first property tested was its ability to bind to purified ribosomes *in vitro*. The IGR IRESs directly bind to purified ribosomes with high affinity (Jan and Sarnow 2002; C. J. Jang, Lo, and Jan* 2008). Using an established filter-binding assay, we measured 80S binding to the aNCV IGR and the CrPV IGR IRES by incubating radiolabeled IGR with increasing amounts of purified salt-washed human 40S and 60S. The fraction of ribosome: IGR complexes was then resolved by monitoring the amount of radioactivity on the nitro-cellular and nylon membrane. As shown previously, the wild type but not the mutant (Δ PKI/II/III) CrPV IGR IRES bound to 80S ribosomes with high affinity (apparent K_D 0.4 nM) (Jan and Sarnow 2002; C. J. Jang, Lo, and Jan* 2008). The mutant CrPV Δ PKI/II/III IRES contains mutations that disrupts all three PK base-pairings (**Figure 2.2A**). Similarly, the wild type aNCV IGR bound to purified ribosomes with a similar affinity as the CrPV IGR IRES (K_D 0.7 nM). To confirm the binding specificity of ribosome:IGR complex interactions, we performed competition assays by addition of excess unlabeled IGR RNAs to the reaction prior to incubating with ribosomes. As expected, adding increasing amounts of unlabeled wild type but not mutant CrPV IGR decreased the levels of radiolabeled CrPV IGR bound to 80S ribosomes (**Figure 2.2B**, left). Similarly, adding increasing amounts of unlabeled aNCV IGR also decreased the fraction of CrPV IGR IRES-ribosome complex formation. In the reverse experiment, excess of unlabeled wild-type CrPV and aNCV IGRs but not mutant CrPV IGR also competed for ribosomes from radiolabeled aNCV (**Figure 2.2B**, right). The apparent K_D measurements of the aNCV-ribosome complex binding in the competition assay were similar to the direct filter binding assay (K_D 0.8 nM). Taken

together, aNCV IGR RNA binds to purified ribosomes with high affinity and is likely to occupy the same sites on the ribosome as the CrPV IGR IRES.

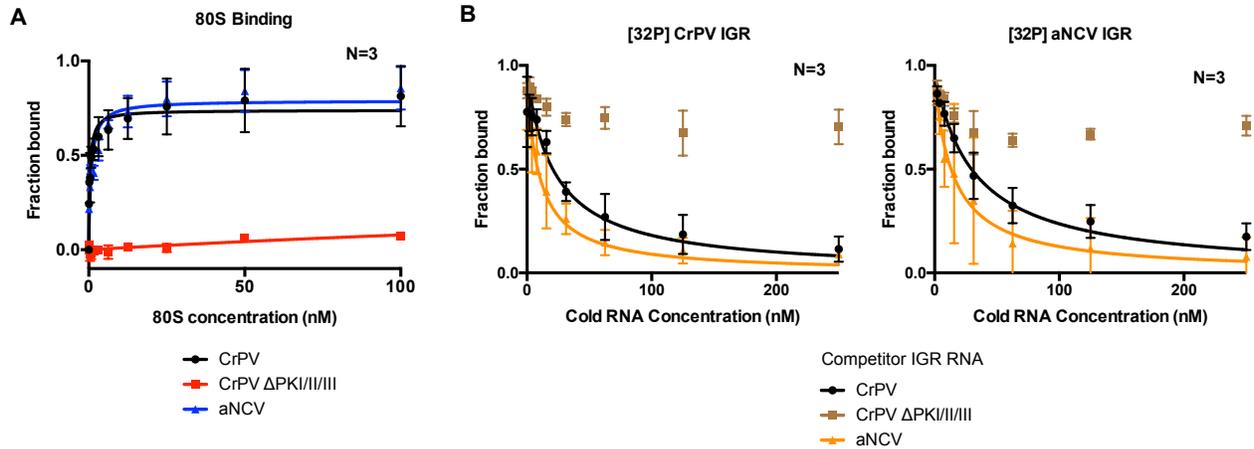


Figure 2.2. Affinity of 80S-aNCV IGR IRES complexes.

(A) Filter binding assays. [32 P]-aNCV IGR IRES, cricket paralysis virus (CrPV) IGR IRES or mutant (Δ PKI/II/III) CrPV IGR IRES (0.5 nM) were incubated with increasing amounts of purified salt-washed 80S. The fractions bound were quantified by phosphor-imager analysis. (B) Competition assays. Quantification of radiolabeled 80S-CrPV IGR IRES (left) or 80S-aNCV IGR IRES (right) complex formation with increasing amounts of cold competitor RNAs (aNCV IGR IRES, CrPV IGR IRES or mutant (Δ PKI/II/III) CrPV IGR IRES). Shown are the averages \pm standard deviation from at least three independent experiments.

2.3.3 RNase protection analysis of aNCV IGR-ribosome complexes

We next examined whether aNCV IGR RNA binds within the inter-subunit space of the ribosome. We previously developed a novel ribosome protection assay, whereby localization of the RNA in the inter-subunit space or solvent side of the ribosome can be inferred by its susceptibility to RNase I-mediated degradation (data not shown). Radiolabeled CrPV IGR IRES in complex with purified ribosomes were incubated with RNase I and then the RNA was isolated and resolved on a urea-PAGE gel. RNase I treatment of radiolabeled CrPV IGR IRES prebound to the ribosome resulted in faster migrating fragments, indicative of sequences that were protected by the ribosome from degradation (**Figure 2.3**). Specifically, compared to the full-length CrPV IGR IRES (188 nucleotides), RNase I treatment led to ribosome-protected fragments ranging from 50–160 nucleotides. Sequencing of these IRES fragments and mapping back to the IRES revealed a signature core domain consisting of PKII and PKIII structures that were primarily protected by the ribosome from RNase I (unpublished work). Importantly, RNase I treatment of the mutant Δ PKI/II/III CrPV IGR IRES (TM) incubated with ribosomes did not lead to protected fragments, indicating that the concentration of RNase I is sufficient to degrade the unprotected RNA. Similar to that observed with the CrPV IGR IRES, RNase I-treatment of aNCV IGR-ribosome complexes resulted in faster migrating fragments (**Figure 2.3**). Compared to the full-length aNCV IGR IRES (291 nucleotides), several protected fragments ranging from 70 to 200 nucleotides were detected. These results indicate that the aNCV IGR binds to the inter-subunit core of the ribosome.

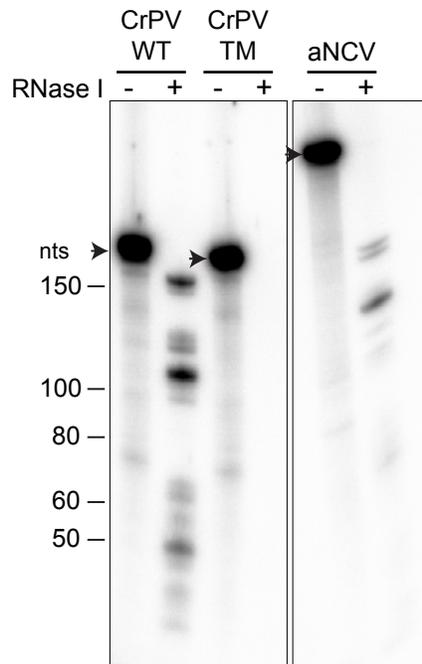


Figure 2.3. aNCV IGR IRES bound to 80S is RNase I-resistant.

Radiolabeled RNAs were incubated with 80S (0.6 μ M) for 20 min before adding RNase I (1 U) for 1 hr. RNAs from RNase I treated/untreated reactions were TRIZOL-extracted and loaded onto urea-PAGE and visualized by phosphor-imager analysis. Shown is a representative gel from at least two independent experiments.

2.3.4 Determination of the aNCV IGR IRES initiation site

We next determined the start site of the aNCV IGR, which is predicted to be at a GCU codon and adjacent to the PKI domain (**Figure 2.1**). To investigate this, we monitored initiating ribosomes assembled on the aNCV IGR in rabbit reticulocyte lysates (RRL) by toeprinting, an established primer extension assay (Nilsen 2013). Briefly, when reverse transcriptase encounters the ribosome assembled on the IRES, a truncated cDNA product is generated, which can be detected on a urea-PAGE gel. As shown previously, ribosomes assembled on the wild-type but not mutant CrPV IGR IRES in RRL in the presence of cycloheximide resulted in toeprints at CC6232-3, which is +20-21 nucleotides from the PKI domain CCU triplet, given that the first C is +1 (**Figure 2.4B**) (Jan and Sarnow 2002). This result showed that the ribosome can translocate on the IGR IRES two cycles in the presence of cycloheximide (J E Wilson, Pestova, et al. 2000; Pestova and Hellen 2003; Jan, Kinzy, and Sarnow 2003). Ribosomes assembled on the aNCV IGR in the presence of cycloheximide resulted in a prominent toeprint at U308, which is +20 nucleotides from the AUC codon, given that A is +1, which is consistent with ribosomes having translocated two cycles (**Figure 2.4B**). To confirm that this toeprint is representative of translocated ribosomes on the aNCV IGR, mutations within the PKI domain were generated that disrupted PKI base-pairing (mPKI #1 and #2) or compensatory (comp) mutations that restored PKI base-pairing (**Figure 2.4B**). Toeprint U308 was only detected when the PKI domain is intact, thus supporting the conclusion that the initiating codon is the GCU alanine adjacent to the PKI of aNCV IGR.

To further confirm direct ribosome binding on the IRES, we performed toeprinting analysis of purified ribosomes assembled on the aNCV IGR IRES. As shown previously, purified ribosomes assembled on the wild-type but not mutant CrPV IGR IRES resulted in toeprints at CA6226-7, which is +13-14 nucleotides from the PKI domain CCU triplet, given that the first C

is +1 (**Figure 2.4C**). Purified ribosomes assembled on the aNCV IGR resulted in a toeprint at C302, which is +13 nucleotides from the AUC codon, given that A is +1 (**Figure 2.4C**), thus supporting the conclusion that ribosomes initially assembled on the aNCV IGR contain the AUC codon in the ribosomal A site and subsequently, after the first pseudo-translocation step, the adjacent start site GCU codon occupies the A site where IRES initiation starts from.

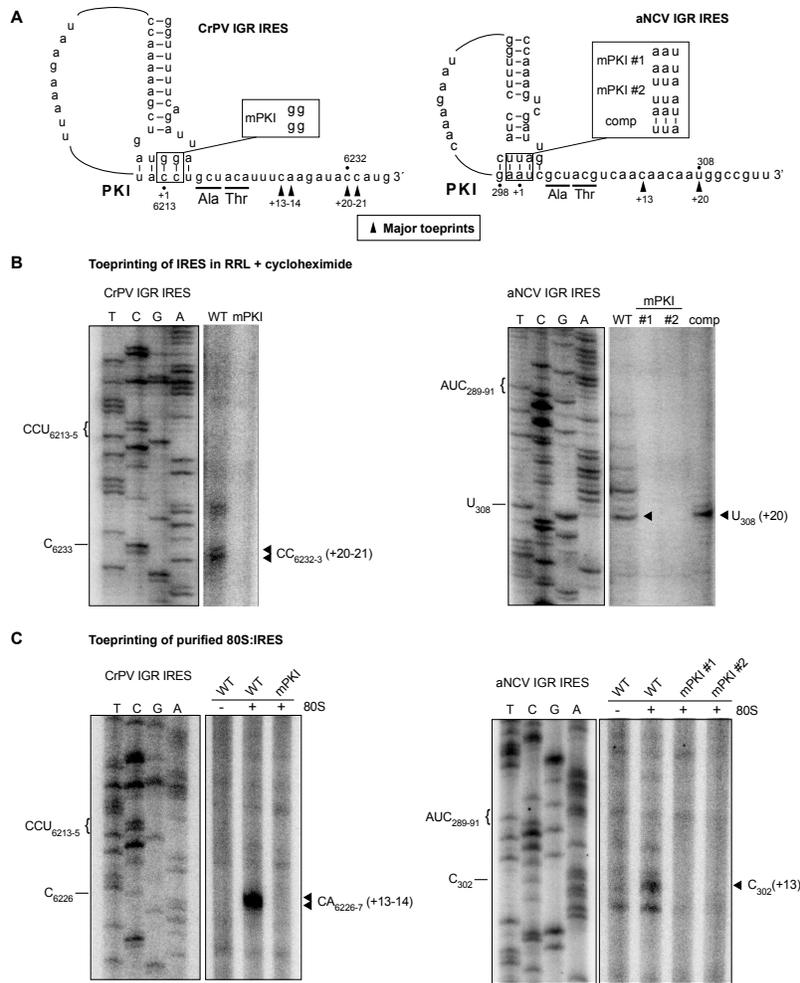


Figure 2.4. Toeprinting analysis of 80S-aNCV IRES complexes.

(A) Schematic of IGR IRES PKI region. Mutations within the PKI region are annotated. (B) Primer extension analysis was performed on in vitro transcribed bicistronic RNAs containing the indicated wild-type or mutant CrPV IRES (left) or aNCV IRES or CrPV IRES (right) incubated in rabbit reticulocyte lysate (RRL) pretreated with cycloheximide (0.68 mg/mL). For the aNCV IRES, the sequencing ladder corresponds to nucleotides 279–331. The major toeprint is indicated at right, which is twenty nucleotides downstream of AUC₂₈₉₋₉₁ of PKI given that the A is +1. (C) Primer extension analysis on in vitro transcribed bicistronic RNAs containing the indicated wild-type or mutant aNCV IRES incubated with purified ribosomes. Sequencing ladder corresponds to aNCV IRES nucleotides 269–319. The major toeprint is indicated at right, which is 13 nucleotides downstream of AUC₂₈₉₋₉₁. Shown is a representative gel from at least three independent experiments.

2.3.5 aNCV IGR IRES directs translation in vitro

To determine whether the aNCV IGR has IRES activity, we inserted the aNCV IGR in a bicistronic reporter RNA construct and assessed translation in RRL (**Figure 2.5A, B**). In vitro transcribed RNA was incubated in RRL in the presence of [³⁵S]-methionine/cysteine to monitor protein expression level. Wild type but not mutant (mPKI) aNCV IGR IRES was active in RRL in vitro (**Figure 2.5A** lane 3). Importantly, mutating the aNCV PKI region in the reporter abolished FLuc activity, indicating that IRES activity is compromised (**Figure 2.5A**, lanes 4 and 5). In contrast, compensatory mutations that restored base-pairing rescued aNCV IGR IRES activity to ~50% of wild type (**Figure 2.5A**, lane 6). Unlike some dicistrovirus IGR IRESs that can direct +1 frame translation (Ren et al. 2012; Kerr et al. 2018), the aNCV IGR IRES does not do this in RRL (data not shown). In summary, the aNCV IGR is a bona fide IRES.

To investigate whether the upstream region of the aNCV IRES is important for IRES activity, we monitored translation of wild type and a mutant aNCV IRES, where the upstream region has been deleted (Δ 1-99). In RRL, both wild type and mutant (Δ 1-99) aNCV can direct IRES activity to similar levels (**Figure 2.5B**), indicating that this upstream region does not contribute to IRES activity in vitro.

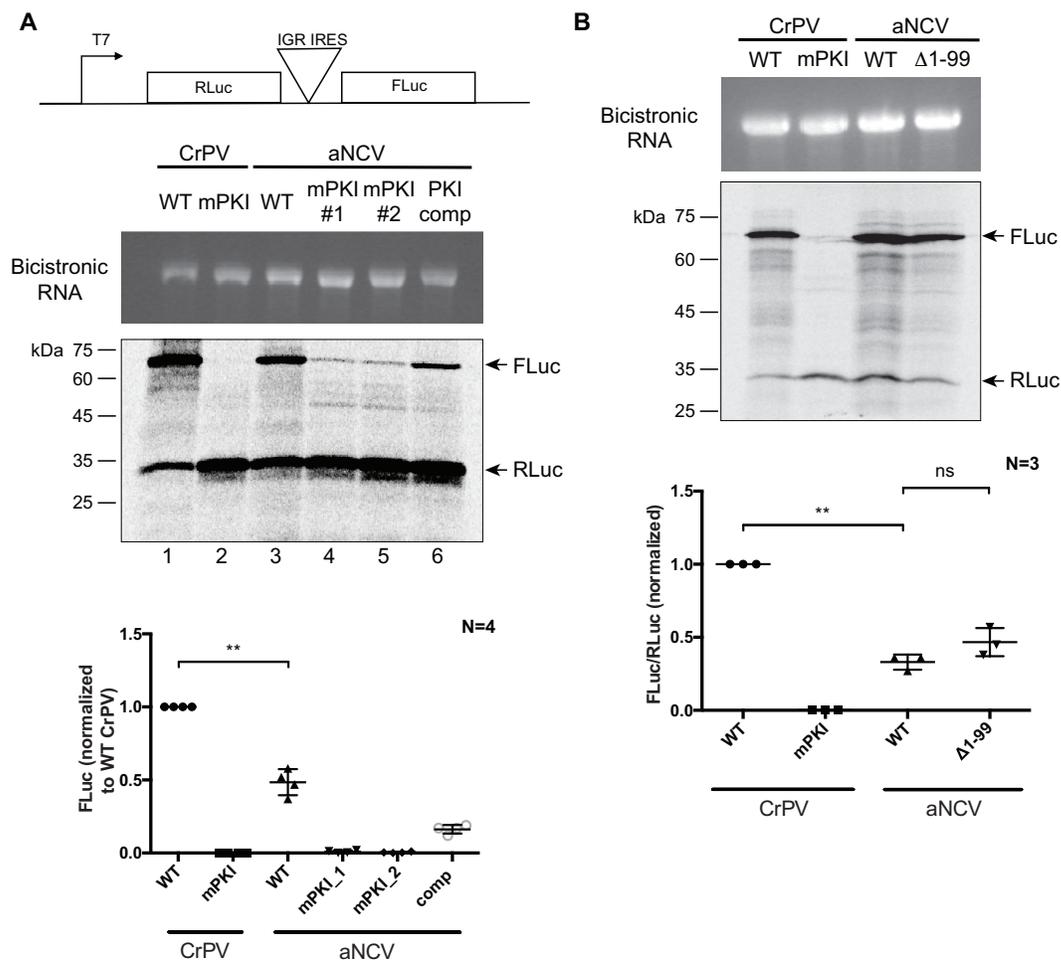


Figure 2.5. In vitro translation assays in RRL.

(A) Schematic of bicistronic reporter construct containing the IRES within the intergenic region. (B) In vitro-transcribed bicistronic reporter RNAs were incubated in rabbit reticulocyte lysate (RRL) for 60 min in the presence of [³⁵S]-methionine/cysteine. The reactions were loaded on an SDS-PAGE gel, which was then dried and imaged by phosphor-imager analysis. (top) Integrity of the in vitro transcribed RNAs are shown. (bottom) Quantification of the radiolabeled Renilla (RLuc) and firefly (FLuc) luciferase proteins. ns, $p > 0.05$; **, $p < 0.005$. Shown are the averages \pm standard deviation from at least three independent experiments.

2.3.6 Chimeric CrPV clone containing the aNCV IGR is infectious

Having demonstrated that the aNCV IGR can bind to ribosomes directly and drive IRES activity from a non-AUG codon, we next examined whether the aNCV can support infection in a more physiological system. Because only a partial sequence of the aNCV genome was recovered in the 700-year-old caribou feces (Ng et al. 2014), we used a heterologous approach by generating chimeric dicistrovirus CrPV clones by replacing IGR IRES with either the full-length aNCV (CrPV-aNCV) or the mutant aNCV (Δ 1-99), where the upstream 99 nucleotides is deleted (Au, Elspass, and Jan 2017; Kerr et al. 2015) (**Figure 2.6A**). We first addressed whether the aNCV IGR can support translation in the infectious clone. In Sf21 extracts, incubation of in vitro transcribed chimeric CrPV-aNCV led to translation of non-structural and structural proteins that are similar to that of the wild-type CrPV infectious clone (**Figure 2.6B**). Of note, the chimeric CrPV-aNCV containing either the full-length or the mutant aNCV (Δ 1-99) led to similar expression of viral proteins in vitro, consistent with previous data showing that the upstream 99 nucleotides do not affect aNCV IRES activity. Compared to the CrPV clone RNA, translation of structural proteins VP1 and VP2/3 was compromised in reactions containing the full-length and the mutant chimeric CrPV-aNCV RNA, indicating a defect in aNCV IGR IRES translation in vitro under these conditions (**Figure 2.6B**). The mutant CrPV containing a stop codon within ORF1 only resulted in expression of the unprocessed ORF2 polyprotein (Kerr et al. 2015).

To determine whether the CrPV-aNCV clone is infectious, we transfected in vitro transcribed RNA into S2 cells and monitored viral protein expression by immunoblotting. Transfection of the wild-type CrPV and CrPV-aNCV containing the full-length aNCV IGR RNA resulted in expression of the CrPV non-structural protein 1A and the structural protein VP2 at 120 h post-transfection (h.p.t) (**Figure 2.6C**), which from previous experience is indicative of

productive infection (Kerr et al. 2015). By contrast, viral proteins were not detected in cells transfected with the CrPV-aNCV (Δ 1-99) clone. To confirm these results, we monitored viral replication by RT-PCR analysis. Mirroring the viral protein expression, CrPV RNA was detected by RT-PCR after transfection of the CrPV-aNCV and CrPV RNA but not the CrPV-aNCV (Δ 1-99) RNA (**Figure 2.6D**). Note that the CrPV RNA was detected at 72 h.p.t. whereas CrPV-aNCV was not detected until 144 h.p.t., suggesting that the replication of the chimeric virus is delayed.

We attempted to propagate the CrPV-aNCV virus from transfected cells by reinfecting and passaging in naïve S2 cells. Despite multiple attempts, the CrPV-aNCV (Δ 1-99) did not lead to productive virus as measured by viral titres. However, the CrPV-aNCV yielded productive virus (1.15×10^{10} FFU/mL). We sequence verified that the aNCV IGR was intact in viral RNA isolated from the propagated CrPV-aNCV (data not shown). To further validate virus production, we infected S2 cells with CrPV-aNCV and monitored viral expression by immunoblotting. Similarly to that observed with CrPV infection, the CrPV 1A protein produced from CrPV-aNCV was detected at 2-10 h.p.i., albeit VP2 expression was reproducibly detected until 4 h.p.i., thus likely reflecting the decreased IRES activity of aNCV compared to CrPV (**Figure 2.6E**). In summary, we have demonstrated that the full-length aNCV IGR can support virus infection in a heterologous infectious clone.

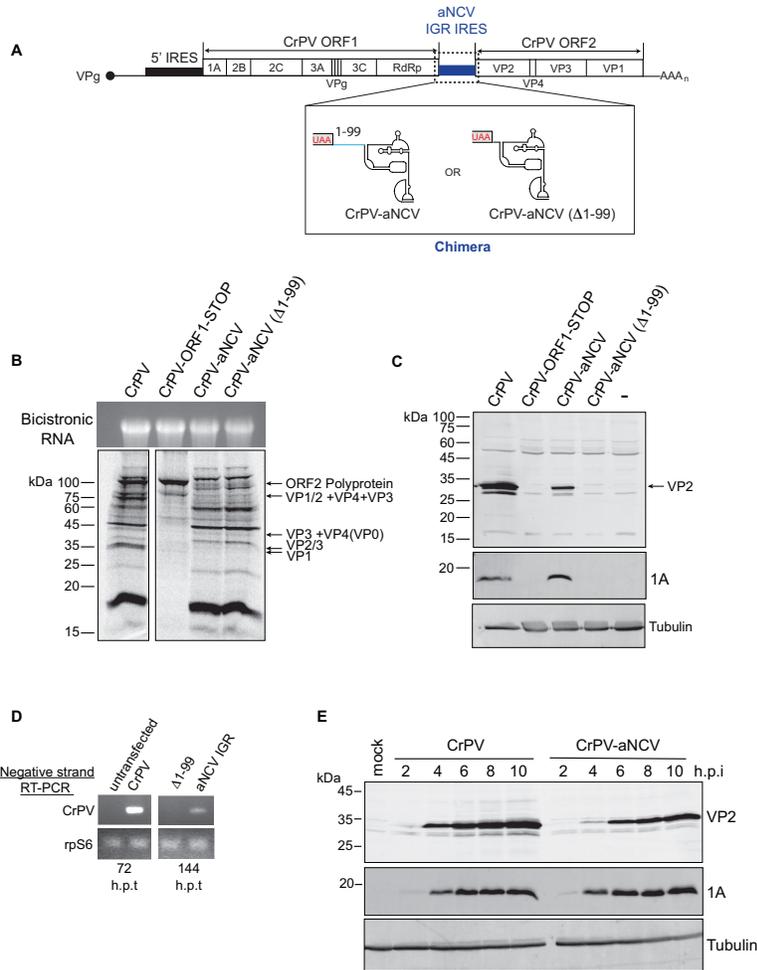


Figure 2.6. Chimeric CrPV-aNCV clone is infectious in *Drosophila* S2 cells.

(A) Schematic of the chimeric CrPV-aNCV clone replacing CrPV IGR IRES with that of the aNCV IRES. (B) In vitro translation of CrPV and CrPV-aNCV RNAs in Sf21 extracts. CrPV-ORF1-STOP contains a stop codon within the N-terminal ORF1, thus preventing expression of the non-structural proteins. Reactions were resolved by SDS-PAGE and visualized by autoradiography. (C) Immunoblotting of CrPV VP2 structural protein and 1A non-structural protein. (D) RT-PCR of viral negative-strand RNA from *Drosophila* S2 cells transfected with the indicated viral clone RNAs at 72 and 144 h after transfection. (E) Immunoblotting of CrPV 1A and VP2 proteins from lysates of *Drosophila* S2 cells infected with CrPV or CrPV-aNCV chimera (MOI 5) at the indicated h.p.i. Shown are representative gels from at least three independent experiments.

2.4 Discussion

Tracking back and identifying ancient RNA viruses may provide insights into the evolution and origin of present-day viruses including the mechanisms that permit virus translation, replication and host virus interactions. In this study, we have examined the IGR of an ancient dicistrovirus that was discovered from 700-year-old ice-preserved caribou feces. To our knowledge, the aNCV genome has not been identified in RNA metagenomic analysis; thus, this study is the first to characterize the oldest IRES to date and provides insights into the origins of this type of IRES. Using molecular and biochemical approaches, we demonstrate that the aNCV IGR possesses IRES activity using a mechanism that is similar to that found in present day Type IV dicistrovirus IRESs. The aNCV IGR can direct ribosome assembly directly and initiates translation from a non-AUG codon. The aNCV IGR IRES can support virus infection in a heterologous infectious clone; thus, a functional aNCV IGR IRES has been resurrected from an ancient RNA virus.

The secondary structure model of the core aNCV IGR resembles classic Type I IGR IRESs containing all three PKs. However, there are subtle differences, including the L1.1 bulge region, responsible for 60S recruitment, which is comprised of a mixture of Type I and II IRESs elements. Additionally, aNCV IGR contains 105 nucleotides upstream of the core IGR IRES, within which several stem-loops are predicted. We showed that this upstream sequence is not required for aNCV IGR IRES translation (**Figure 2.5, 2.6**) but is required for virus infection using the CrPV-aNCV. RT-PCR analysis of the viral RNA suggests that the upstream region has a role in replication; however, it may have other roles in the viral life cycle such as RNA stability or replication. Of note, a few dicistrovirus IGRs also contain sequences beyond the core IRES structure. For example,

the Rhopalosiphum padi virus (RhPV) contains an IGR that is over 500 nucleotides. Similar to the aNCV IRES, the extra sequences within the RhPV IGR do not affect the Type I RhPV IGR IRES (Domier, McCoppin, and D'arcy 2000). It will be interesting to investigate in more detail how these extra IGR sequences play a role in the viral life cycle.

To date, there are increasing numbers of dicistrovirus-like genomes identified via metagenomic studies (Shi et al. 2016; Wolfe, Dunavan, and Diamond 2007). From limited analysis, it is clear that the IGR IRES structures can be distinct (i.e., Type I and II) and can have novel functions (i.e., +1 reading frame selection). For example, the Halastavi árva virus IRES uses a unique but similar streamlined mechanism as the CrPV IRES, where it bypasses the first pseudo-translocation event to direct translation initiation (Abaeva et al. 2020). Moreover, the IGR IRES can direct translation in multiple species (Colussi et al. 2015; Thompson, Gulyas, and Sarnow 2001; J E Wilson, Powell, et al. 2000), suggesting that the IGR IRES mechanism may be a molecular fossil of an RNA-based translational control strategy that existed in ancient viruses. IGR IRESs are likely to have evolved through recombination of independent functional domains; the PKI domains can be functionally swapped between Type I and II IRESs (C. J. Jang and Jan 2010; Hertz and Thompson 2011). Furthermore, IRES elements in unrelated viruses appear to have disseminated between these viruses via horizontal gene transfer (Arhab et al. 2020), further supporting the idea that dicistrovirus IGR IRESs may have originated via recombination events.

Given that the aNCV IGR IRES structure resembles an IGR IRES found in contemporary dicistrovirus genomes, it was not surprising that it functions by a similar mechanism. Relatively speaking, the IGR IRES from a 700-year-old RNA viral genome is far from the presumed primordial IRES. However, this study provides proof-in-concept that this viral RNA translation mechanism can be resurrected and investigated to provide context in the evolution of viral

mechanisms. The challenge in identifying more ancient RNA viral genomes and mechanisms has been and will always be in capturing intact RNA viral genomes as RNA, compared to DNA virus counterparts, are unstable and apt to degrade over time. The remarkable discoveries by Ng et al. and others of ancient RNA viral genomes (Ng et al. 2014; Smith et al. 2014; Peyambari et al. 2019) provide hope in the pursuit of identifying more ancient viruses that shed light into the origins of contemporary viral mechanisms.

Chapter 3: The search for *Drosophila* RNAs that have viral IGR IRES-like properties

3.1 Introduction

The first IRESs were discovered in PV and EMCV in 1988 (S. K. Jang et al. 1988; Pelletier and Sonenberg 1988). Given that important cellular mechanisms including splicing were first discovered from virus systems (Berget, Moore, and Sharp 1977; L. T. Chow et al. 1977), it was reasonable to hypothesize that cellular IRESs exist. Indeed, the first cellular IRES was identified from the mRNA encoding the immunoglobulin heavy chain binding protein (Sarnow 1989; Macejak and Sarnow 1991). Furthermore, the search for IRES elements in *Drosophila* genes lead to the discovery of an IRES within the homeotic gene *Antennapedia*, which controls the formation of legs during embryonic development (Oh, Scott, and Sarnow 1992). The detailed mechanism of how cellular IRESs initiate translation remains to be investigated. However, it has been reported that eIF4E or intact eIF4G is not required in some cellular IRESs mediated translation (Spriggs et al. 2009). Furthermore, regulation of the translation factor, eIF2 α via phosphorylation, also did not inhibit protein synthesis of certain mRNAs containing IRESs (Gerlitz, Jagus, and Elroy-Stein 2002; Subkhankulova, Mitchell, and Willis 2001).

Host and virus are proposed to co-evolve via genetic recombination. Many viral elements are found to be imbedded in the host genome including fungus, plant and animal (Stedman 2015). The insertion of viral elements in cellular genomes appears to be a relatively recent event but likely continuously evolved over a million years (Horie et al. 2010; 2013). It has been hypothesized that by incorporating a virus element within the host genome, this may provide an antiviral defense

such that the host is resistant to that virus infection (Stedman 2015). Honey bee genomes contain parts of the Israeli acute paralysis virus (IAPV) genome that are thought to provide virus-resistance (Maori, Tanne, and Sela 2007). *Drosophila* containing cDNA fragments of viruses inserted within the host genome can modulate virus infection via the RNA interference pathway (Goic et al. 2013). It was also shown that viral RNA can combine with the host RNA to increase pathogenicity (Khatchikian, Orlich, and Rott 1989; Meyers et al. 1991).

The IGR IRES, which contains the most streamlined translation mechanism, has only been identified exclusively in the dicistrovirus family. So far, no cellular IGR IRESs have been discovered. However, it is possible that fragments of the IGR IRES-like structure RNA may exist in host mRNAs/genomes (Hatakeyama et al. 2004). Computational approaches have attempted to identify IRES elements in eukaryotic mRNAs, but given that IRESs typically do not share a sequence similarity, it has been challenging to use bioinformatics to identify them (Mignone et al. 2005; J. Wang and Gribskov 2019). One alternative approach would be to exploit the unique properties of the IGR IRES, in particular direct IRES binding to the ribosome within the intersubunit conserved ribosomal core. Towards this, previous work by a postdoc in the Jan lab, Dr. Qing S Wang, used an adapted selected evolution approach (SELEX) to identify RNA elements in the *Drosophila* genome that can bind to purified human ribosomes. This approach was inspired by a similar SELEX-adapted screen that successfully identified a Hepatitis D virus ribozyme-like sequence, a self-cleaving catalytic RNA, in the human *CEPB3* gene (Author et al. 2006). Addressing whether a sequence within the host genome has an IGR IRES-like property may reveal a common core mechanism for ribosome recruitment and may provide insights into the evolution of this unusual mechanism. Therefore, my hypothesis is that cellular genomes contain IGR IRES-like elements. This chapter will focus on characterizing and validating the candidates identified from

the screen using biochemical approaches. Among candidates being validated, three of them showed promising results, which are named RNA3, RNA5, RNA7 respectively. All of three candidates showed tight binding affinity to human ribosomes, and only RNA5 can compete with CrPV IGR IRES for ribosome binding. Using RNA5 as a model, I generated mutations that disrupted predicted stem-loops to determine whether these structures are important for ribosome binding. Finally, I tested whether RNA5 could affect translation of reporter RNAs in an *in vitro* system. In summary, I have developed a pipeline consisting of biochemical techniques that can systematically test candidate for properties that are IGR IRES-like.

3.2 Materials and methods

3.2.1 Plasmids and constructs

The candidate sequences were synthesized (Integrated DNA Technologies) and cloned into pCRTM4-TOPOTM vector (Thermofisher), per the manufacturer's instructions. Mutant RNA5 constructs were generated using PCR-based mutagenesis. All constructs were verified by sequencing.

3.2.2 *In vitro* transcription

Plasmids containing candidates were linearized with NotI or PstI for positive- or negative sense RNA synthesis, respectively. RNAs were *in vitro* transcribed using a bacteriophage T7 RNA polymerase reaction. Radiolabeled RNAs were bulk labelled by incorporating α - [³²P] UTP (3000 Ci/mmol). RNA was purified using a RNeasy Kit (Qiagen). The integrity and purity of RNAs were confirmed by agarose gel analysis. For purifying full-length RNA, *in vitro* transcription reactions were loaded onto 6% (w/v) polyacrylamide/8M urea gel and following phosphorimage analysis (Typhoon, GE life sciences), the main RNA band was cut out and gel eluted by incubating in 4X

gel volume of 300mM NaOAc at 4 °C rotating overnight. RNA was precipitated in 2.5X volume of ethanol containing 1 µL RNA grade glycogen at -20 °C for 2 hr. After centrifugation, the RNA pellet was washed 2X 75% ethanol and then resuspended in DEPC-treated water.

3.2.3 Purification of the 40S and 60S subunits

Ribosomal subunits were purified from HeLa cell pellets (National Cell Culture Centre) as described (Jan and Sarnow 2002). In brief, HeLa cells were lysed in a lysis buffer (15 mM Tris-HCl (pH 7.5), 300 mM NaCl, 6 mM MgCl₂, 1% (v/v) Triton X-100, 1 mg/ml heparin). Debris was removed by centrifuging at 23,000 g and the supernatant was layered on a 30% (w/w) cushion of sucrose in 0.5 M KCl and centrifuged at 100,000 g to pellet crude ribosomes. Ribosomes were gently resuspended in buffer B (20 mM Tris-HCl (pH 7.5), 6 mM magnesium acetate, 150 mM KCl, 6.8% (w/v) sucrose, 1 mM DTT) at 4 °C, treated with puromycin (final 2.3 mM) to release ribosomes from mRNA, and KCl (final 500 mM) was added to wash and separate 80S ribosomes into 40S and 60S. The dissociated ribosomes were then separated on a 10%–30% (w/w) sucrose gradient. The 40S and 60S peaks were detected by measuring the absorbance at 260 nm. Corresponding fractions were pooled, concentrated using Amicon Ultra spin concentrators (Millipore) in buffer C (20 mM Tris-HCl (pH 7.5), 0.2 mM EDTA, 10 mM KCl, 1 mM MgCl₂, 6.8% sucrose). The concentration of 40S and 60S subunits was determined by spectrophotometry, using the conversions 1 A_{260 nm} = 50 nM for 40S subunits, and 1 A_{260 nm} = 25 nM for 60S subunits.

3.2.4 Filter-binding assays

The radiolabeled candidate RNAs (final 0.5 nM) were preheated at 65 °C for 3 min, followed by the addition of 1× Buffer E (final concentration: 20 mM Tris pH 7.5, 100 mM KCl, 2.5 mM MgOAc, 0.25 mM Spermidine, and 2 mM DTT) and slowed cooling in a water bath

preheated to 60 °C for 20 min. The prefolded RNAs, with 50 ng/μL of non-competitor RNAs, were incubated with increasing amounts of 40S subunits from 0.1 nM to 100 nM, and 1.5 fold excess of 60S subunits for 20 min at room temperature. Non-competitor RNAs were *in vitro* transcribed from the pcDNA3 vector 880-948 nucleotides. Reactions were then loaded onto a Bio-Dot filtration apparatus (Bio-Rad) including a double membrane of nitrocellulose and nylon pre-washed with buffer E. Membranes were then washed three times with buffer E, dried, and the radioactivity were imaged and quantified by phosphorimager analysis.

Radiolabeled candidate RNAs (final 0.5 nM), IRES competitors, and non-competitor RNAs (50 ng/μL) were pre-folded in buffer E. Unlabeled IRES competitors were added in increasing concentrations from 2 nM to 250 nM. RNAs were then incubated with 25 nM 40S and 37.5 nM 60S subunits at room temperature for 20 min. Reactions were then loaded onto the Bio-Dot filtration apparatus, and data were fitted to the Linn-Riggs equation that describes competitive ligand binding to the target.

3.2.5 Ribosome protection assay

[³²P]-labeled RNAs (final 0.1 μM) were pre-folded in buffer E and incubated with 0.6 μM 40S and 0.9 μM 60S at room temperature for 20 min. 1 μL of 1U/μL of RNase I (Ambion) was added to the mixture and incubated at 20 °C for 1 hr. RNAs without RNase I treatment were incubated at 20 °C with the same time length. RNAs from mixtures with or without RNase I treatment were TRIZOL-extracted and loaded onto 6% (w/v) polyacrylamide/8M urea gels to separate them. RNA Ladders (Ambion) were synthesized per manufacturer's instruction. Gels were dried and imaged by phosphorimager analysis.

3.2.6 Translation competition in RRL

Uncapped bicistronic RNAs and candidate RNAs were first pre-folded by heating at 65 °C for 3 min, followed by the addition of 1× Buffer E and slowed cooling at room temperature for 10 min. The pre-folded candidate RNAs (0.6 μM, 1.3 μM, 2.6 μM) and the pre-folded bicistronic RNAs (0.02 μM) were incubated in rabbit reticulocyte lysate (RRL) translation extract (Promega) which contained 8 U Ribolock inhibitor (Fermentas), 20 μM amino acid mix minus methionine, 0.3 μl [35S]-methionine/cysteine (PerkinElmer, >1000 Ci/mmol), and 75 mM KOAc pH 7.5 at 30 °C for 1 hr. The translated proteins were resolved using SDS-PAGE and analyzed by phosphorimager analysis.

3.3 Results

3.3.1 SELEX screen identifies potential RNA candidates that can bind directly to ribosomes

The IGR IRES can drive translation by an unprecedented mechanism (Kerr and Jan 2016). Two unique properties of IGR IRESs are direct binding to ribosomes with high affinity and binding within the ribosomal conserved core consisting of the space that tRNAs normally occupy during translation (Jan and Sarnow 2002; Fernández et al. 2014). The IGR IRES mimics tRNA interactions with the ribosome including the anticodon:codon domain. However, the IRES also contains unique domains that interact with the ribosome which are not observed with tRNA binding. PKII and PKIII make extensive interactions with 40S and 60S responsible for recruiting ribosomes. For example, SLIV and SLV from PKIII interact with rps7 and rps25 of 40S, and L1.1 interacts with L1 stalk of 60S (Kerr and Jan 2016). Thus, the IRES has evolved RNA structural elements that can interact and manipulate the ribosome. The origin of these RNA domains is not known. To search for RNA elements that have similar viral IGR IRES properties, previous work

by a postdoc in the lab (Dr. Qing S Wang) used an adapted selected evolution approach (SELEX) to identify elements in the *Drosophila* genome that can bind to purified ribosomes specifically. Here, the *Drosophila* genome was extracted from *Drosophila Schneider* line 2 (S2) cells and randomly digested to short fragments (150-250 nucleotides) by DNase I, which is in the size range of dicistrovirus IGR IRESs (**Figure 3.1A**). The digested fragments were then *in vitro* transcribed into RNAs and incubated with purified human 40S and 60S ribosomal subunits, followed by sucrose gradient centrifugation to isolate the fraction containing 80S ribosomes and thus RNAs that bind to 80S. Fractions containing 80S were collected and RNAs were TRIZOL-extracted. The extracted RNAs underwent RT-PCR to generate cDNAs which were used for the next round selection. In summary, ten rounds of selection were performed. As a proof of principle, after each round of selection, half of amplified candidate RNAs were radiolabeled by α - [³²P] CTP and mixed with purified ribosomes followed by sucrose gradient centrifugation to confirm their binding to 80S ribosomes. As shown in **Figure 3.1B**, in the initial round, almost all of the radioactivity was located in the non-ribosome bound fractions. After the 5th round of selection, ~10% of radiolabeled RNAs were found in fractions 16-19, the 80S ribosome fractions. After round 10, the percentage of radiolabeled RNAs binding to 80S ribosomes was doubled to 20%. This result indicated that RNAs were enriched that have the ability to bind to purified 80S ribosomes. After the last round, the extracted RNAs were divided equally into two pools. RNAs from the first pool were RT-PCR amplified and sequenced by Ion Torrent technology to generate a candidate list that potentially have high binding affinity to ribosomes. RNAs from the second pool were transformed into cDNA using RT-PCR, *in-vitro* transcribed, mixed with ribosomes, and treated with RNase I. As previously shown, IGR IRESs bind to the ribosomal core and are protected from RNase I degradation (X. Wang et al. 2021). With RNase I treatment, only the candidate RNAs that bind

within ribosomes will be protected and any RNAs that bind the solvent side of ribosomes will be degraded. The protected RNA fragments were sequenced by next-generation RNA-Sequencing (RNA-Seq) to generate a second candidate list (Ingolia et al. 2012). Ribosomal RNAs (rRNAs) were removed during cDNA library preparation.

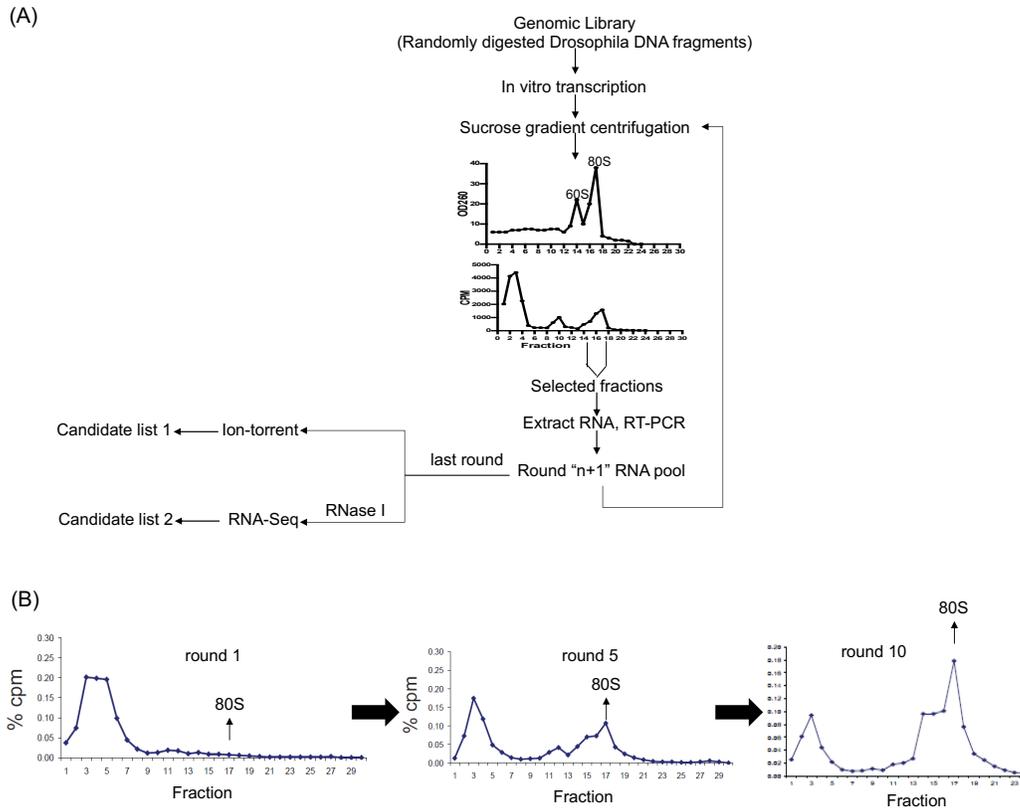


Figure 3.1. Selection IGR IRES-like RNA elements in genomic RNA library

(A) Pipeline of an adapted SELEX of genomic RNAs that bind to 80S ribosomes. Candidates were generated by extracting *Drosophila* genome which was digested into short fragments (average 200 nt) followed by *in vitro* transcription. Candidates that bind to 80S ribosomes were selected after sucrose gradient centrifugation, extracted, reverse-transcribed to cDNA and PCR-amplified using T7 promoter-containing primers for the next round of selection. From the last round, selected candidates together with 80S were either sequenced by Ion-torrent or subjected with RNase I followed by RNA-Seq. (B) Enrichment of genomic RNAs that can bind to 80S ribosomes by sucrose gradient centrifugation. Purified human ribosomes incubated with [³²P]-radiolabeled enriched RNAs were separated by sucrose gradient centrifugation followed by fractionation. Radioactivity in each fraction was normalized to the total radioactivity in all fractions. For screening, RNAs from the 80S ribosomes (fractions 16-18) were isolated and purified. The screen and the sucrose gradient validation were performed by Dr. Qing Wang.

A total of 694 candidates were identified from Ion Torrent and RNA-Seq technologies that identify RNAs that are solvent accessible and RNAs that bind within the ribosomal core. 157 candidates were identified by Ion Torrent technology which represent enriched RNAs that are solvent-accessible and that can bind within the ribosomal core. 591 candidates were identified by RNA-Seq which represent RNA fragments that are protected by the ribosome from RNase I digestion. 54 candidates were found from both methods (**Figure 3.2A**). **Figure 3.2B** summarizes the genome position of candidates from the top 20 RNAs from each method. The candidates were ranked based on the number of reads mapped back to the genome.

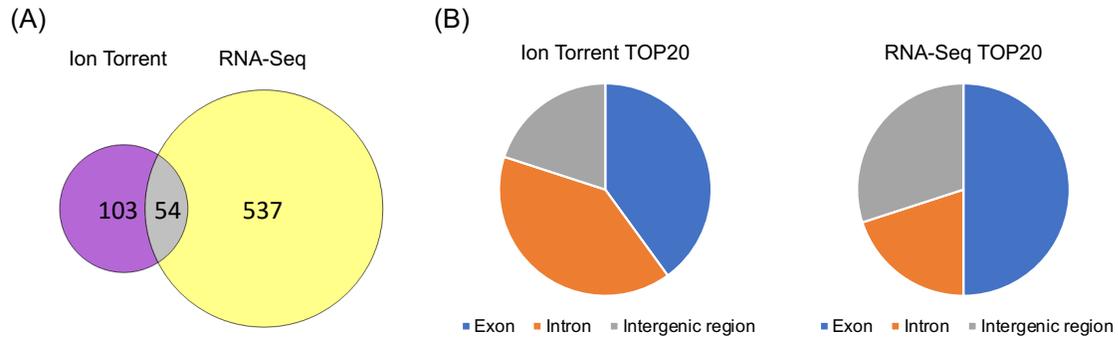


Figure 3.2. Summary of SELEX screen.

(A) Total number of candidates identified from each sequencing method and the overlapping number of candidates discovered by both methods. (B) Venn diagram showing the genomic location of the top 20 candidates from each sequencing method.

A few tRNAs including Glu-tRNA, Ala-tRNA, Lys-tRNA and Leu-tRNA were detected in the enriched RNA bound within the ribosomal core and sequenced by the RNA-Seq approach. This result provided confidence that bona fide RNAs bound within the ribosomal core are enriched. Indeed, it has been shown that tRNAs can bind to ribosome under conditions with high salt requiring no translation factors (Fahlman and Uhlenbeck 2004). However, no tRNAs were detected from the Ion Torrent approach which identified RNAs binding to solvent side or intersubunit of ribosome, which is likely due to the limited sequencing depth of this method (Liu et al. 2012).

To begin validating this approach, I selected high-ranking (i.e., highest number of reads) candidates that were identified by both sequencing methods. The candidate DNAs encoding the RNAs were cloned into a TOPO™ vector flanked with T7 and T3 promoters. I extended the sequence of the candidates by 60 nucleotides as it is possible that RNase I treatment likely digested the full length structure RNA to result in a minimal core RNA fragment that is protected by the ribosome. Both sense and antisense strands of RNAs were synthesized by using either the T7 or T3 promoter. We performed several assays including the ribosome binding assay, the ribosome competition assay and the ribosome protection assay to test whether candidate RNAs act similarly to the viral IGR IRES. We also selected the highest ranked candidate RNAs from the RNase I-protected RNA-seq approach. Their biological function in terms of the effect on translation was tested using an *in vitro* translation system, which will be discussed in **section 3.3.6**.

3.3.2 Ribosome binding of candidate RNAs

One of the most unique characteristics of the IGR IRES mechanism is its high binding affinity to the ribosome (Jan and Sarnow 2002; C. J. Jang and Jan 2010). To test whether the candidate RNAs have this similar property, *in vitro* ribosome binding assays were performed (C.

J. Jang and Jan 2010). RNA was synthesized by *in vitro* transcription reactions using T7 or T3 polymerase. Similar to what had been done in section **section 2.3.2**, [³²P]-labeled RNA was incubated with increasing amounts of purified salt-washed human 40S and 60S ribosomal subunits and the fraction of binding was monitored by filter binding assay (C. J. Jang and Jan 2010). The fraction of ribosome:RNA complexes was calculated as described in section **section 2.3.2**. Briefly, the dissociation constant was calculated from the concentration of ribosomes when the fraction of RNAs bound to the ribosomes was half of the saturated fraction.

In summary, a total of 13 candidate RNAs were tested. Their genomic information including gene position and length of candidates are summarized in **Table 3.1**. Seven of the thirteen RNAs were from the combined list which includes candidates detected from both Ion-Torrent that sequenced candidates that are solvent-accessible and that can bind within the ribosomal core and RNA-Seq that sequenced candidates protected within the ribosomal core from RNase degradation, named as RNA2, RNA3, RNA5, RNA7, RNA8, RNA15 and RNA48. The other 6 candidates were from the list generated by RNA-Seq, named as cut3, cut6, cut8, cut14, cut22, cut28. To maximize the structural integrity of each candidate, I included 30 nucleotides upstream and downstream of the candidate sequence. Only three of the thirteen candidate RNAs, RNA3, RNA5 and RNA7, showed appreciative ribosome binding affinities (apparent K_D 5.9 nM, 3.9 nM, 2.1 nM) and were considered within a similar range as the CrPV IGR IRES-ribosome complex (0.5 nM) (**Figure 3.3A**). RNA3 and RNA7 have a similar saturation point, which is around 0.4 whereas RNA5 has a higher saturation point at 0.6, all of which are lower than that of the CrPV IGR IRES. It is possible that these candidate RNAs are not fully folded properly to allow ribosome binding. As a negative control, I generated the antisense-strand of RNA3, RNA5 and RNA7. None of the antisense RNAs showed any binding to ribosomes and had similar binding as

the defective mutant (Δ PKI/II/III) CrPV IGR IRES, thus demonstrating that the sense candidate RNAs, RNA3, RNA5 and RNA7 bind specifically to the ribosome (**Figure 3.3B**).

Candidate	Gene	Position	Detected Length (bp)
RNA2	CG31145	exon	54
RNA3	Plexus	intron	45
RNA5	Nicotinic acetylcholine receptor	intron	91
RNA7	Serrate	exon	95
RNA8	Mekk1	exon	60
RNA15	Milton	intron	91
RNA48	CG9452	exon	64
cut3		Intergenic region	45
cut6		Intergenic region	38
cut8	Big bang	intron	45
cut14	O/E associated zinc finger protein	exon	27
cut22	CG2852	exon	111
cut28	Heat shock 70-kDa protein cognate 3	exon	104

Table 3.1. Genomic information of candidates being tested from ribosome binding assay.

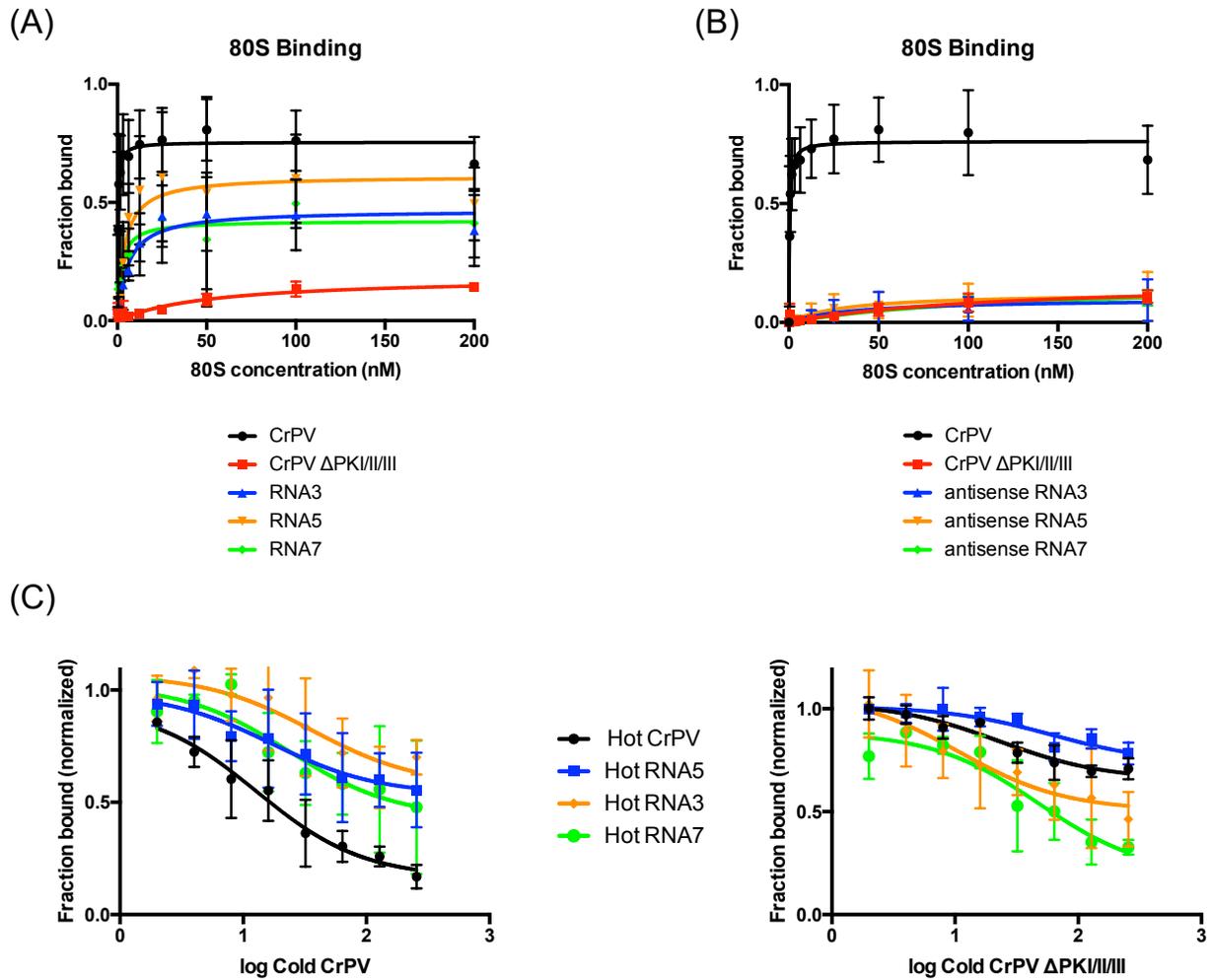


Figure 3.3. Affinity of 80S-RNA candidate complexes.

80S-RNA binding by filter binding assays of (A) [32 P]-candidate RNAs, (B) anti-sense candidate RNAs, wild-type or mutant (Δ PKI/II/III) CrPV IGR IRES (0.5 nM) were incubated with increasing amounts of purified salt-washed 80S (y-axis). (C) Competition binding assays. Quantification of radiolabeled 80S-CrPV IGR IRES or 80S-candidate complex formation (12 nM 80S and 0.5 nM radiolabeled RNA) with increasing amounts of cold competitor RNAs CrPV IGR IRES (left) or mutant (Δ PKI/II/III) CrPV IGR IRES (right) (y-axis). The fraction bound is normalized to reactions with no competitor RNA (x-axis). Shown are the averages \pm standard deviation from at least three independent experiments.

3.3.3 Ribosome competition assays

To further confirm the binding specificity of 80S ribosome-candidate RNA interactions, we applied ribosome competition assays by incubating excess unlabeled wild-type or mutant (Δ PKI/II/III) CrPV IGR IRES to the reaction prior to incubating with ribosomes. As shown previously, incubating increasing amount of competitor wild-type but not mutant (Δ PKI/II/III) CrPV IGR IRES was able to compete for 80S from the radiolabeled CrPV IGR IRES (**Figure 3.3C**, black line). A similar trend was observed in reactions with increasing amounts of wild-type CrPV IGR IRES and radiolabeled RNA5; the fraction of 80S-RNA5 decreased, however, the degree of competition was less (**Figure 3.3C**, blue line). The lack of competition with increasing mutant (Δ PKI/II/III) CrPV IGR IRES for 80S ribosomes from RNA5 indicated that RNA5 binds specifically with 80S (**Figure 3.3C**, right). However, the incomplete competition between wild-type CrPV IGR and RNA5 may suggest that the binding interface of RNA5 on the ribosome may not fully overlap with that of the CrPV IGR IRES. On the other hand, increasing competitor wild-type or mutant (Δ PKI/II/III) CrPV IGR IRES competed for 80S from RNA3 and RNA7 (**Figure 3.3C**, orange/green line), thus indicating that they bound to ribosomes with non-specific interactions. Therefore, as RNA5 showed specific binding to 80S ribosomes, we focused our studies on this candidate RNA.

3.3.4 Mutations in RNA5 alter its binding affinity to ribosomes

The CrPV IGR IRES adopts a structure comprising of overlapping pseudoknots (PKI, II and III) and apical loops of stem-loops to interact specifically with distinct regions of the ribosome (Jan and Sarnow 2002; C. J. Jang and Jan 2010). To assess the possible key ribosome-interaction regions within RNA5, we attempted to predict a secondary structure model of the RNA5

sequence. RNA5 is found within an intron region of nicotinic acetylcholine receptor alpha1. This receptor is involved in responses to the neurotransmitter acetylcholine, and the subunit alpha1 is associated with resistance to neonicotinoid insecticides (Somers et al. 2017). The RNA5 sequence is 91 nucleotides in length. A sequence alignment of RNA5 was also performed to identify sequence conservation among *Drosophila* family and other species (**Figure 3.4**). The RNA5 sequence was uploaded onto mFold, which predicts possible secondary structures of an RNA based on energy minimization (Zuker 2003) (**Figure 3.5A**). Overall, RNA5 has a predicted structure including five SL structures. I chose four different stem-loop regions from the predicted RNA5 structure and generated mutations at these four sites to disrupt basepairing or within the apical loops using site-directed mutagenesis, which are UUG42-4, GCG70-2, CGC121-13, and UUA129-31 (**Figure 3.5A**). The four mutant constructs were tested for ribosome binding using the filter binding assay. As shown, all mutant IRESs were still able to bind to various extents as compared to the wild-type RNA5 (**Figure 3.5B**). Interestingly, mutations at CGC121-3 (apparent K_D 11 ± 2 nM) and UUA129-31 (K_D 9 ± 2 nM) slightly increased the binding affinity compared to wild-type RNA5 (K_D 17 ± 3 nM) and mutation at UUG42-4 (K_D 47 ± 10 nM) slightly decreased binding, whereas mutation at GCG70-2 (K_D 19 ± 3 nM) did not alter 80S binding compared to the wild-type RNA5. This result suggested that the stem at UUG42-4 contributes to interactions with ribosomes.

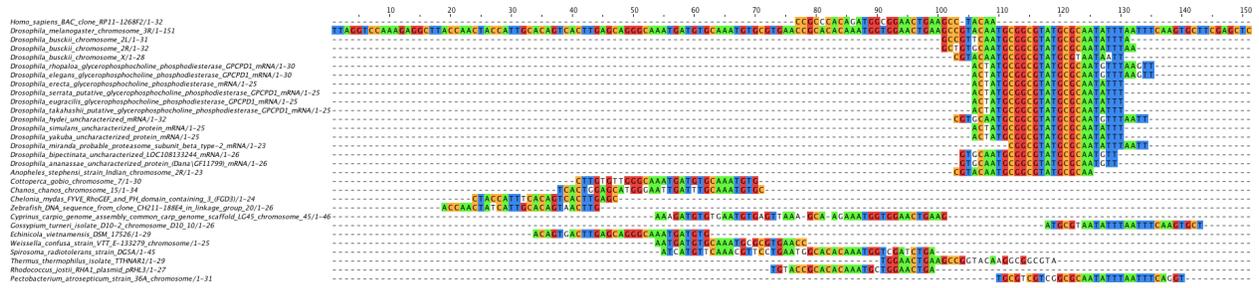


Figure 3.4. Alignment of RNA5 within *Drosophila* and other species.

RNA5 sequence was blasted and aligned. The colored regions represent conserved sequences with RNA5.

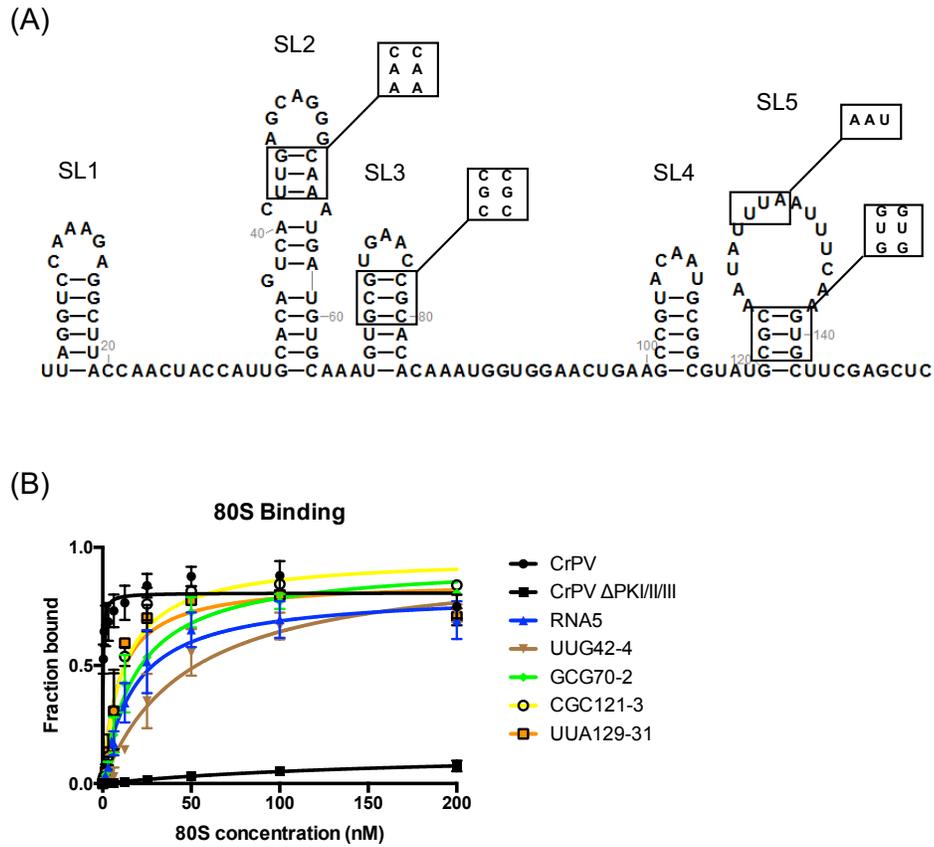


Figure 3.5. Mutations at certain stem and loop regions of RNA5 affect its ribosome binding affinity.

(A) Secondary structure of RNA5 predicted by mFold. Four different locations within RNA5 are mutated to the alternate Watson-crick bases and denoted in the box. (B) Filter binding assays of RNA5 with mutations created at UUG42-4, GCG70-2, CGC121-3, UUA129-31. Shown are the averages \pm standard deviation from at least two independent experiments.

3.3.5 None of the candidates show protection by ribosomes from RNase I

As mentioned earlier, one of the characteristics of IGR IRES is that it binds within the intersubunit space of the ribosome (Fernández et al. 2014; X. Wang et al. 2021). We previously developed a new biochemical assay to monitor structured RNAs that can bind to the ribosomal core (X. Wang et al. 2021). Briefly, radiolabeled RNAs in complex with purified ribosomes were incubated with RNase I, and the protected RNA fragments were isolated and resolved on a urea-PAGE gel. I performed the RNase protection analysis to test whether the candidate RNAs bind within the ribosomal core. I tested all of the candidates that were validated in the ribosome binding assay (RNA3, RNA5, RNA7). All candidate RNAs were not protected by the ribosome from RNase I treatment. It was surprising that RNA5 was not protected as it was shown to bind to ribosomes and compete with CrPV IGR (**Figure 3.6**). This result suggested that RNA5 or part of RNA 5 might be solvent accessible when bound to the ribosome.

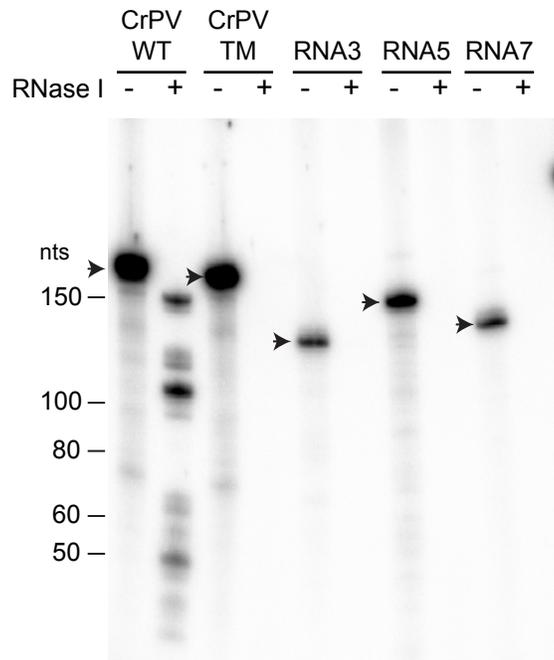


Figure 3.6. RNA candidates bound to the 80S subunit are not resistant to RNase I.

[³²P]-labeled CrPV IGR IRES, mutant (Δ PKI/II/III) CrPV IGR IRES and indicated candidate RNAs were incubated with purified 80S ribosome subunit (0.6 μ M) for 20 min prior to the addition of RNase I (1 U) for 1 hr. RNAs from RNase I treated/untreated reactions were TRIZOL-extracted and resolved on a urea-PAGE and the bands were visualized using phosphorimager. The expected band size for each RNA is pointed. Shown is a representative gel from at least three independent experiments.

3.3.6 RNA5 does not affect translation level in vitro

As RNA5 showed binding to ribosomes, I want to test whether it could have any biological function on translation. There is precedent that RNAs can regulate translation by binding to ribosomes (Pircher et al. 2014; Bakowska-Zywicka, Kasprzyk, and Twardowski 2016). To address this question, a bicistronic construct including a T7 promoter directing canonical Renilla luciferase and a CrPV IGR IRES directing cap-independent Firefly luciferase was used. A fixed amount of bicistronic RNA and increasing amounts of the candidate RNAs were incubated, and the translation level of RLuc/FLuc in the RRL cell-free translation system was monitored using [³⁵S]-methionine/cysteine (**Figure 3.7A**, top). CrPV IGR with mutation at PKI was used as a positive control, as this mutant IRES can bind to ribosomes but is defective in translation initiation. The mutant (Δ PKI/II/III) CrPV IGR was also used as the negative control.

Adding increasing amounts of CrPV IGR Δ PKI (**Figure 3.7A**) decreased the levels of both RLuc and FLuc expression, suggesting that the CrPV IGR IRES can block and sequester ribosomes for translation. The decreased level of RLuc was comparable to the decreased level of FLuc (**Figure 3.7B**). This result suggests that the excess of CrPV IGR sequesters the free pool of ribosomes, and as a result decreases both scanning-dependent/cap-independent translation. I then tested both the positive-sense of RNA5 which showed binding to the ribosome and its negative-sense transcript which was shown not to bind to ribosomes (**Figure 3.7A**, right). Incubation of increasing amounts of RNA5 did not decrease FLuc expression (**Figure 3.7B**). The antisense of RNA5 did not affect FLuc expression neither. Although incubation of RNA5 reduced RLuc expression, the decrease was comparable to that when antisense-RNA5 was incubated, thus indicating that the effect of RNA5 is non-specific. In summary, RNA5 did not affect both scanning-dependent and cap-independent translation in the RRL system.

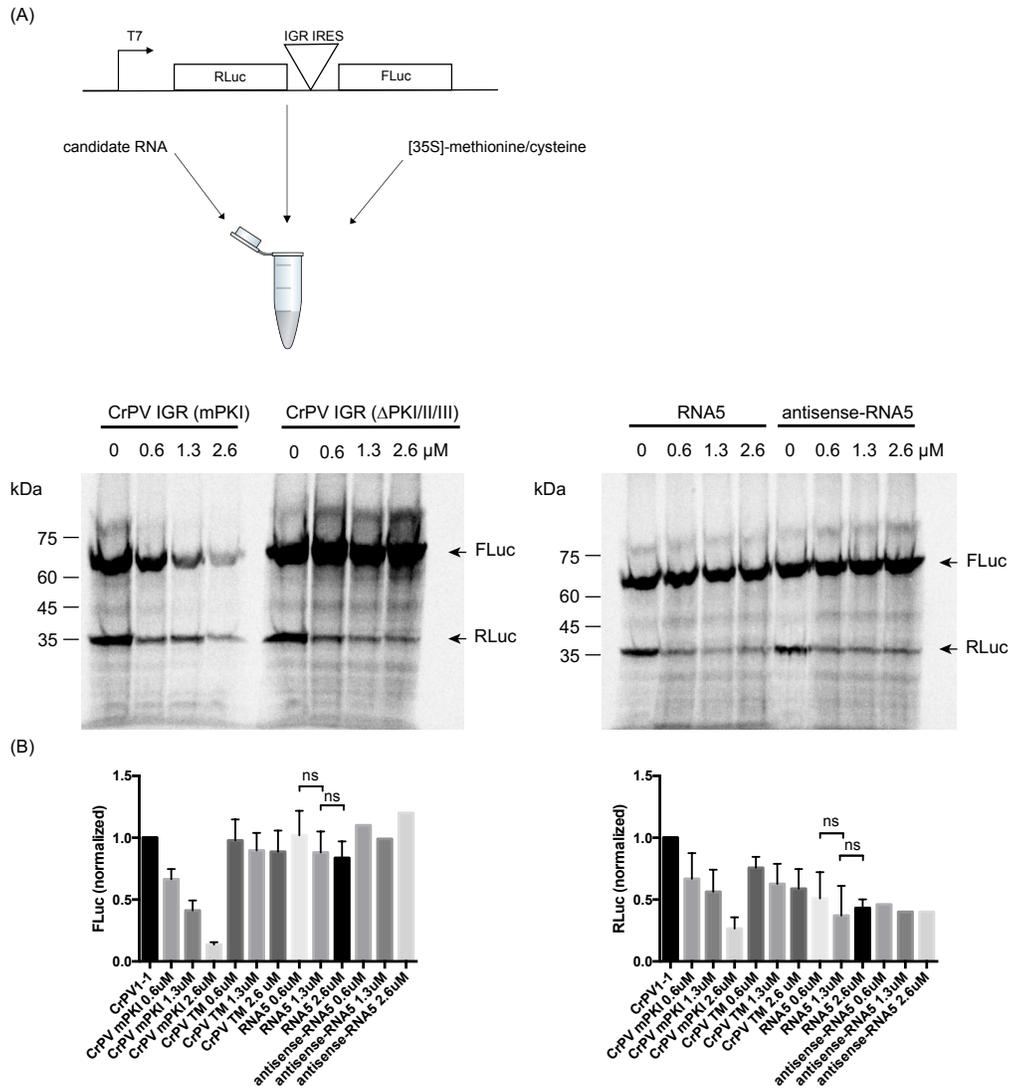


Figure 3.7. Incubation of RNA5 in an in-vitro translation reaction of bicistronic reporter has no impact on reporter RNA translation in RRL.

(A) *Top*. Schematic of bicistronic reporter incubating with RNA5. The upstream RLuc and the downstream FLuc are expressed by scanning-mediated and IRES-mediated translation respectively. *Bottom*. Uncapped bicistronic reporter RNAs (0.02 μ M) with increasing amount of CrPV IGR mPKI, CrPV IGR Δ PKI/II/III, RNA5 or antisense-RNA5 (0.6–2.6 μ M) incubating in RRL for 60 min in the presence of [³⁵S]-methionine/cysteine. (B) Quantification of the radiolabeled Renilla (RLuc) and firefly (FLuc) luciferase proteins, normalized to the luciferase level where no candidate was added. ns, $p > 0.05$. Shown are the averages \pm standard deviation from at least three independent experiments.

3.4 Discussion

The IGR IRES possesses the most streamlined translation mechanism and discovering cellular IGR IRES-like element may lead to insights into the co-evolution between the host and the virus and shed light onto the origin of this unique mechanism. Here we used an adapted selected evolution approach (SELEX) to identify cellular RNA elements that can bind to ribosomes specifically and possibly identify cellular IGR IRES-like element. From the candidate lists generated from the screen, candidates with high read counts were tested. Only RNA3, RNA5 and RNA7 showed binding to purified human ribosomes, and only RNA5 binds specifically to the ribosome. However, RNA5 most likely binds to the solvent side of ribosome.

Although the screen so far did not yield bona fide RNAs that can bind within the ribosome core, I did find RNA5 that can bind to the ribosome and be competed by the CrPV IRES, thus suggesting that part of RNA5 binds to the ribosomal core. Since RNA5 is located in one of the introns of nicotinic acetylcholine receptor alpha1, it is less likely to be directly involved in initiating translation. However, it might have other roles including mediating translation by interacting with the ribosome.

It is well established that the structure of IGR IRESs directs ribosome binding. RNA can bind to ribosomes via interacting with ribosomal proteins and hybridizing with rRNA (Matsuda and Mauro 2014; Argetsinger Steitz and Jakes 1975). In **section 3.3.4** site-directed mutagenesis at different sites within RNA5 was performed to test whether the mutated site is a key ribosome interaction region. Most mutations did not change the binding affinity significantly. UUG42-4 doubled the apparent dissociation constant of RNA5 which suggested a potential interacting region. Searching for the sequence conservation of RNA5 among different species might hint which region

within RNA5 is important. From a Blast search result (**Figure 3.4**), sequences near the 3' end of RNA5 were found to be conserved within *Drosophila* family including *Drosophila busckii* which is native to North America, *Drosophila rhopaloea*, *Drosophila elegans* and so on. Sequences near the 5' end or in the middle were also found to be conserved with other species not belonging to *Drosophila* such as bony fishes, turtles and bacteria that across both animal and bacteria. A partial sequence that was conserved in mosquitos locates at the conserved region within *Drosophila*. Sequences conserved among *Drosophila* species could be interesting, and indeed there is a SL within that region and was not mutated in the first trial. In the first trial of site-directed mutagenesis, UUG42-4, GCG70-2, CGC121-3, and UUA129-31 were mutated. The SL where UUG42-4 is located has multiple basepairing regions, and only mutating three nucleotides at the top of the SL might not be strong enough to break the structure. Mutating both UCA38-40 and UUG42-4 may be needed to completely unfold the SL. There are many possible structures and higher order of structures such as pseudoknots which might not be captured by mFold. Instead, a structural probing approach should be performed in the presence or absence of ribosomes, which will provide structural data analysis and can infer possible conformational changes upon binding to the ribosome.

None of the candidates tested were protected by ribosomes after RNase I digestion and this can be explained by several reasons. The RNAs being detected from the screen after RNase I treatment are fragments of the RNA, thus it is possible that the structured RNAs that can bind to the ribosome is larger than this fragment. To circumvent this, I included ~30 nucleotides upstream and downstream of the RNA fragment but it is possible that long distance tertiary structures may not be captured by this approach. The next step would be expanding the sequences to be included by 100-200 nucleotides upstream and downstream of the candidate. The amount of RNase I and

the incubation time can also be optimized to different candidates since they were based on the titration result from CrPV IGR IRES. Depending on the stability of the folded RNA structure, different candidates could be more susceptible to RNase I degradation if the condition is not optimized even though they bind to ribosomes.

RNAs have been known to regulate translation by targeting mRNAs, including microRNA and small interference RNA. Over the last few decades, various non-protein-coding RNAs (ncRNAs) have been identified as emerging translation regulators by interfering with ribosomes. By directly interacting with ribosomes or polysomes, ncRNAs inhibit global or mRNA specific translation, by interfering with tRNA binding to ribosomal P site or competing with mRNA binding to block translation for example (Pircher et al. 2014; Himeno, Kurita, and Muto 2014; Gebetsberger et al. 2017). ncRNAs have also been shown to stimulate protein synthesis by promoting the recruitment of mRNA during translation initiation (Barbosa, Calhoun, and Wieden 2020). These mechanisms are found in Archaea and eukaryotic systems. There is also a hypothesis that ncRNAs can regulate translation by binding to ribosomes unspecifically and thus represent biological noise (Pircher, Gebetsberger, and Polacek 2014). As I showed RNA5 binding to ribosomes but possibly not fully within the core (**section 3.3.2** and **section 3.3.5**), we wondered whether RNA5 can regulate translation. Incubating RNA5 in RRL translation extracts did not specifically affect IGR IRES or scanning dependent translation. Considering that both the SELEX screen and the binding assay used human ribosomes, it is quite possible that the RNA5 does not bind to rabbit ribosomes, thus a future direction would be to test RNA5 in a human translation system (Hela lysate). ncRNAs can regulate translation by binding to ribosomes under specific stress conditions (Pircher et al. 2014; Bakowska-Zywicka, Kasprzyk, and Twardowski 2016;

Gebetsberger et al. 2012). Thus another direction would be to test RNA5 regulating translation under different stress conditions.

Another finding from this assay is that adding excess of CrPV IRES Δ PKI can inhibit both scanning-dependent/cap-independent translation. This result is intriguing as it suggests that the IGR IRES can sequester ribosomes. In cells, CrPV IGR IRES translation is 0.8% of cap-dependent translation, indicating that CrPV IGR is weak (Q. S. Wang and Jan 2014). Under virus infection, overall cap-dependent translation is shutoff and CrPV IGR IRES translation is stimulated (Garrey et al. 2010). A future direction would be to test the effect of CrPV IGR IRES in cells by transfecting excess of mutant (Δ PKI) CrPV IGR RNA into cells with or without CrPV infection. One idea that we had was to see if we can sequester ribosomes to the CrPV IGR IRES through an RNA with repeated IGR IRESs. This will in effect increase the unit mole of IRES ribosome binding sites on the IRES per RNA. Several challenges exist when designing this RNA, such as how to ensure each IGR IRES folds properly and how long the linker sequence between each CrPV IGR should be. Cloning with multiple repeating sequences is also challenging. Despite all of these concerns, synthesizing an RNA containing multiple CrPV IGR IRESs can be a potential tool to regulate translation during virus infection.

Chapter 4: Conclusions and Future Directions

Viruses have evolved mechanisms to hijack the host translation machinery to synthesize viral proteins. One such mechanism is via recruiting ribosome internally by IRESs to initiate translation (Mailliot and Martin 2018). IGR IRESs from dicistroviruses use the most streamlined mechanism by directly recruiting the ribosome without any initiation factors and initiate the downstream ORF at a non-AUG codon site (J E Wilson, Pestova, et al. 2000; J. Sasaki and Nakashima 2000). The mechanisms of IGR IRESs interacting with ribosomes and initiating translation have been well-studied. The two distinctive domains of IGR IRESs function independently as recruiting ribosomes (PKII and PKIII) and positioning the IRES at the ribosomal A site (PKI) via anticodon:codon mimicry (Mailliot and Martin 2018; Kerr et al. 2016; Pisareva, Pisarev, and Fernández 2018). IGR IRESs have been determined to function similarly across species including yeast, plant, insect and mammalian systems (Colussi et al. 2015; Thompson, Gulyas, and Sarnow 2001; J E Wilson, Powell, et al. 2000). IGR IRESs have also been shown to function in bacteria system, though the mechanism is different (Colussi et al. 2015). As the most unprecedented mechanism by circumventing the requirements for initiation factors, the origin of the IGR IRES ribosome binding structure remains elusive.

With increasing diversity of dicistro-like viral genomes identified by metagenomic approaches (Shi et al. 2016; Wolfe, Dunavan, and Diamond 2007), it is now apparent that there may be different mechanisms utilized by the IGR IRESs. For example, the Halatavi árva virus IRES uses a simpler mechanism, where it bypasses the first pseudotranslocation event to direct translation initiation (Abaeva et al. 2020). Thus indicating that IGR IRESs could have enormous structural change during evolution. Identifying ancient RNA genomes could provide a historical

framework in the importance of this viral translational mechanism. In Chapter 2, I characterized an IGR IRES from a divergent dicistrovirus RNA genome, aNCV, extracted from 700-year-old caribou feces. From the structure prediction, the aNCV IGR adopts a secondary structure similar to contemporary IGR IRESs. Although it has a chimeric sequence of Type I and II IGR IRESs, aNCV IGR IRES overall resembles a Type I IGR IRES. An extra 105 nucleotides upstream of the aNCV IRES was also discovered with unknown function. Sequence alignment of aNCV IGR IRES with classic Type I and II IGR IRESs uncovered the conserved regions including L1.1B and L3 resembling to Type I IGR IRESs and L1.1A resembling to Type II IGR IRESs. Besides the conserved regions, aNCV IGR IRES also exhibits a unique feature at its SLIV with an AUUA sequence which is invariant AUUU sequence in other classified IGR IRESs. Through biomolecular analysis, we proved that aNCV IGR functions as a *bona fide* IRES. Using filter binding assays and competition assays, aNCV IRES was shown to bind to purified human ribosomes specifically with high binding affinity (K_D 0.7 nM). The start site of the downstream ORF was determined to be a GCU alanine codon site via toeprinting analysis using both purified ribosomes and RRL system. By incorporating aNCV IGR into the bicistronic reporter, its ability of initiating translation *in vitro* was confirmed. Finally, aNCV IGR was successfully resurrected by creating a chimeric infectious clone swapping in the aNCV IRES which could both translate and generate infectious virions. Noticeably, only the intact chimera CrPV-aNCV was infectious. The chimera CrPV-aNCV with deleted upstream sequences (Δ 1-99) within IGR IRES failed to produce virions. RT-PCR could only detect the negative strand RNA of the intact CrPV-aNCV. Given that the chimera CrPV-aNCV (Δ 1-99) can support translation *in vitro*, it is suggested that the upstream region may have other roles in the viral life cycle such as RNA stability or replication. It will be interesting to

explore how the upstream sequences affect viral replication by designing a chimera CrPV-aNCV replicon.

Virus-host interaction has been suggested to be a coevolutionary process supporting both host immune adaptation and viral escape strategies. Genetic recombination, which was referred to DNA rejoining, also refers to RNA recombination after evidence of genetic recombination of RNA-based viruses with their host genomes has been provided (Bujarski and Kaesberg 1986; Khatchikian, Orlich, and Rott 1989; Meyers et al. 1991; Goic et al. 2013). Given the tRNA-mimicry of the IGR IRESs, it is intriguing to consider that the IGR IRESs may have emerged through recombination events with the host. However, the sequence diversity within the IRESs is the main challenge to identify IRES-like element in the host genomes using computational approach (Mignone et al. 2005; J. Wang and Gribskov 2019). Given the unique properties of the IGR IRES including the direct binding to the ribosome within the intersubunit conserved ribosomal core, a previous postdoc, Dr. Qing S Wang, used an adapted selected evolution approach (SELEX) to identify RNA elements in the *Drosophila* genome that can bind to purified ribosomes specifically. In Chapter 3, candidates identified from the screen which potentially have properties similar to the IGR IRESs were analyzed, and candidates with high coverage reading numbers were selected for characterizing and validating their binding to the ribosome. Using filter binding assays, the binding affinity of candidates bound to purified human ribosomes was measured and among candidates being tested, RNA3, RNA5 and RNA7 showed high binding affinity (K_D 5.9 nM, 3.9 nM and 2.1 nM). With increasing amounts of competitors added to compete for the ribosome binding, only RNA5 showed decreased binding when wild-type but not mutant (Δ PKI/II/III) CrPV IGR IRES was added. RNA3 and RNA7 ribosome bindings were affected by both wild-type and mutant CrPV IGR IRES to a similar level, which suggested their interactions with ribosomes were non-specific. From the

mutagenesis on RNA5 sequence, the mutations could only affect RNA5 binding affinity moderately. From the four mutations generated, CGC121-3 (K_D 11 ± 2 nM) and UUA129-31 (K_D 9 ± 2 nM) slightly increased RNA5 binding affinity (K_D 17 ± 3 nM), and UUG42-4 (K_D 47 ± 10 nM) decreased its binding affinity by two times difference, suggesting a potential interaction between the stem at UUG42-4 with the ribosome. To address whether RNA5 binds to the ribosomal core, RNase I resistant assay was performed on RNA5 bound to ribosomes. No RNA fragments were detected after incubating RNA5:ribosome complex with RNase I, suggesting that RNA5 likely binds to the solvent face of the ribosome. Finally, the modulation effect of RNA5 on RNA translation was tested, given that RNA fragments targeting ribosomes have been shown to regulate translation through direct interaction (Gebetsberger et al. 2012; Pircher et al. 2014; Fricker et al. 2019). Using an *in-vitro* translation system, the translation level of a bicistronic RNA containing Renilla luciferase directed by a T7 promoter and Firefly luciferase directed by CrPV IGR IRES was measured with increasing amounts of RNA5 added. Overall, RNA5 showed no effect on both scanning-dependent/cap-independent translation using RRL as the translation system. In general, we have set up a comprehensive validating system to characterize ribosome binding of potential candidates and address their effect on translation.

The original SELEX screen might lead to false positive results. The first issue is that the preliminary screen used different species components. Purified human ribosomes were used with the genome library extracted from *Drosophila*. This is because the RNA structural core is universally conserved across the kingdom of life, and translation of IGR IRES was tested in different cell-free systems from mammal to wheat (Colussi et al. 2015; Thompson, Gulyas, and Sarnow 2001; J E Wilson, Powell, et al. 2000). However, there are still differences in ribosomes from different species, including ribosomal proteins (Ban et al. 2014). Therefore, using a uniform

system with both genome and ribosomes extracted from *Drosophila* should be more optimal. The second issue is the choice of reverse transcriptase. The one used in the screen, AMV-RT, conducted several years ago, may not be able to unwind strong secondary RNA structures and thus biasing the result. Recently, reverse transcriptase with high fidelity and processivity to unwind strong secondary RNA structures has been developed including TGIRT and marathonRT. It is being widely used in RNA sequencing applications, including profiling human RNAs, giving full-length reads of tRNAs and structural probing (Zhao, Liu, and Pyle 2018; Xu et al. 2019). The use of a reverse transcriptase with higher processivity will ensure generation of full-length cDNAs and increase the enrichment of the genomic RNA library after each round of selection. The third issue is the identification of the RNAs bound within the ribosomal core after RNase I treatment. RNase I treatment in the screen leads to rRNA fragments that may obscure RNA-Seq reads of the enriched candidate RNAs. In the original screen, after RNA-Seq, 80% of reads were aligned to rRNA and only 0.7% of reads aligned to *Drosophila* genome (data not shown). To circumvent this, 4-thiouridine can be incorporated in the last *in vitro* transcription reaction. After RNase I digestion, ribosome-protected RNAs containing 4-thiouridine can be purified by a biotinylation approach followed by pulldown using streptavidin beads (Garibaldi, Carranza, and Hertel 2017). With an optimized system, the SELEX screen should be able to generate more promising data.

An outstanding question still exists: where and how did the IGR IRES evolve? The tRNA-mimicry region (PKI) of IGR IRESs points to the possibility that one approach of dicistroviruses acquiring IGR IRESs is from the host. The RNA world hypothesis posits that there was a period of time in primitive Earth's history when the primary living substances were RNA (Higgs and Lehman 2015). Given that the cap-dependent translation requires a complicated mechanism involving many factors activities, it is unlikely to be utilized by an early life. By contrast, IRES-

dependent translation requires less factors, and depending on different types, IRES can acquire the most streamlined mechanism which is possible to be utilized by the primitive life. Considering the accessibility of regulatory control presented by cap-dependent translation, it is plausible that the IRES-dependent translation is taken over during evolution.

The ribosome binding domain and the tRNA-like positioning domain of the IGR IRES have been shown to be modular and functionally independent (C. J. Jang and Jan 2010; Hertz and Thompson 2011), suggesting that each domain evolved from different origins and that this chimeric IGR IRES structure was generated through recombination events. Besides previous knowledge about the IGR IRES-dependent mechanism from the well-studied viruses within *Dicistroviridae*, recent novel IGR IRESs with less conserved structures or a non-canonical ORF1 stop codon within the IGR IRES have been identified. For example, the Halastavi árva virus IRES (HaIV) lacks PKIII region which normally is required for the IGR IRES to interact with the 40S subunit. Instead, the HaIV IRES only binds to preformed 80S instead of 40S (Abaeva et al. 2020). Another example is the IGR from Wenling picorna-like virus 4, which has the upstream ORF1 stop codon presented within the IGR sequence. Preliminary data from our lab showed that with the upstream stop codon which potentially will unfold partial structure during translation, this IGR can still function and direct the translation internally, suggesting that this IGR IRES might mediate translation by a different mechanism. The discovery of IGR IRESs with novel strategies to direct the translation will help to comprehend the current knowledge and shed light on the evolution of this viral strategy.

References

- Abaeva, Irina S., Quentin Vicens, Anthony Bochler, Heddy Soufari, Angelita Simonetti, Tatyana V. Pestova, Yaser Hashem, and Christopher U.T. Hellen. 2020. “The Halastavi árva Virus Intergenic Region IRES Promotes Translation by the Simplest Possible Initiation Mechanism.” *Cell Reports* 33 (10): 108476. <https://doi.org/10.1016/j.celrep.2020.108476>.
- Abeyrathne, Priyanka D., Cha San Koh, Timothy Grant, Nikolaus Grigorieff, and Andrei A. Korostelev. 2016. “Ensemble cryo-EM uncovers inchworm-like translocation of a viral IRES through the ribosome.” *ELife* 5 (May): e14874. <https://doi.org/10.7554/eLife.14874>.
- Argetsinger Steitz, J., and K. Jakes. 1975. “How ribosomes select initiator regions in mRNA: base pair formation between the 3’ terminus of 16S rRNA and the mRNA during initiation of protein synthesis in Escherichia coli.” *PNAS* 72 (12): 4734–38. <https://doi.org/10.1073/pnas.72.12.4734>.
- Arhab, Yani, Alexander G. Bulakhov, Tatyana V. Pestova, and Christopher U.T. Hellen. 2020. “Dissemination of Internal Ribosomal Entry Sites (IRES) Between Viruses by Horizontal Gene Transfer.” *Viruses* 12 (6): 612. <https://doi.org/10.3390/v12060612>.
- Au, Hilda H. T., Valentina M. Elspass, and Eric Jan. 2017. “Functional Insights into the Adjacent Stem-Loop in Honey Bee Dicistroviruses That Promotes Internal Ribosome Entry Site-Mediated Translation and Viral Infection.” *Journal of Virology* 92 (2): e01725-17. <https://doi.org/10.1128/jvi.01725-17>.
- Au, Hilda H.T., and Eric Jan. 2012. “Insights into Factorless Translational Initiation by the tRNA-Like Pseudoknot Domain of a Viral IRES.” *PLoS ONE* 7 (12). <https://doi.org/10.1371/journal.pone.0051477>.

- Author, Gene, Kourosh Salehi-Ashtiani, Andrej Lupták, Alexander Litovchick, and Jack W Szostak. 2006. “A Genomewide Search for Ribozymes Reveals an HDV-like Sequence in the Human CPEB3.” *New Series*. Vol. 313. <https://www-jstor-org.ezproxy.library.ubc.ca/stable/pdf/20031364.pdf?refreqid=excelsior%3A7f3186b9a49367b32eafcf10a685010a>.
- Bakowska-Zywicka, Kamilla, Marta Kasprzyk, and Tomasz Twardowski. 2016. “TRNA-derived short RNAs bind to *Saccharomyces cerevisiae* ribosomes in a stress-dependent manner and inhibit protein synthesis in vitro.” *FEMS Yeast Research* 16 (6). <https://doi.org/10.1093/femsyr/fow077>.
- Baltimore, D. 1971. “Expression of animal virus genomes.” *Bacteriological Reviews* 35 (3): 235–41. <https://doi.org/10.1128/membr.35.3.235-241.1971>.
- Ban, Nenad, Roland Beckmann, Jamie H.D. Cate, Jonathan D. Dinman, François Dragon, Steven R. Ellis, Denis L.J. Lafontaine, et al. 2014. “A new system for naming ribosomal proteins.” *Current Opinion in Structural Biology* 24 (1): 165–69. <https://doi.org/10.1016/j.sbi.2014.01.002>.
- Barbosa, Cristina Carvalho, Sydnee H. Calhoun, and Hans Joachim Wieden. 2020. “Non-coding RNAs: What are we missing?” *Biochemistry and Cell Biology*. <https://doi.org/10.1139/bcb-2019-0037>.
- Belov, G. A., V. Nair, B. T. Hansen, F. H. Hoyt, E. R. Fischer, and E. Ehrenfeld. 2012. “Complex Dynamic Development of Poliovirus Membranous Replication Complexes.” *Journal of Virology* 86 (1): 302–12. <https://doi.org/10.1128/jvi.05937-11>.
- Belsham, G J, and J K Brangwyn. 1990. “A region of the 5’ noncoding region of foot-and-mouth disease virus RNA directs efficient internal initiation of protein synthesis within cells:

- involvement with the role of L protease in translational control.” *Journal of Virology* 64 (11): 5389–95. <https://doi.org/10.1128/jvi.64.11.5389-5395.1990>.
- Berget, S. M., C. Moore, and P. A. Sharp. 1977. “Spliced segments at the 5’ terminus of adenovirus 2 late mRNA.” *Proceedings of the National Academy of Sciences of the United States of America* 74 (8): 3171–75. <https://doi.org/10.1073/pnas.74.8.3171>.
- Bienz, Kurt, Denise Egger, and Luis Pasamontes. 1987. “Association of polioviral proteins of the P2 genomic region with the viral replication complex and virus-induced membrane synthesis as visualized by electron microscopic immunocytochemistry and autoradiography.” *Virology* 160 (1): 220–26. [https://doi.org/10.1016/0042-6822\(87\)90063-8](https://doi.org/10.1016/0042-6822(87)90063-8).
- Bird, Sara W., and Karla Kirkegaard. 2015. “Escape of non-enveloped virus from intact cells.” *Virology* 479–480 (May): 444–49. <https://doi.org/10.1016/j.virol.2015.03.044>.
- Bonami, Jean Robert, Kenneth W. Hasson, Jocelyne Mari, Bonnie T. Poulos, and Donald V. Lightner. 1997. “Taura syndrome of marine penaeid shrimp: Characterization of the viral agent.” *Journal of General Virology* 78 (2): 313–19. <https://doi.org/10.1099/0022-1317-78-2-313>.
- Bonnal, Sophie, Céline Schaeffer, Laurent Créancier, Simone Clamens, Hervé Moine, Anne Catherine Prats, and Stéphane Vagner. 2003. “A single internal ribosome entry site containing a G quartet RNA structure drives fibroblast growth factor 2 gene expression at four alternative translation initiation codons.” *Journal of Biological Chemistry* 278 (41): 39330–36. <https://doi.org/10.1074/jbc.M305580200>.
- Borman, Andrew, and Richard J. Jackson. 1992. “Initiation of translation of human rhinovirus RNA: Mapping the internal ribosome entry site.” *Virology* 188 (2): 685–96.

[https://doi.org/10.1016/0042-6822\(92\)90523-R](https://doi.org/10.1016/0042-6822(92)90523-R).

- Brandenburg, Boerries, Lily Y. Lee, Melike Lakadamyali, Michael J. Rust, Xiaowei Zhuang, and James M. Hogle. 2007. "Imaging poliovirus entry in live cells." *PLoS Biology* 5 (7): 1543–55. <https://doi.org/10.1371/journal.pbio.0050183>.
- Bujarski, Jozef J., and Paul Kaesberg. 1986. "Genetic recombination between RNA components of a multipartite plant virus." *Nature* 321 (6069): 528–31. <https://doi.org/10.1038/321528a0>.
- Bushell, Martin, Mark Stoneley, Yi Wen Kong, Tiffany L. Hamilton, Keith A. Spriggs, Helen C. Dobbyn, Xiaoli Qin, Peter Sarnow, and Anne E. Willis. 2006. "Polypyrimidine Tract Binding Protein Regulates IRES-Mediated Gene Expression during Apoptosis." *Molecular Cell* 23 (3): 401–12. <https://doi.org/10.1016/j.molcel.2006.06.012>.
- Chen, Chang You, and Peter Sarnow. 1995. "Initiation of protein synthesis by the eukaryotic translational apparatus on circular RNAs." *Science* 268 (5209): 415–17. <https://doi.org/10.1126/science.7536344>.
- Chen, Ying Han, Wenli Du, Marne C. Hagemeijer, Peter M. Takvorian, Cyrilla Pau, Ann Cali, Christine A. Brantner, et al. 2015. "Phosphatidylserine vesicles enable efficient en bloc transmission of enteroviruses." *Cell* 160 (4): 619–30. <https://doi.org/10.1016/j.cell.2015.01.032>.
- Cherry, Sara, and Norbert Perrimon. 2004. "Entry is a rate-limiting step for viral infection in a *Drosophila melanogaster* model of pathogenesis." *Nature Immunology* 5 (1): 81–87. <https://doi.org/10.1038/ni1019>.
- Chiu, Shang Yi, Fabrice Lejeune, Aparna C. Ranganathan, and Lynne E. Maquat. 2004. "The pioneer translation initiation complex is functionally distinct from but structurally overlaps with the steady-state translation initiation complex." *Genes and Development* 18 (7): 745–

54. <https://doi.org/10.1101/gad.1170204>.

Cho, Michael W., Natalya Teterina, Denise Egger, Kurt Bienz, and Ellie Ehrenfeld. 1994.

“Membrane rearrangement and vesicle induction by recombinant poliovirus 2C and 2BC in human cells.” *Virology* 202 (1): 129–45. <https://doi.org/10.1006/viro.1994.1329>.

Chow, Louise T., James M. Roberts, James B. Lewis, and Thomas R. Broker. 1977. “A map of

cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids.” *Cell* 11 (4): 819–36. [https://doi.org/10.1016/0092-8674\(77\)90294-X](https://doi.org/10.1016/0092-8674(77)90294-X).

Chow, M., J. F.E. Newman, D. Filman, J. M. Hogle, D. J. Rowlands, and F. Brown. 1987.

“Myristylation of picornavirus capsid protein VP4 and its structural significance.” *Nature* 327 (6122): 482–86. <https://doi.org/10.1038/327482a0>.

Colussi, Timothy M., David A. Costantino, Jianyu Zhu, John Paul Donohue, Andrei A.

Korostelev, Zane A. Jaafar, Terra Dawn M. Plank, Harry F. Noller, and Jeffrey S. Kieft.

2015. “Initiation of translation in bacteria by a structured eukaryotic IRES RNA.” *Nature* 519 (7541): 110–13. <https://doi.org/10.1038/nature14219>.

Costantino, David A., Jennifer S. Pfingsten, Robert P. Rambo, and Jeffrey S. Kieft. 2008.

“tRNA-mRNA mimicry drives translation initiation from a viral IRES.” *Nature Structural and Molecular Biology* 15 (1): 57–64. <https://doi.org/10.1038/nsmb1351>.

Cox-Foster, Diana L., Sean Conlan, Edward C. Holmes, Gustavo Palacios, Jay D. Evans, Nancy

A. Moran, Phenix Lan Quan, et al. 2007. “A metagenomic survey of microbes in honey bee colony collapse disorder.” *Science* 318 (5848): 283–87.

<https://doi.org/10.1126/science.1146498>.

Culley, Alexander I., Andrew S. Lang, and Curtis A. Suttle. 2003. “High diversity of unknown

picorna-like viruses in the sea.” *Nature* 424 (6952): 1054–57.

<https://doi.org/10.1038/nature01886>.

Dhindwal, Sonali, Bryant Avila, Shanshan Feng, and Reza Khayat. 2019. “Porcine Circovirus 2 Uses a Multitude of Weak Binding Sites To Interact with Heparan Sulfate, and the Interactions Do Not Follow the Symmetry of the Capsid.” *Journal of Virology* 93 (6).
<https://doi.org/10.1128/jvi.02222-18>.

Domier, Leslie L., Nancy K. McCoppin, and Cleora J. D’arcy. 2000. “Sequence requirements for translation initiation of *Rhopalosiphum padi* virus ORF2.” *Virology* 268 (2): 264–71.
<https://doi.org/10.1006/viro.2000.0189>.

Edery, I., and N. Sonenberg. 1985. “Cap-dependent RNA splicing in a HeLa nuclear extract.” *Proceedings of the National Academy of Sciences of the United States of America* 82 (22): 7590–94. <https://doi.org/10.1073/pnas.82.22.7590>.

Etchison, Diane, Susan C Milburn, I. Edery, N. Sonenberg, and J. W. Hershey. 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.” *Journal of Biological Chemistry* 257 (24): 14806–10.
[https://doi.org/10.1016/s0021-9258\(18\)33352-0](https://doi.org/10.1016/s0021-9258(18)33352-0).

Fahlman, Richard P., and Olke C. Uhlenbeck. 2004. “Contribution of the esterified amino acid to the binding of aminoacylated tRNAs to the ribosomal P- and A-sites.” *Biochemistry* 43 (23): 7575–83. <https://doi.org/10.1021/bi0495836>.

Fernández, Israel S, Xiao-Chen Bai, Garib Murshudov, Sjors H W Scheres, and V Ramakrishnan. 2014. “Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome.” *Cell* 157 (4): 823–31.

<https://doi.org/10.1016/j.cell.2014.04.015>.

Fernandez, James, Ibrahim Yaman, Peter Sarnow, Martin D. Snider, and Maria Hatzoglou. 2002.

“Regulation of internal ribosomal entry site-mediated translation by phosphorylation of the translation initiation factor eIF2 α .” *Journal of Biological Chemistry* 277 (21): 19198–205.

<https://doi.org/10.1074/jbc.M201052200>.

Firth, Andrew E., and Ian Brierley. 2012. “Non-canonical translation in RNA viruses.” *Journal*

of General Virology 93 (PART 7): 1385–1409. <https://doi.org/10.1099/vir.0.042499-0>.

Fresco, Lucille D., and Stephen Buratowski. 1996. “Conditional mutants of the yeast mRNA

capping enzyme show that the cap enhances, but is not required for, mRNA splicing.” *RNA* 2 (6): 584–96.

Fricker, Roger, Rebecca Brogli, Hannes Luidalepp, Leander Wyss, Michel Fasnacht, Oliver Joss,

Marek Zywicki, et al. 2019. “A tRNA half modulates translation as stress response in *Trypanosoma brucei*.” *Nature Communications* 10 (1): 118. <https://doi.org/10.1038/s41467-018-07949-6>.

Fricks, C E, and J M Hogle. 1990. “Cell-induced conformational change in poliovirus:

externalization of the amino terminus of VP1 is responsible for liposome binding.” *Journal of Virology* 64 (5): 1934–45. <https://doi.org/10.1128/jvi.64.5.1934-1945.1990>.

Garibaldi, Angela, Francisco Carranza, and Klemens J. Hertel. 2017. “Isolation of newly

transcribed rna using the metabolic label 4-thiouridine.” In *Methods in Molecular Biology*, 1648:169–76. Humana Press Inc. https://doi.org/10.1007/978-1-4939-7204-3_13.

Garrey, Julianne L., Yun-Young Lee, Hilda H. T. Au, Martin Bushell, and Eric Jan. 2010. “Host

and viral translational mechanisms during cricket paralysis virus infection.” *Journal of Virology* 84 (2): 1124–38. <https://doi.org/10.1128/JVI.02006-09>.

- Gebetsberger, Jennifer, Leander Wyss, Anna M. Mleczko, Julia Reuther, and Norbert Polacek. 2017. "A tRNA-derived fragment competes with mRNA for ribosome binding and regulates translation during stress." *RNA Biology* 14 (10): 1364–73.
<https://doi.org/10.1080/15476286.2016.1257470>.
- Gebetsberger, Jennifer, Marek Zywicki, Andrea Künzi, and Norbert Polacek. 2012. "tRNA-derived fragments target the ribosome and function as regulatory non-coding RNA in *Haloferax volcanii*." *Archaea* 2012. <https://doi.org/10.1155/2012/260909>.
- Gerlitz, Gabi, Rosemary Jagus, and Orna Elroy-Stein. 2002. "Phosphorylation of initiation factor-2 α is required for activation of internal translation initiation during cell differentiation." *European Journal of Biochemistry* 269 (11): 2810–19.
<https://doi.org/10.1046/j.1432-1033.2002.02974.x>.
- Glass, Michael J., Xi Yu Jia, and Donald F. Summers. 1993. "Identification of the hepatitis a virus internal ribosome entry site: In vivo and in vitro analysis of bicistronic mRNAs containing the hcv 5' noncoding region." *Virology* 193 (2): 842–52.
<https://doi.org/10.1006/viro.1993.1193>.
- Goic, Bertsy, Nicolas Vodovar, Juan A Mondotte, Clément Monot, Lionel Frangeul, Hervé Blanc, Valérie Gausson, Jorge Vera-Otarola, Gael Cristofari, and Maria-Carla Saleh. 2013. "RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila*." *Nature Immunology* 14 (4): 396–403.
<https://doi.org/10.1038/ni.2542>.
- Goodfellow, Ian. 2011. "The genome-linked protein VPg of vertebrate viruses - A multifaceted protein." *Current Opinion in Virology* 1 (5): 355–62.
<https://doi.org/10.1016/j.coviro.2011.09.003>.

- Gradi, Alessandra, Yuri V. Svitkin, Hiroaki Imataka, and Nahum Sonenberg. 1998. "Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not EIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection." *Proceedings of the National Academy of Sciences of the United States of America* 95 (19): 11089–94. <https://doi.org/10.1073/pnas.95.19.11089>.
- Gribble, Jennifer, Andrea J Pruijssers, Maria L Agostini, Jordan Anderson-Daniels, James D Chappell, Xiaotao Lu, Laura J Stevens, Andrew Laurence Routh, and Mark R Denison. 2021. "The coronavirus proofreading exoribonuclease mediates extensive viral recombination." *PLoS Pathogens* 17 (1): e1009226. <https://doi.org/10.1101/2020.04.23.057786>.
- Grover, R., M. M. Candeias, R. Fhraeus, and S. Das. 2009. "P53 and little brother p53/47: Linking IRES activities with protein functions." *Oncogene* 28 (30): 2766–72. <https://doi.org/10.1038/onc.2009.138>.
- Hashem, Yaser, Amedee Des Georges, Vidya Dhote, Robert Langlois, Hstau Y. Liao, Robert A. Grassucci, Tatyana V. Pestova, Christopher U.T. Hellen, and Joachim Frank. 2013. "Hepatitis-C-virus-like internal ribosome entry sites displace eIF3 to gain access to the 40S subunit." *Nature* 503 (7477): 539–43. <https://doi.org/10.1038/nature12658>.
- Hatakeyama, Yoshinori, Norihiro Shibuya, Takashi Nishiyama, and Nobuhiko Nakashima. 2004. "Structural variant of the intergenic internal ribosome entry site elements in dicistroviruses and computational search for their counterparts." *RNA (New York, N.Y.)* 10 (5): 779–86. <https://doi.org/10.1261/rna.5208104>.
- Hellen, C. U.T., and P. Sarnow. 2001. "Internal ribosome entry sites in eukaryotic mRNA molecules." *Genes and Development* 15 (13): 1593–1612.

<https://doi.org/10.1101/gad.891101>.

- Hertz, Marla I., and Sunnie R. Thompson. 2011. “In vivo functional analysis of the Dicistroviridae intergenic region internal ribosome entry sites.” *Nucleic Acids Research* 39 (16): 7276–88. <https://doi.org/10.1093/nar/gkr427>.
- Higgs, Paul G., and Niles Lehman. 2015. “The RNA World: Molecular cooperation at the origins of life.” *Nature Reviews Genetics* 16 (1): 7–17. <https://doi.org/10.1038/nrg3841>.
- Himeno, Hyouta, Daisuke Kurita, and Akira Muto. 2014. “TmRNA-mediated trans-translation as the major ribosome rescue system in a bacterial cell.” *Frontiers in Genetics* 5: 66. <https://doi.org/10.3389/fgene.2014.00066>.
- Hinnebusch, Alan G., and Jon R. Lorsch. 2012. “The mechanism of eukaryotic translation initiation: New insights and challenges.” *Cold Spring Harbor Perspectives in Biology* 4 (10): a011544. <https://doi.org/10.1101/cshperspect.a011544>.
- Horie, Masayuki, Tomoyuki Honda, Yoshiyuki Suzuki, Yuki Kobayashi, Takuji Daito, Tatsuo Oshida, Kazuyoshi Ikuta, et al. 2010. “Endogenous non-retroviral RNA virus elements in mammalian genomes.” *Nature* 463 (7277): 84–87. <https://doi.org/10.1038/nature08695>.
- Horie, Masayuki, Yuki Kobayashi, Yoshiyuki Suzuki, and Keizo Tomonaga. 2013. “Comprehensive analysis of endogenous bornavirus-like elements in eukaryote genomes.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 368 (1626): 20120499. <https://doi.org/10.1098/rstb.2012.0499>.
- Ingolia, Nicholas T., Gloria A. Brar, Silvia Rouskin, Anna M. McGeachy, and Jonathan S. Weissman. 2012. “The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments.” *Nature Protocols* 7 (8): 1534–50. <https://doi.org/10.1038/nprot.2012.086>.

- Isken, Olaf, Yoon Ki Kim, Nao Hosoda, Greg L. Mayeur, John W.B. Hershey, and Lynne E. Maquat. 2008. "Upf1 Phosphorylation Triggers Translational Repression during Nonsense-Mediated mRNA Decay." *Cell* 133 (2): 314–27. <https://doi.org/10.1016/j.cell.2008.02.030>.
- Jaafar, Zane A., and Jeffrey S. Kieft. 2018. "Viral RNA structure-based strategies to manipulate translation." *Nature Reviews Microbiology* 17 (December): 110–23. <https://doi.org/10.1038/s41579-018-0117-x>.
- Jaafar, Zane A., Akihiro Oguro, Yoshikazu Nakamura, and Jeffrey S. Kieft. 2016. "Translation initiation by the hepatitis C virus IRES requires eIF1A and ribosomal complex remodeling." *ELife* 5 (December): e21198. <https://doi.org/10.7554/eLife.21198>.
- Jackson, Richard J. 2013. "The current status of vertebrate cellular mRNA IRESs." *Cold Spring Harbor Perspectives in Biology* 5 (2). <https://doi.org/10.1101/cshperspect.a011569>.
- Jackson, Richard J., Christopher U.T. Hellen, and Tatyana V. Pestova. 2010. "The mechanism of eukaryotic translation initiation and principles of its regulation." *Nature Reviews Molecular Cell Biology* 11 (2): 113–27. <https://doi.org/10.1038/nrm2838>.
- Jacobs, Jonathan L., and Jonathan D. Dinman. 2004. "Systematic analysis of bicistronic reporter assay data." *Nucleic Acids Research* 32 (20): e160–e160. <https://doi.org/10.1093/nar/gnh157>.
- Jan, Eric, Terri Goss Kinzy, and Peter Sarnow. 2003. "Divergent tRNA-like element supports initiation, elongation, and termination of protein biosynthesis." *Proceedings of the National Academy of Sciences of the United States of America* 100 (26): 15410–15. <https://doi.org/10.1073/pnas.2535183100>.
- Jan, Eric, and Peter Sarnow. 2002. "Factorless Ribosome Assembly on the Internal Ribosome Entry Site of Cricket Paralysis Virus." *Journal of Molecular Biology* 324: 889–902.

[https://doi.org/10.1016/S0022-2836\(02\)01099-9](https://doi.org/10.1016/S0022-2836(02)01099-9).

Jang, Christopher J., and Eric Jan. 2010. "Modular domains of the Dicistroviridae intergenic internal ribosome entry site." *RNA (New York, N.Y.)* 16 (6): 1182–95.

<https://doi.org/10.1261/rna.2044610>.

Jang, Christopher J, Miranda C Y Lo, and Eric Jan*. 2008. "Conserved Element of the Dicistrovirus IGR IRES that Mimics an E-site tRNA/Ribosome Interaction Mediates Multiple Functions." *Journal of Molecular Biology* 387: 42–58.

<https://doi.org/10.1016/j.jmb.2009.01.042>.

Jang, S K, H G Kräusslich, M J Nicklin, G M Duke, A C Palmenberg, and E Wimmer. 1988. "A segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during in vitro translation." *Journal of Virology* 62 (8): 2636–2643.

<https://doi.org/10.1128/jvi.62.8.2636-2643.1988>.

Jesus, Nidia H. De. 2007. "Epidemics to eradication: The modern history of poliomyelitis."

Virology Journal 4: 70. <https://doi.org/10.1186/1743-422X-4-70>.

Joachims, Michelle, Pieter C. Van Breugel, and Richard E. Lloyd. 1999. "Cleavage of Poly(A)-Binding Protein by Enterovirus Proteases Concurrent with Inhibition of Translation In Vitro." *Journal of Virology* 73 (1): 718–27. <https://doi.org/10.1128/jvi.73.1.718-727.1999>.

Johannes, G, and P Sarnow. 1998. "Cap-independent polysomal association of natural mRNAs encoding c-myc, BiP, and eIF4G conferred by internal ribosome entry sites." *RNA (New York, N.Y.)* 4 (12): 1500–1513. <http://www.ncbi.nlm.nih.gov/pubmed/9848649>.

Jopling, Catherine L., Keith A. Spriggs, Sally A. Mitchell, Mark Stoneley, and Anne E. Willis. 2004. "L-Myc protein synthesis is initiated by internal ribosome entry." *RNA* 10 (2): 287–98. <https://doi.org/10.1261/rna.5138804>.

- Kapp, Lee D., and Jon R. Lorsch. 2004. "GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2." *Journal of Molecular Biology* 335 (4): 923–36. <https://doi.org/10.1016/j.jmb.2003.11.025>.
- Kerr, Craig H., Qing S. Wang, Kathleen Keatings, Anthony Khong, Douglas Allan, Calvin K. Yip, Leonard J. Foster, and Eric Jan. 2015. "The 5' Untranslated Region of a Novel Infectious Molecular Clone of the Dicistrovirus Cricket Paralysis Virus Modulates Infection." *Journal of Virology* 89 (11): 5919–34. <https://doi.org/10.1128/jvi.00463-15>.
- Kerr, Craig H., Qing S. Wang, Kyung Mee Moon, Kathleen Keatings, Douglas W. Allan, Leonard J. Foster, and Eric Jan. 2018. "IRES-dependent ribosome repositioning directs translation of a +1 overlapping ORF that enhances viral infection." *Nucleic Acids Research* 46 (22): 11952–67. <https://doi.org/10.1093/nar/gky1121>.
- Kerr, Craig H., and Eric Jan. 2016. "Commandeering the Ribosome: Lessons Learned from Dicistroviruses about Translation." *Journal of Virology* 90 (12): 5538–40. <https://doi.org/10.1128/JVI.00737-15>.
- Kerr, Craig H., Zi Wang Ma, Christopher J Jang, Sunnie R Thompson, and Eric Jan. 2016. "Molecular analysis of the factorless internal ribosome entry site in Cricket Paralysis virus infection." *Scientific Reports* 6: 37319. <https://doi.org/10.1038/srep37319>.
- Khatchikian, Detlev, Michaela Orlich, and Rudolf Rott. 1989. "Increased viral pathogenicity after insertion of a 28S ribosomal RNA sequence into the haemagglutinin gene of an influenza virus." *Nature* 340 (6229): 156–57. <https://doi.org/10.1038/340156a0>.
- Khong, Anthony, Jennifer M. Bonderoff, Ruth V. Spriggs, Erik Tammperre, Craig H. Kerr, Thomas J. Jackson, Anne E. Willis, and Eric Jan. 2016. "Temporal regulation of distinct internal ribosome entry sites of the dicistroviridae cricket paralysis virus." *Viruses* 8 (1).

<https://doi.org/10.3390/v8010025>.

Kim, Dongwan, Joo Yeon Lee, Jeong Sun Yang, Jun Won Kim, V Narry Kim, and Hyeshik Chang. 2020. “The architecture of SARS-CoV-2 transcriptome.” *Cell* 181 (4): 914–21. <https://doi.org/10.1101/2020.03.12.988865>.

Kim, Joon Hyun, Sung Mi Park, Ji Hoon Park, Sun Ju Keum, and Sung Key Jang. 2011. “EIF2A mediates translation of hepatitis C viral mRNA under stress conditions.” *EMBO Journal* 30 (12): 2454–64. <https://doi.org/10.1038/emboj.2011.146>.

Koh, Cha San, Axel F. Brilot, Nikolaus Grigorieff, and Andrei A. Korostelev. 2014. “Taura syndrome virus IRES initiates translation by binding its tRNA-mRNA-like structural element in the ribosomal decoding center.” *PNAS* 111 (25): 9139–44. <https://doi.org/10.1073/pnas.1406335111>.

Komar, Anton A., and Maria Hatzoglou. 2011. “Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states.” *Cell Cycle* 10 (2): 229–40. <https://doi.org/10.4161/cc.10.2.14472>.

Komar, Anton A., Thierry Lesnik, Christophe Cullin, William C. Merrick, Hans Trachsel, and Michael Altmann. 2003. “Internal initiation drives the synthesis of Ure2 protein lacking the prion domain and affects [URE3] propagation in yeast cells.” *The EMBO Journal* 22 (5): 1199–1209. <https://doi.org/10.1093/emboj/cdg103>.

Konarska, Maria M., Richard A. Padgett, and Phillip A. Sharp. 1984. “Recognition of cap structure in splicing in vitro of mRNA precursors.” *Cell* 38 (3): 731–36. [https://doi.org/10.1016/0092-8674\(84\)90268-X](https://doi.org/10.1016/0092-8674(84)90268-X).

Kozak, Marilyn. 1986. “Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes.” *Cell* 44 (2): 283–92.

- [https://doi.org/10.1016/0092-8674\(86\)90762-2](https://doi.org/10.1016/0092-8674(86)90762-2).
- — — . 1991. “Structural features in eukaryotic mRNAs that modulate the initiation of translation.” *Journal of Biological Chemistry* 266 (30): 19867–70.
- — — . 2003. “Alternative ways to think about mRNA sequences and proteins that appear to promote internal initiation of translation.” *Gene* 318 (1–2): 1–23.
- [https://doi.org/10.1016/S0378-1119\(03\)00774-1](https://doi.org/10.1016/S0378-1119(03)00774-1).
- — — . 2005. “A second look at cellular mRNA sequences said to function as internal ribosome entry sites.” *Nucleic Acids Research* 33 (20): 6593–6602.
- <https://doi.org/10.1093/nar/gki958>.
- Kräusslich, H G, M J Nicklin, H Toyoda, D Etchison, and E Wimmer. 1987. “Poliovirus proteinase 2A induces cleavage of eucaryotic initiation factor 4F polypeptide p220.” *Journal of Virology* 61 (9): 2711.
- Kuhle, Bernhard, and Ralf Ficner. 2014. “Structural insight into the recognition of aminoacylated initiator tRNA by eIF5B in the 80S initiation complex.” *BMC Structural Biology* 14 (1). <https://doi.org/10.1186/s12900-014-0020-2>.
- Kühn, R, N Luz, and E Beck. 1990. “Functional analysis of the internal translation initiation site of foot-and-mouth disease virus.” *Journal of Virology* 64 (10): 4625.
- Kuyumcu-Martinez, N. Muge, Marc E. Van Eden, Patrick Younan, and Richard E. Lloyd. 2004. “Cleavage of Poly(A)-Binding Protein by Poliovirus 3C Protease Inhibits Host Cell Translation: a Novel Mechanism for Host Translation Shutoff.” *Molecular and Cellular Biology* 24 (4): 1779–90. <https://doi.org/10.1128/mcb.24.4.1779-1790.2004>.
- Landry, Dori M., Marla I. Hertz, and Sunnie R. Thompson. 2009. “RPS25 is essential for translation initiation by the Dicistroviridae and hepatitis C viral IRESs.” *Genes and*

- Development* 23 (23): 2753–64. <https://doi.org/10.1101/gad.1832209>.
- Lauring, Adam S., and Raul Andino. 2010. “Quasispecies Theory and the Behavior of RNA Viruses.” Edited by Marianne Manchester. *PLoS Pathogens* 6 (7): e1001005. <https://doi.org/10.1371/journal.ppat.1001005>.
- Lee, Y. F., A. Nomoto, B. M. Detjen, and E. Wimmer. 1977. “A protein covalently linked to poliovirus genome RNA.” *PNAS* 74 (1): 59–63. <https://doi.org/10.1073/pnas.74.1.59>.
- Lewis, Joe D., and Elisa Izaurralde. 1997. “The role of the cap structure in RNA processing and nuclear export.” *European Journal of Biochemistry* 247 (2): 461–69. <https://doi.org/10.1111/j.1432-1033.1997.00461.x>.
- Liu, Lin, Yinhu Li, Siliang Li, Ni Hu, Yimin He, Ray Pong, Danni Lin, Lihua Lu, and Maggie Law. 2012. “Comparison of next-generation sequencing systems.” *Journal of Biomedicine and Biotechnology*. <https://doi.org/10.1155/2012/251364>.
- Llácer, Jose L., Tanweer Hussain, Laura Marler, Colin Echeverría Aitken, Anil Thakur, Jon R. Lorsch, Alan G. Hinnebusch, and V. Ramakrishnan. 2015. “Conformational Differences between Open and Closed States of the Eukaryotic Translation Initiation Complex.” *Molecular Cell* 59 (3): 399–412. <https://doi.org/10.1016/j.molcel.2015.06.033>.
- Luke, Garry A., Pablo de Felipe, Alexander Lukashev, Susanna E. Kallioinen, Elizabeth A. Bruno, and Martin D. Ryan. 2008. “Occurrence, function and evolutionary origins of ‘2A-like’ sequences in virus genomes.” *Journal of General Virology* 89 (4): 1036–42. <https://doi.org/10.1099/vir.0.83428-0>.
- Maag, David, Christie A. Fekete, Zygmunt Gryczynski, and Jon R. Lorsch. 2005. “A conformational change in the eukaryotic translation preinitiation complex and release of eIF1 signal recognition of the start codon.” *Molecular Cell* 17 (2): 265–75.

<https://doi.org/10.1016/j.molcel.2004.11.051>.

Macejak, Dennis G., and Peter Sarnow. 1991. "Internal initiation of translation mediated by the 5' leader of a cellular mRNA." *Nature* 353 (6339): 90–94.

<https://doi.org/10.1038/353090a0>.

Mailliot, Justine, and Franck Martin. 2018. "Viral internal ribosomal entry sites: four classes for one goal." *Wiley Interdisciplinary Reviews: RNA* 9 (2): e1458.

<https://doi.org/10.1002/wrna.1458>.

Maori, Eyal, Edna Tanne, and Ilan Sela. 2007. "Reciprocal sequence exchange between non-retro viruses and hosts leading to the appearance of new host phenotypes." *Virology* 362 (2): 342–49. <https://doi.org/10.1016/j.virol.2006.11.038>.

Maquat, Lynne E., Woan Yuh Tarn, and Olaf Isken. 2010. "The pioneer round of translation: Features and functions." *Cell* 142 (3): 368–74. <https://doi.org/10.1016/j.cell.2010.07.022>.

Martineau, Yvan, Christine Le Bec, Laurent Monbrun, Valérie Allo, Ing-Ming Chiu, Olivier Danos, Hervé Moine, Hervé Prats, and Anne-Catherine Prats. 2004. "Internal Ribosome Entry Site Structural Motifs Conserved among Mammalian Fibroblast Growth Factor 1 Alternatively Spliced mRNAs." *Molecular and Cellular Biology* 24 (17): 7622–35.

<https://doi.org/10.1128/mcb.24.17.7622-7635.2004>.

Matsuda, Daiki, and Vincent P. Mauro. 2014. "Base pairing between hepatitis C virus RNA and 18S rRNA is required for IRES-dependent translation initiation in vivo." *PNAS* 111 (43): 15385–89. <https://doi.org/10.1073/pnas.1413472111>.

Mendelsohn, Cathy L., Eckard Wimmer, and Vincent R. Racaniello. 1989. "Cellular receptor for poliovirus: Molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily." *Cell* 56 (5): 855–65. <https://doi.org/10.1016/0092->

8674(89)90690-9.

Meurs, E F, Y Watanabe, S Kadereit, G N Barber, M G Katze, K Chong, B R Williams, and A G Hovanessian. 1992. "Constitutive expression of human double-stranded RNA-activated p68 kinase in murine cells mediates phosphorylation of eukaryotic initiation factor 2 and partial resistance to encephalomyocarditis virus growth." *Journal of Virology* 66 (10): 5805–14. <https://doi.org/10.1128/jvi.66.10.5805-5814.1992>.

Meyers, Gregor, Norbert Tautz, Edward J. Dubovi, and Heinz Jürgen Thiel. 1991. "Viral cytopathogenicity correlated with integration of ubiquitin-coding sequences." *Virology* 180 (2): 602–16. [https://doi.org/10.1016/0042-6822\(91\)90074-L](https://doi.org/10.1016/0042-6822(91)90074-L).

Mignone, Flavio, Giorgio Grillo, Flavio Licciulli, Michele Iacono, Sabino Liuni, Paul J Kersey, Jorge Duarte, Cecilia Saccone, and Graziano Pesole. 2005. "UTRdb and UTRsite: A collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs." *Nucleic Acids Research* 33 (December): D141–46. <https://doi.org/10.1093/nar/gki021>.

Minskaia, Ekaterina, Tobias Hertzog, Alexander E. Gorbalenya, Valérie Campanacci, Christian Cambillau, Bruno Canard, and John Ziebuhr. 2006. "Discovery of an RNA virus 3'→5' exoribonuclease that is critically involved in coronavirus RNA synthesis." *PNAS* 103 (13): 5108–13. <https://doi.org/10.1073/pnas.0508200103>.

Miranda, Joachim R. de, Guido Cordoni, and Giles Budge. 2010. "The Acute bee paralysis virus-Kashmir bee virus-Israeli acute paralysis virus complex." *Journal of Invertebrate Pathology* 103 (SUPPL. 1): S30–47. <https://doi.org/10.1016/j.jip.2009.06.014>.

Mitchell, Sally A., Keith A. Spriggs, Mark J. Coldwell, Richard J. Jackson, and Anne E. Willis. 2003. "The Apaf-1 internal ribosome entry segment attains the correct structural

- conformation for function via interactions with PTB and unr.” *Molecular Cell* 11 (3): 757–71. [https://doi.org/10.1016/S1097-2765\(03\)00093-5](https://doi.org/10.1016/S1097-2765(03)00093-5).
- Moore, Norman F., Anne Kearns, and Jim S. K. Pullin. 1980. “Characterization of Cricket Paralysis Virus-Induced Polypeptides in *Drosophila* Cells.” *Journal of Virology* 33 (1): 1–9. <https://doi.org/10.1128/jvi.33.1.1-9.1980>.
- Muhs, Margarita, Hiroshi Yamamoto, Jochen Ismer, Hiroaki Takaku, Masayuki Nashimoto, Toshio Uchiumi, Nobuhiko Nakashima, et al. 2011. “Structural basis for the binding of IRES RNAs to the head of the ribosomal 40S subunit.” *Nucleic Acids Research* 39 (12): 5264–75. <https://doi.org/10.1093/nar/gkr114>.
- Murray, Jason, Christos G Savva, Byung-Sik Shin, Thomas E Dever, V Ramakrishnan, and Israel S Fernández. 2016. “Structural characterization of ribosome recruitment and translocation by type IV IRES.” *ELife* 5. <https://doi.org/10.7554/eLife.13567>.
- Nakashima, Nobuhiko, and Toshio Uchiumi. 2009. “Functional analysis of structural motifs in dicistroviruses.” *Virus Research* 139 (2): 137–47. <https://doi.org/10.1016/j.virusres.2008.06.006>.
- Ng, Terry Fei Fan, Li Fang Chen, Yanchen Zhou, Beth Shapiro, Mathias Stiller, Peter D. Heintzman, Arvind Varsani, et al. 2014. “Preservation of viral genomes in 700-y-old caribou feces from a subarctic ice patch.” *PNAS* 111 (47): 16842–47. <https://doi.org/10.1073/pnas.1410429111>.
- Nilsen, Timothy W. 2013. “Toeprinting.” *Cold Spring Harbor Protocols* 2013 (9): 896–99. <https://doi.org/10.1101/pdb.prot077180>.
- Nishiyama, Takashi, Hiroshi Yamamoto, Toshio Uchiumi, and Nobuhiko Nakashima. 2007. “Eukaryotic ribosomal protein RPS25 interacts with the conserved loop region in a

- dicistroviral intergenic internal ribosome entry site.” *Nucleic Acids Research* 35 (5): 1514–21. <https://doi.org/10.1093/nar/gkl1121>.
- Nomoto, A., Y. Fon Lee, and E. Wimmer. 1976. “The 5’ end of poliovirus mRNA is not capped with m⁷G(5’)ppp(5’)Np.” *PNAS* 73 (2): 375–80. <https://doi.org/10.1073/pnas.73.2.375>.
- Oh, S. K., M. P. Scott, and P. Sarnow. 1992. “Homeotic gene Antennapedia mRNA contains 5’-noncoding sequences that confer translational initiation by internal ribosome binding.” *Genes and Development* 6 (9): 1643–53. <https://doi.org/10.1101/gad.6.9.1643>.
- Ohlmann, T., M. Rau, V. M. Pain, and S. J. Morley. 1996. “The C-terminal domain of eukaryotic protein synthesis initiation factor (eIF) 4G is sufficient to support cap-independent translation in the absence of eIF4E.” *The EMBO Journal* 15 (6): 1371–82. <https://doi.org/10.1002/j.1460-2075.1996.tb00479.x>.
- Passmore, Lori A., T. Martin Schmeing, David Maag, Drew J. Applefield, Michael G. Acker, Mikkel A A. Algire, Jon R. Lorsch, and V. Ramakrishnan. 2007. “The Eukaryotic Translation Initiation Factors eIF1 and eIF1A Induce an Open Conformation of the 40S Ribosome.” *Molecular Cell* 26 (1): 41–50. <https://doi.org/10.1016/j.molcel.2007.03.018>.
- Paulin, Fiona E.M., Linda E. Campbell, Kirsty O’Brien, Jane Loughlin, and Christopher G. Proud. 2001. “Eukaryotic translation initiation factor 5 (eIF5) acts as a classical GTPase-activator protein.” *Current Biology* 11 (1): 55–59. [https://doi.org/10.1016/S0960-9822\(00\)00025-7](https://doi.org/10.1016/S0960-9822(00)00025-7).
- Pelletier, Jerry, and Nahum Sonenberg. 1988. “Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA.” *Nature* 334 (6180): 320–25. <https://doi.org/10.1038/334320a0>.
- Pestova, Tatyana V., and Christopher U.T. Hellen. 2003. “Translation elongation after assembly

of ribosomes on the Cricket paralysis virus internal ribosomal entry site without initiation factors or initiator tRNA.” *Genes and Development* 17 (2): 181–86.

<https://doi.org/10.1101/gad.1040803>.

— — —. 2005. “Reconstitution of eukaryotic translation elongation in vitro following initiation by internal ribosomal entry.” *Methods* 36 (3): 261–69.

<https://doi.org/10.1016/j.ymeth.2005.04.004>.

Pestova, Tatyana V., Ivan B. Lomakin, Joon H. Lee, Sang Ki Choi, Thomas E. Dever, and Christopher U.T. Hellen. 2000. “The joining of ribosomal subunits in eukaryotes requires eIF5B.” *Nature* 403 (6767): 332–35. <https://doi.org/10.1038/35002118>.

Pestova, Tatyana V., Ivan N. Shatsky, Simon P. Fletcher, Richard J. Jackson, and Christopher U.T. Hellen. 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 and Development* 12 (1): 67–83.

<https://doi.org/10.1101/gad.12.1.67>.

Petersen, Christian P., Marie Eve Bordeleau, Jerry Pelletier, and Phillip A. Sharp. 2006. “Short RNAs repress translation after initiation in mammalian cells.” *Molecular Cell* 21 (4): 533–42. <https://doi.org/10.1016/j.molcel.2006.01.031>.

Peyambari, Mahtab, Sylvia Warner, Nicholas Stoler, Drew Rainer, Marilyn J. Roossinck, and J Roossinck. 2019. “A 1,000-Year-Old RNA Virus.” *Journal of Virology* 93 (October): e01188-18. <https://doi.org/10.1128/jvi.01188-18>.

Pfingsten, Jennifer S., David A. Costantino, and Jeffrey S. Kieft. 2006. “Structural basis for ribosome recruitment and manipulation by a viral IRES RNA.” *Science* 314 (5804): 1450–54. <https://doi.org/10.1126/science.1133281>.

- Pinkstaff, Jason K., Stephen A. Chappell, Vincent P. Mauro, Gerald M. Edelman, and Les A. Krushel. 2001. "Internal initiation of translation of five dendritically localized neuronal mRNAs." *PNAS* 98 (5): 2770–75. <https://doi.org/10.1073/pnas.051623398>.
- Pircher, Andreas, Kamilla Bakowska-Zywicka, Lukas Schneider, Marek Zywicki, and Norbert Polacek. 2014. "An mRNA-Derived Noncoding RNA Targets and Regulates the Ribosome." *Molecular Cell* 54 (1): 147. <https://doi.org/10.1016/J.MOLCEL.2014.02.024>.
- Pircher, Andreas, Jennifer Gebetsberger, and Norbert Polacek. 2014. "Ribosome-associated ncRNAs: An emerging class of translation regulators." *RNA Biology* 11 (11): 1335–39. <https://doi.org/10.1080/15476286.2014.996459>.
- Pisareva, Vera P, Andrey V Pisarev, and Israel S Fernández. 2018. "Dual tRNA mimicry in the Cricket Paralysis Virus IRES uncovers an unexpected similarity with the Hepatitis C Virus IRES." *ELife* 7 (June): e34062. <https://doi.org/10.7554/eLife.34062>.
- Plank, Terra Dawn M., and Jeffrey S. Kieft. 2012. "The structures of nonprotein-coding RNAs that drive internal ribosome entry site function." *Wiley Interdisciplinary Reviews: RNA* 3 (2): 195–212. <https://doi.org/10.1002/wrna.1105>.
- Qin, Xiaoli, and Peter Sarnow. 2004. "Preferential Translation of Internal Ribosome Entry Site-containing mRNAs during the Mitotic Cycle in Mammalian Cells." *Journal of Biological Chemistry* 279 (14): 13721–28. <https://doi.org/10.1074/jbc.M312854200>.
- Quesne, John P.C. Le, Mark Stoneley, Graham A. Fraser, and Anne E. Willis. 2001. "Derivation of a structural model for the c-myc IRES." *Journal of Molecular Biology* 310 (1): 111–26. <https://doi.org/10.1006/JMBI.2001.4745>.
- Ren, Qian, Qing S. Wang, Andrew E. Firth, Mandy M.Y. Chan, Joost W. Gouw, M. Marta Guarna, Leonard J. Foster, John F. Atkins, and Eric Jan. 2012. "Alternative reading frame

- selection mediated by a tRNA-like domain of an internal ribosome entry site.” *PNAS* 109 (11): e630-639. <https://doi.org/10.1073/pnas.1111303109>.
- Riley, Alura, Jordan Lindsay, E, and Martin Holcik. 2010. “Distinct 5’ UTRs Regulate XIAP Expression Under Normal Growth Conditions and During Cellular Stress.” *Nucleic Acids Research* 38 (14). <https://doi.org/10.1093/NAR/GKQ241>.
- Roos, W. H., I. L. Ivanovska, A. Evilevitch, and G. J.L. Wuite. 2007. “Viral capsids: Mechanical characteristics, genome packaging and delivery mechanisms.” *Cellular and Molecular Life Sciences*. Springer. <https://doi.org/10.1007/s00018-007-6451-1>.
- Ruehle, Marisa D, Haibo Zhang, Ryan M Sheridan, Somdeb Mitra, Yuanwei Chen, Ruben L Gonzalez, Barry S Cooperman, and Jeffrey S Kieft. 2015. “A dynamic RNA loop in an IRES affects multiple steps of elongation factor-mediated translation initiation.” *ELife* 4 (November). <https://doi.org/10.7554/elife.08146>.
- Sanjuán, Rafael, Miguel R. Nebot, Nicola Chirico, Louis M. Mansky, and Robert Belshaw. 2010. “Viral Mutation Rates.” *Journal of Virology* 84 (19): 9733–48. <https://doi.org/10.1128/jvi.00694-10>.
- Sarnow, P. 1989. “Translation of glucose-regulated protein 78/immunoglobulin heavy-chain binding protein mRNA is increased in poliovirus-infected cells at a time when cap-dependent translation of cellular mRNAs is inhibited.” *PNAS* 86 (15): 5795–99. <https://doi.org/10.1073/PNAS.86.15.5795>.
- Sasaki, J., and N. Nakashima. 2000. “Methionine-independent initiation of translation in the capsid protein of an insect RNA virus.” *PNAS* 97 (4): 1512–15. <https://doi.org/10.1073/pnas.010426997>.
- Sasaki, J, and N Nakashima. 1999. “Translation initiation at the CUU codon is mediated by the

- internal ribosome entry site of an insect picorna-like virus in vitro.” *Journal of Virology* 73 (2): 1219–26. <http://www.ncbi.nlm.nih.gov/pubmed/9882324>.
- Sasaki, Jun, and Nobuhiko Nakashima. 2000. “Methionine-independent initiation of translation in the capsid protein of an insect RNA virus.” *PNAS* 97 (4): 1512–15. <https://doi.org/10.1073/pnas.010426997>.
- Schüler, Martin, Sean R. Connell, Aurelie Lescoute, Jan Giesebrecht, Marylena Dabrowski, Birgit Schroeer, Thorsten Mielke, Pawel A. Penczek, Eric Westhof, and Christian M.T. Spahn. 2006. “Structure of the ribosome-bound cricket paralysis virus IRES RNA.” *Nature Structural and Molecular Biology* 13 (12): 1092–96. <https://doi.org/10.1038/nsmb1177>.
- Sena, John de, and Benjamin Mandel. 1977. “Studies on the in vitro uncoating of poliovirus II. Characteristics of the membrane-modified particle.” *Virology* 78 (2): 554–66. [https://doi.org/10.1016/0042-6822\(77\)90130-1](https://doi.org/10.1016/0042-6822(77)90130-1).
- Shi, Mang, Xian Dan Lin, Jun Hua Tian, Liang Jun Chen, Xiao Chen, Ci Xiu Li, Xin Cheng Qin, et al. 2016. “Redefining the invertebrate RNA virosphere.” *Nature* 540 (7634): 539–43. <https://doi.org/10.1038/nature20167>.
- Shuman, Stewart. 2002. “What messenger RNA capping tells us about eukaryotic evolution.” *Nature Reviews Molecular Cell Biology* 3 (8): 619–25. <https://doi.org/10.1038/nrm880>.
- Simonetti, Angelita, Ewelina Guca, Anthony Bochler, Lauriane Kuhn, and Yaser Hashem. 2020. “Structural Insights into the Mammalian Late-Stage Initiation Complexes.” *Cell Reports* 31 (1): 107497. <https://doi.org/10.1016/j.celrep.2020.03.061>.
- Smith, Oliver, Alan Clapham, Pam Rose, Yuan Liu, Jun Wang, and Robin G. Allaby. 2014. “A complete ancient RNA genome: Identification, reconstruction and evolutionary history of archaeological Barley Stripe Mosaic Virus.” *Scientific Reports* 4 (1): 4003.

<https://doi.org/10.1038/srep04003>.

- Snijder, Eric J., Peter J. Bredenbeek, Jessika C. Dobbe, Volker Thiel, John Ziebuhr, Leo L.M. Poon, Yi Guan, Mikhail Rozanov, Willy J.M. Spaan, and Alexander E. Gorbalenya. 2003. “Unique and conserved features of genome and proteome of SARS-coronavirus, an early split-off from the coronavirus group 2 lineage.” *Journal of Molecular Biology* 331 (5): 991–1004. [https://doi.org/10.1016/S0022-2836\(03\)00865-9](https://doi.org/10.1016/S0022-2836(03)00865-9).
- Somers, Jason, Hang Ngoc Bao Luong, Judith Mitchell, Philip Batterham, and Trent Perry. 2017. “Pleiotropic effects of loss of the D α 1 subunit in *Drosophila melanogaster*: Implications for insecticide resistance.” *Genetics* 205 (1): 263–71. <https://doi.org/10.1534/genetics.116.195750>.
- Spahn, Christian M.T., Eric Jan, Anke Mulder, Robert A. Grassucci, Peter Sarnow, and Joachim Frank. 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 (4): 465–75. <https://doi.org/10.1016/j.cell.2004.08.001>.
- Spriggs, Keith A., Laura C. Cobbold, Catherine L. Jopling, Rebecca E. Cooper, Lindsay A. Wilson, Mark Stoneley, Mark J. Coldwell, et al. 2009. “Canonical Initiation Factor Requirements of the Myc Family of Internal Ribosome Entry Segments.” *Molecular and Cellular Biology* 29 (6): 1565–74. <https://doi.org/10.1128/mcb.01283-08>.
- Stedman, Kenneth M. 2015. “Deep Recombination: RNA and ssDNA Virus Genes in DNA Virus and Host Genomes.” *Annual Review of Virology* 2 (November): 203–17. <https://doi.org/10.1146/annurev-virology-100114-055127>.
- Stoneley, Mark, and Anne E. Willis. 2004. “Cellular internal ribosome entry segments: Structures, trans-acting factors and regulation of gene expression.” *Oncogene* 23 (18):

- 3200–3207. <https://doi.org/10.1038/sj.onc.1207551>.
- Subkhankulova, T., S. A. Mitchell, and A. E. Willis. 2001. “Internal ribosome entry segment-mediated initiation of c-Myc protein synthesis following genotoxic stress.” *Biochemical Journal* 359 (1): 183–92. <https://doi.org/10.1042/0264-6021:3590183>.
- Suhy, David A., Thomas H. Giddings, and Karla Kirkegaard. 2000. “Remodeling the Endoplasmic Reticulum by Poliovirus Infection and by Individual Viral Proteins: an Autophagy-Like Origin for Virus-Induced Vesicles.” *Journal of Virology* 74 (19): 8953–65. <https://doi.org/10.1128/jvi.74.19.8953-8965.2000>.
- Suttle, Curtis A. 2007. “Marine viruses - Major players in the global ecosystem.” *Nature Reviews Microbiology* 5 (10): 801–12. <https://doi.org/10.1038/nrmicro1750>.
- Sweeney, Trevor R., Irina S. Abaeva, Tatyana V. Pestova, and Christopher U.T. Hellen. 2014. “The mechanism of translation initiation on type 1 picornavirus IRESs.” *EMBO Journal* 33 (1): 76–92. <https://doi.org/10.1002/emboj.201386124>.
- Tate, John, Lars Liljas, Paul Scotti, Peter Christian, Lin Tianwei, and John E. Johnson. 1999. “The crystal structure of cricket paralysis virus: The first view of a new virus family.” *Nature Structural Biology* 6 (8): 765–74. <https://doi.org/10.1038/11543>.
- Taubenberger, Jeffery K., Ann H. Reid, Amy E. Krafft, Karen E. Bijwaard, and Thomas G. Fanning. 1997. “Initial genetic characterization of the 1918 ‘Spanish’ influenza virus.” *Science* 275 (5307): 1793–96. <https://doi.org/10.1126/science.275.5307.1793>.
- Taubenberger, Jeffery K., Ann H. Reid, Raina M. Lourens, Ruixue Wang, Guozhong Jin, and Thomas G. Fanning. 2005. “Characterization of the 1918 influenza virus polymerase genes.” *Nature* 437 (7060): 889–93. <https://doi.org/10.1038/nature04230>.
- Terenin, Ilya M., Sergey E. Dmitriev, Dmitry E. Andreev, and Ivan N. Shatsky. 2008.

- “Eukaryotic translation initiation machinery can operate in a bacterial-like mode without eIF2.” *Nature Structural and Molecular Biology* 15 (8): 836–41.
<https://doi.org/10.1038/nsmb.1445>.
- Thompson, Sunnie R. 2012. “So you want to know if your message has an IRES?” *Wiley Interdisciplinary Reviews: RNA* 3 (5): 697–705. <https://doi.org/10.1002/wrna.1129>.
- Thompson, Sunnie R., Keith D. Gulyas, and Peter Sarnow. 2001. “Internal initiation in *Saccharomyces cerevisiae* mediated by an initiator tRNA/eIF2-independent internal ribosome entry site element.” *PNAS* 98 (23): 12972–77.
<https://doi.org/10.1073/pnas.241286698>.
- Tosteson, M T, and M Chow. 1997. “Characterization of the ion channels formed by poliovirus in planar lipid membranes.” *Journal of Virology* 71 (1): 507–11.
<https://doi.org/10.1128/jvi.71.1.507-511.1997>.
- Valles, S. M., Y. Chen, A. E. Firth, D. M.A. Guérin, Y. Hashimoto, S. Herrero, J. R. de Miranda, and E. Ryabov. 2017. “ICTV virus taxonomy profile: Dicistroviridae.” *Journal of General Virology* 98 (3): 355–56. <https://doi.org/10.1099/jgv.0.000756>.
- vanEngelsdorp, Dennis, Jerry Hayes, Robyn M. Underwood, and Jeffery Pettis. 2008. “A Survey of Honey Bee Colony Losses in the U.S., Fall 2007 to Spring 2008.” Edited by Nick Gay. *PLoS ONE* 3 (12): e4071. <https://doi.org/10.1371/journal.pone.0004071>.
- Venkataraman, Sangita, Burra V.L.S. Prasad, and Ramasamy Selvarajan. 2018. “RNA dependent RNA polymerases: Insights from structure, function and evolution.” *Viruses* 10 (2): 76.
<https://doi.org/10.3390/v10020076>.
- Virgen-Slane, Richard, Janet M. Rozovics, Kerry D. Fitzgerald, Tuan Ngo, Wayne Chou, Gerbrand J. Van Der Heden Van Noort, Dmitri V. Filippov, Paul D. Gershon, and Bert L.

- Semler. 2012. "An RNA virus hijacks an incognito function of a DNA repair enzyme." *PNAS* 109 (36): 14634–39. <https://doi.org/10.1073/pnas.1208096109>.
- Vogt, Dorothee A., and Raul Andino. 2010. "An RNA element at the 5'-end of the poliovirus genome functions as a general promoter for RNA synthesis." *PLoS Pathogens* 6 (6). <https://doi.org/10.1371/journal.ppat.1000936>.
- Wang, Junhui, and Michael Gribskov. 2019. "IRESpy: An XGBoost model for prediction of internal ribosome entry sites." *BMC Bioinformatics* 20 (1): 409. <https://doi.org/10.1186/s12859-019-2999-7>.
- Wang, Qing S., Hilda H.T. Au, and Eric Jan. 2013. "Methods for studying IRES-mediated translation of positive-strand RNA viruses." *Methods* 59 (2): 167–79. <https://doi.org/10.1016/j.ymeth.2012.09.004>.
- Wang, Qing S., and Eric Jan. 2014. "Switch from Cap- to Factorless IRES-Dependent 0 and +1 Frame Translation during Cellular Stress and Dicistrovirus Infection." Edited by Thomas Preiss. *PLoS ONE* 9 (8): e103601. <https://doi.org/10.1371/journal.pone.0103601>.
- Wang, Xinying, Marli Vlok, Stephane Flibotte, and Eric Jan. 2021. "Resurrection of a Viral Internal Ribosome Entry Site from a 700 Year Old Ancient Northwest Territories Cripavirus." *Viruses* 13 (3): 493. <https://doi.org/10.3390/v13030493>.
- Warsaba, Reid, Jibin Sadasivan, and Eric Jan. 2020. "Dicistrovirus-Host Molecular Interactions." *Current Issues in Molecular Biology* 34: 83–112. <https://doi.org/10.21775/cimb.034.083>.
- White, J. P., L. C. Reineke, and R. E. Lloyd. 2011. "Poliovirus Switches to an eIF2-Independent Mode of Translation during Infection." *Journal of Virology* 85 (17): 8884–93. <https://doi.org/10.1128/jvi.00792-11>.
- Wilson, J E, T V Pestova, C U Hellen, and P Sarnow. 2000. "Initiation of protein synthesis from

the A site of the ribosome.” *Cell* 102 (4): 511–20.

<http://www.ncbi.nlm.nih.gov/pubmed/10966112>.

Wilson, J E, M J Powell, S E Hoover, and P Sarnow. 2000. “Naturally occurring dicistronic cricket paralysis virus RNA is regulated by two internal ribosome entry sites.” *Molecular and Cellular Biology* 20 (14): 4990–99. <http://www.ncbi.nlm.nih.gov/pubmed/10866656>.

Wilson, Joan E., Tatyana V. Pestova, Christopher U.T. Hellen, and Peter Sarnow. 2000.

“Initiation of protein synthesis from the A site of the ribosome.” *Cell* 102 (4): 511–20.

[https://doi.org/10.1016/S0092-8674\(00\)00055-6](https://doi.org/10.1016/S0092-8674(00)00055-6).

Wilson, Joan E., Marguerite J. Powell, Susan E. Hoover, and Peter Sarnow. 2000. “Naturally Occurring Dicistronic Cricket Paralysis Virus RNA Is Regulated by Two Internal Ribosome Entry Sites.” *Molecular and Cellular Biology* 20 (14): 4990–99.

<https://doi.org/10.1128/mcb.20.14.4990-4999.2000>.

Wolfe, Nathan D., Claire Panosian Dunavan, and Jared Diamond. 2007. “Origins of major human infectious diseases.” *Nature* 447 (7142): 279–83.

<https://doi.org/10.1038/nature05775>.

Xu, Hengyi, Jun Yao, Douglas C. Wu, and Alan M. Lambowitz. 2019. “Improved TGIRT-seq methods for comprehensive transcriptome profiling with decreased adapter dimer formation and bias correction.” *Scientific Reports* 9 (1): 1–17. <https://doi.org/10.1038/s41598-019-44457-z>.

Yamamoto, Hiroshi, Marianne Collier, Justus Loerke, Jochen Ismer, Andrea Schmidt, Tarek Hilal, Thiemo Sprink, et al. 2015. “Molecular architecture of the ribosome-bound Hepatitis C Virus internal ribosomal entry site .” *The EMBO Journal* 34 (24): 3042–58.

<https://doi.org/10.15252/emj.201592469>.

Yaman, Ibrahim, James Fernandez, Haiyan Liu, Mark Caprara, Anton A. Komar, Antonis E.

Koromilas, Lingyin Zhou, et al. 2003. "The zipper model of translational control: A small upstream ORF is the switch that controls structural remodeling of an mRNA leader." *Cell* 113 (4): 519–31. [https://doi.org/10.1016/S0092-8674\(03\)00345-3](https://doi.org/10.1016/S0092-8674(03)00345-3).

Yu, Yingpu, Assen Marintchev, Victoria G. Kolupaeva, Anett Unbehaun, Tatyana Veryasova, Shao Chiang Lai, Peng Hong, Gerhard Wagner, Christopher U.T. Hellen, and Tatyana V. Pestova. 2009. "Position of eukaryotic translation initiation factor eIF1A on the 40S ribosomal subunit mapped by directed hydroxyl radical probing." *Nucleic Acids Research* 37 (15): 5167–82. <https://doi.org/10.1093/nar/gkp519>.

Zhao, Chen, Fei Liu, and Anna Marie Pyle. 2018. "An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron." *RNA* 24 (2): 183–85. <https://doi.org/10.1261/rna.063479.117>.

Zuker, Michael. 2003. "Mfold web server for nucleic acid folding and hybridization prediction." *Nucleic Acids Research* 31 (13): 3406–15. <https://doi.org/10.1093/nar/gkg595>.