Conservation of Translation Initiation Between Prokaryotes and Eukaryotes

by

Nathalie Puchacz

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Science

in

Biology

Carleton University
Ottawa, Ontario

Abstract

While the fundamental goals of translation initiation are the same for all cells, it is the most phylogenetically diverse step within translation. Until recently, there has been no evidence of a molecular mechanism that can initiate translation in both prokaryotes and eukaryotes. Colussi et al. (2015) reported that the eukaryotic PSIV IGR IRES can initiate prokaryotic translation and effectively circumvent domain-specific divergences. This novel discovery led us to investigate whether this IRES is unique in its ability to initiate prokaryotic translation and attempt to elucidate any distinguishing characteristics that allow for its mechanism of action to occur. Our results indicate that this IRES is unique in its ability to initiate prokaryotic translation. It also appears that the structural integrity of this IRES is less critical in prokaryotes than eukaryotes, and complementary interactions between Domain III and the 16S rRNA may contribute to this IRES' mechanism of action.

Acknowledgments

So many people deserve recognition for their contribution to this thesis. Above all, to Dr. Ashkan Golshani, thank you. Not only for the opportunity to conduct this research, but for also accepting me into your lab, for your constant support (both in the lab and life) and for furthering me as a researcher.

To my fellow Golshani lab members, past and present, thank you for your support, your advice and your acceptance. Thank you for becoming my family.

To my friends and family, thank you for your support, patience and understanding when life got tough.

And finally, to James. Words can't even begin to describe how grateful I am for your presence in my life. Thank you for your support through failed experiments, the struggles of writing, the pressures of life and the-day-to-day mood swings. I don't think I could have done this without you.

Table of Contents

List of Tables

List of Illustrations

List of Abbreviations

A – adenine
ANOVA – analysis of variance
APAF-1 – apoptotic protease activating factor 1
A-Site – aminoacyl/acceptor site
BCL-2 – B-cell lymphoma 2
BVDV – *Bovine viral diarrhea virus*
C – cytosine
CaMV – *Cauliflower mosaic virus*
c-myc – c-myelocytoma
CrPV – *Cricket paralysis virus*
CSFV – *Classical swine fever virus*
DNA – deoxyribonucleic acid
eIFs – eukaryotic initiation factors
EMCV – *Encephalomyocarditis virus*
eRF1 – eukaryotic release factor 1
E-Site – exit site
FLuc – Firefly luciferase
FMDV – *Foot-and-Mouth disease virus*
G – guanine
GDP – guanosine diphosphate
GTP – guanosine triphosphate
HCV – *Hepatitis C virus*
HIF-1α – Hypoxia-inducible factor 1-alpha
HPV – *Human papillomavirus*
IAPV – *Isreali acute paralysis virus*
IF – initiation factor
IGF1R – insulin-like growth factor 1
IGR – intergenic
Indel – insertion/deletion
IPTG– Isopropyl-ß-D-thiogalactopyranoside
IRES – internal ribosome entry site
ITAF – IRES trans acting factor
lacZ – β-galactosidase
LB – Luria Broth
Met – methionine
MFE – minimum free energy
mRNA – messenger ribonucleic acid
mRNP – messenger ribonucleoprotein
ORF – open reading frame

PABP – poly-A binding protein
PBS – phosphate buffered saline
PCPB-2 – poly(C)-binding protein 2
PCR – polymerase chain reaction
PIC – pre-initiation complex
PK – pseudoknot
PLB – passive lysis buffer
PMT – photomultiplier tube
Post-TCs – post-termination complexes
P-Site – peptidyl site
PSIV – *Plautia stali intestinal virus*
PTB – polypyrimidine-tract binding protein
PTC – peptidyl transferase center
PV – *Poliovirus*
RBS – ribosome binding site
RLU – relative light units
RLuc – Renilla luciferase
RNA – ribonucleic acid
rRNA – ribosomal ribonucleic acid
RTBV – *Rice tungro bacilliform virus*
RV – *Rhinovirus*
SD – Shine-Dalgarno
SL – stem-loop
sORF – short open reading frame
T – thymine
tRNA – transfer ribonucleic acid
U – uracil
uORF – upstream open reading frame
UTR – untranslated region

1. General Introduction

1.1. Translation

The translation of messenger RNA (mRNA) into protein is the final phase within the process of gene expression. Translation is a cyclical process that can be subdivided into four steps: initiation, elongation, termination and ribosome recycling.

Initiation marks the beginning of translation, whereby the ribosomal subunits are assembled at the initiation codon on the mRNA, with initiator transfer RNA (tRNA) bound in the ribosomal P-site and base paired with the start codon present on the mRNA (Allen & Frank, 2007). During elongation, the mRNA codons paired with their respective tRNA are repetitively decoded and peptide bonds are formed. Termination occurs once a stop codon is encountered and the newly completed polypeptide is released from the ribosome. Within the final stage, recycling, the ribosomal subunits are dissociated, releasing the mRNA and deacylated tRNAs, which will be reused for another round of initiation (Kapp and Lorsch, 2004).

1.1.1. Ribosomes

In both prokaryotic and eukaryotic cells, the functional machinery operating during the process of translation is the ribosome. Ribosomes are dynamic ribonucleoprotein complexes that catalyze the conversion of genetic information on an mRNA template into a nascent polypeptide chain (Laursen *et al.*, 2005). Consisting of proteins organized on a scaffolding of RNA, ribosomes are composed of a large and small subunit, each respectively possessing functional significance (Melnikov *et al.*, 2012). The small subunit contains the conduit that mRNA is conducted along during translation, the decoding center,

and the tRNA binding sites (A, P and E-sites) (Klinge *et al.*, 2011; de la Cruz, Karbstein and Woolford, 2015). This subunit is the functional site of the decoding process, whereby tRNA is paired to its appropriate mRNA codon and promotes adequate translation fidelity by the monitoring of base pairing (Ramakrishnan, 2002). The large subunit acts as the active site of the ribosome. Significant functional sites include the peptidyl transferase center (PTC), where peptide bonds are catalyzed, the tRNA binding sites and the peptide exit tunnel, which extends throughout the subunit (de la Cruz, Karbstein and Woolford, 2015). The entirety of the ribosomal catalytic activity is carried out by the ribosomal RNA (rRNA), with ribosomal proteins largely residing on the surface of the PTC, acting to stabilize the structure (Lafontaine and Tollervey, 2001).

Ribosomes can be somewhat heterogeneous and possess distinct structural and compositional differences (**Figure 1**). Between species, disparities can include size, structure, protein to RNA proportions, but can also be more exclusive, including domain-specific proteins, mutations within conserved proteins, and expanded fragments of rRNAs (Melnikov *et al.*, 2012). Regardless of the differences, a structurally conserved portion of the ribosome is present in cells from all domains (Ben-Shem *et al.*, 2011). The core is composed of the aforementioned functional sites, such as the decoding site, PTC and tRNA-binding sites. These sites are comprised of 34 conserved proteins, nearly evenly split between the small and large subunit (15 and 19 proteins, respectively) and over four thousand RNA bases (Melnikov *et al.*, 2012). Many of the expanded rRNA segments and protein moieties cluster on the solvent-exposed surface of the ribosome, surrounding the conserved core in an arrangement that does not directly interrupt the functional centers of ribosomes and allows them to be accessible to potential molecular interactors, such as

chaperones and translation factors (Spahn *et al.*, 2001; Melnikov *et al.*, 2012; Ramesh and Woolford, 2016).

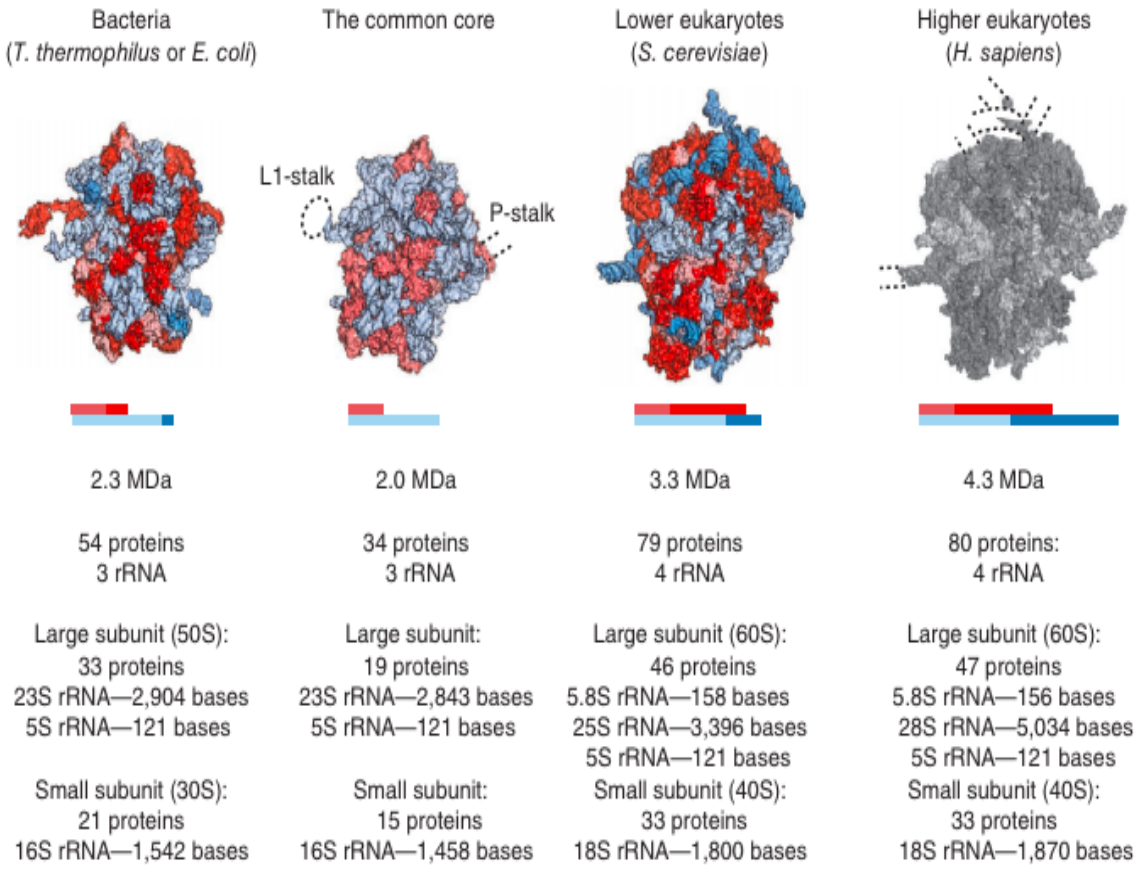

Figure 1. Composition of bacterial and eukaryotic ribosomes, and their shared conserved core. The conserved core is composed of RNA (light blue) and proteins (light red), with additions to conserved proteins (red) and extensions of rRNA (blue). Reproduced with permission from (Melnikov *et al.*, 2012).

In bacteria, the 70S ribosome is composed of the small 30S ribosomal subunit and the large 50S ribosomal subunit (Ramakrishnan, 2002). The small subunit is composed of 21 ribosomal proteins attached to the 16S rRNA chain and organized into three distinct structural subdomains: the body, containing the 5′ domain of the 16S rRNA; the platform, containing the central domain; and the head, which contains the 3′ major domain (Lafontaine and Tollervey, 2001). The large subunit is composed of 33 ribosomal proteins

accompanying two rRNA molecules: the 5S rRNA and the 23S rRNA (Ben-Shem *et al.*, 2011). Surrounding the conserved core, the 70S ribosome possesses 20 bacteria-specific ribosomal proteins (6 and 14 within the small and large subunits, respectively), along with their own extensions within conserved proteins and rRNA (Melnikov *et al.*, 2012).

In eukaryotes, the 80S ribosome is composed of the small 40S ribosomal subunit and the large 60S ribosomal subunit (Wool, 1979). The small subunit is composed of 33 ribosomal proteins attached to the 18S rRNA chain and organized into three distinct structural subdomains: the body, containing the 5′ domain of the 18S rRNA; the platform, containing the central domain; and the head, which contains the 3′ major domain (de la Cruz, Karbstein and Woolford, 2015). The large subunit is composed of 46 ribosomal proteins accompanying three rRNA molecules: 25S, 5.8S and 5S (Ben-Shem *et al.*, 2011).

Compared to its bacterial counterpart, the eukaryotic ribosome is more complex. This statement is attributable to the fact that the eukaryotic ribosome accommodates an additional 46 eukaryotic-specific ribosomal proteins (18 and 28 within the small and large subunits, respectively), with homologous proteins within the conserved core having acquired supplementary insertions and extensions. Similarly, the rRNA also possesses additional blocks of sequences, known as expansion segments, interspersed throughout the core rRNA secondary structure at unambiguous, conserved sites (Jenner *et al.*, 2012; Melnikov *et al.*, 2012; Ramesh and Woolford, 2016). These numerous additions of expanded ribosomal elements account for the approximate 40% larger size of the eukaryotic ribosome compared to the bacterial ribosome, as well as the increased level of complexity within the pathway of eukaryotic translation (Ben-Shem *et al.*, 2011; Melnikov *et al.*, 2012).

1.2. Translation Initiation

Gene expression is a highly controlled process and translation initiation is a critical control point during protein synthesis. Translation initiation acts as the rate-limiting step in translation, influencing subsequent efficiency, mRNA stability and kinetics (Gray and Wickens, 1998; Jackson, Hellen and Pestova, 2010; Benelli and Londei, 2011). Though the fundamental goal of translation initiation is the same in bacteria and eukaryotes, it is the most phylogenetically diverse step within the process of translation (Laursen *et al.*, 2005; Benelli and Londei, 2011).

1.2.1. Eukaryotic Cap-Dependent Mechanism of Translation Initiation

Translation initiation in eukaryotes is a highly regulated and complex process (**Figure 2**). Over time, eukaryotes have evolved sophisticated mechanisms, conformations and factors to initiate the process of translation. The formation of novel mRNA structures, such as the 5' methyl-guanosine cap (^{7}mGpppN) and the poly-adenosine (poly-A) tail, enabled the ability of mRNA to recruit ribosomes in a multi-step process using protein-RNA and protein-protein interactions, while simultaneously maintaining control over gene expression.

Eukaryotic canonical translation initiation is a cyclical process whereby rounds of translation initiation follow the recycling of post-termination complexes (post-TCs) present from the previous rounds of translation (Jackson, Hellen and Pestova, 2010). These complexes are composed of mRNA-bound 80S ribosomes, P-site deacylated tRNA and a minimum of one release factor, commonly eukaryotic release factor 1 (eRF1). At low free Mg^{2+} concentrations, recycling is facilitated by eukaryotic initiation factors (eIFs). eIF3,

alongside eIF1 and eIF1A, mediate the release of the ligands and the disassociation of ribosomes into available 60S subunits, and mRNA-bound and tRNA-bound 40S subunits. Particularly, eIF1 stimulates the release of tRNA, while eIF3j mediates mRNA release, and eIF3, in concert with eIF1 and eIF1A, remain associated with the recycled 40S subunits to prevent their re-engagement with the 60S subunits (Jackson, Hellen and Pestova, 2010).

To mark the onset of initiation, a ternary complex containing eIF2, Met-tRNA$_i^{Met}$ and GTP is formed. This complex subsequently binds to the small ribosomal subunit through the coordinated efforts of eIF1, eIF1A, and eIF3 to yield the 43S pre-initiation complex (PIC) (Kapp and Lorsch, 2004). The 43S PIC is recruited to the 5' cap in a process coordinated by eIF3, poly-A-binding-protein (PABP) and the eIF4F complex, which is composed of eIF4E, eIF4A and eIF4G (Sonenberg and Hinnebusch, 2009). eIF4E is considered the obligatory cap-binding protein. eIF4A is a RNA helicase that couples the hydrolysis of ATP to the binding of RNA and duplex separation, which promotes the "activation" of mRNA via unwinding of 5' untranslated region (5'UTR) secondary structures. eIF4G is a scaffold protein that possesses binding domains that aid in the coordinated assembly of the PIC and its attachment to the template mRNA (López-Lastra, Rivas and Barría, 2005; López-Lastra *et al.*, 2010). The interaction of eIF4G with the mRNA facilitates synchronization amongst the 5' cap, poly-A tail, and mRNA sequence to assemble into a stable, circular messenger ribonucleoprotein (mRNP). A "protein bridge" is then established between the mRNP and the 43S PIC to promote the attachment of the 43S to the mRNA, with aid from eIF4A (Sonenberg and Hinnebusch, 2009).

Once assembled at the cap, the 5'UTR secondary structures are unwound and the energy-dependent 5' to 3' scanning of the mRNA begins, looking to identify an initiation

codon in a favorable sequence context, known as the Kozak consensus sequence (Kapp and Lorsch, 2004). The ribosome halts when stable base-pairing between the anticodon of Met-tRNAi (CAU) and the initiator AUG in the P-site of the 40S subunit occurs (Sonenberg and Hinnebusch, 2009). The recognition of the initiation codon, promotedby eIF1 and eIF1A, induces a conformational change that leads to the displacement of eIF1 and causes hydrolysis of the eIF2-bound GTP within the ternary complex to a GDP-bound state by eIF5, a GTPase-activating protein (López-Lastra, Rivas and Barría, 2005). The subsequent joining of the large ribosomal subunit is mediated by eIF5B and catalyzes the mass dissociation of initiation factors eIF1, eIF3 and any residual eIF2-GDP from the complex (Kapp and Lorsch, 2004; López-Lastra, Rivas and Barría, 2005). Dissociation of eIF1A and eIF5B then occurs following the hydrolysis of eIF5B-bound GTP, which frees up the A-site and results in the formation of an elongation-competent 80S ribosome (López-Lastra, Rivas and Barría, 2005; Sonenberg and Hinnebusch, 2009; Jackson, Hellen and Pestova, 2010). This event is thought to mark the conclusion of translation initiation, though supplementary steps may be required prior to the establishment of a fully active and peptide-bond forming complex (Kapp and Lorsch, 2004).

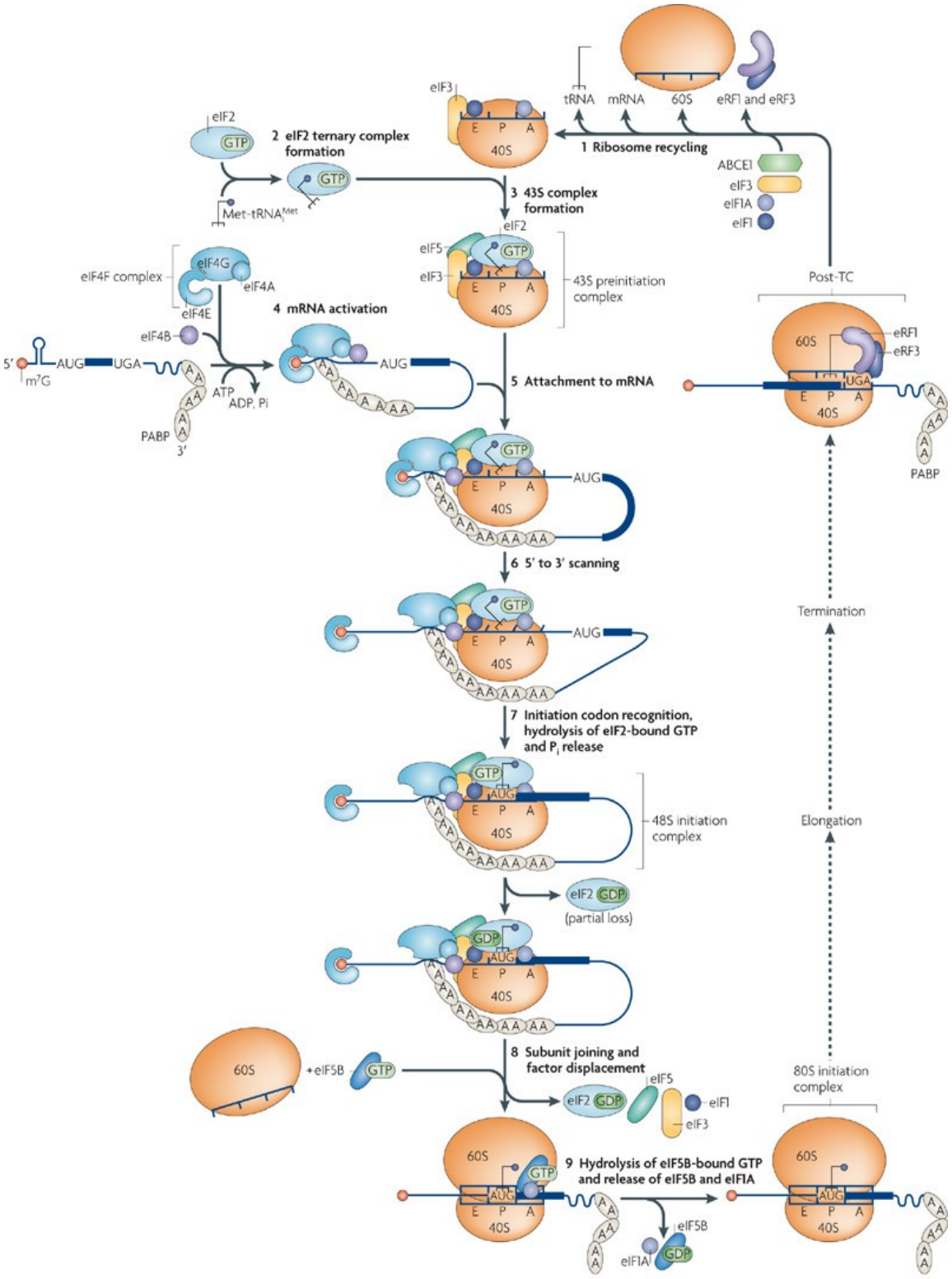

Figure 2. Schematic diagram of cap-dependent translation initiation in eukaryotes. Details of process described in the text. Reproduced with permission from (Jackson, Hellen and Pestova, 2010).

1.2.2. Eukaryotic Cap-Independent Mechanism of Translation Initiation

While most mRNAs are initiated via the canonical cap-dependent mechanism of translation initiation, there are mechanisms that do not rely on the 5' cap to initiate translation. These mechanisms tend to be employed to overcome limitations enacted by the mRNA and cell state, and including leaky scanning, ribosome shunting, termination-reinitiation and internal initiation.

1.2.2.1. Leaky Scanning

The selection of a suitable start codon during the initiation of translation relies heavily on two parameters; the position effect and the Kozak consensus sequence. The position effect refers to how the first AUG encountered is favored as the initiation codon (Kozak, 2002). This has been demonstrated; 1) when a mutation generates an AUG upstream of the authentic start codon and initiation is consequently observed to shift to this upstream site, and 2) the removal of the authentic start codon prompts initiation to shift downstream to the subsequent AUG encountered (Bergenhem *et al.*, 1992; Mével-Ninio *et al.*, 1996; Kozak, 2002).

While an important parameter, codon positioning does not always dictate initiation preference; the Kozak consensus sequence also exerts influence on the decision. The Kozak consensus sequence specifies that the nucleotide composition surrounding the initiation codon, experimentally verified as 5′-GCCGCC(**A/G**)CCAUG**G**-3′ in vertebrates, will considerably impact the affinity of the ribosome for the AUG in question (Kozak, 1987). Thus, if the primary AUG is presented in an unfavorable context, such as lacking a purine at the third base upstream of the AUG (**-3**) as well as a G one base downstream of the AUG

(**+1**), then the ribosome might bypass it and continue scanning further downstream for a more suitable start codon (López-Lastra *et al.*, 2010). This results in initiation from both AUG codons (although, not equally) present in these mRNA sequences (Ryabova, Pooggin and Hohn, 2006). Accordingly, leaky scanning can allow for an mRNA transcript to encode for two different proteins, such that; a) dissimilar proteins are produced if the start codons are out of frame, or b) isoforms with varying N-termini lengths are generated should the start codon remain in-frame (Kozak, 1991a).

Importantly, while leaky scanning predominantly occurs due to a weak consensus sequence, it may also occur if the primary AUG codon is positioned too proximally, within a range of 12 nucleotides, of the 5' cap (Kozak, 1991b). Leaky scanning is a mechanism commonly exploited by mammalian viruses, including but not limited to; *Human papillomavirus* (HPV), *Influenza A* and *Influenza B* (Stacey *et al.*, 2000; Jaafar and Kieft, 2019).

1.2.2.2. Ribosome Shunting

Ribosome shunting is a unique process that incorporates aspects of cap-dependent initiation with internal initiation. Ribosomes are recruited to the mRNA in a cap-dependent approach but scanning through the 5' UTR occurs in a discontinuous manner. The 40S ribosome will bypass complex secondary structures within the leader regions by being shunted from the upstream donor site to a downstream acceptor site as to initiate translation (Hertz *et al.*, 2013). This mechanism allows for the avoidance of inhibitory elements, which simultaneously maintains the ease of scanning and preserves the integrity of potential RNA regulatory elements (Ryabova, Pooggin and Hohn, 2006; López-Lastra *et al.*, 2010).

Ribosome shunting was first discovered in 1993, and is now most extensively documented in *Cauliflower mosaic virus* (CaMV), which was then followed by *Rice tungro bacilliform virus* (RTBV) in 1996, both plant pararetroviruses (Ryabova, Pooggin and Hohn, 2006). Both viruses possess a leader sequence composed of multiple short open reading frames (sORFs) followed by a stable hairpin, which obstructs the option for linear scanning to occur. In both cases, the 5'-proximal sORF within the leader is translated prior to encountering the hairpin structure, which then forces the ribosome to shunt to a downstream portion of the mRNA sequence (pairing ribosome shunting and reinitiation, which is mentioned below). While identified in plant viruses, ribosomal shunting also occurs in animal viruses and in cellular mRNAs (Pooggin and Ryabova, 2018).

1.2.2.3. Reinitiation

Reinitiation is a form of internal initiation that allows for the expression of multiple cistrons from a single mRNA. This occurs when ribosomes have translated a short upstream open reading frame (uORF) and remain associated with the mRNA, with initiation resuming at a downstream start codon within vicinity of the stop codon (López-Lastra *et al.*, 2010; Zinoviev, Hellen and Pestova, 2015). The efficiency of reinitiation varies depending on several features, including the following three. Firstly, reinitiation efficiency is increased as the intercistonic length is increased, but only in the absence of complex secondary structures, as it is believed that the ribosomes reacquire lost initiation factors prior to reinitiation (Kozak, 1987). Secondly, the length of the uORF directly impacts reinitiation efficiency, with shorter uORFs (<30 codons) increasing the rate of occurrence (Ryabova, Pooggin and Hohn, 2006). Thirdly, composition of the uORF

impacts reinitiation efficiency. For example, the presence of complex secondary structures within the uORF, such as a pseudoknot, would decrease efficiency (Somogyi *et al.*, 1993). Overall, the regulatory control over the process of reinitiation is quite complex and presently not well understood, with a multitude of additional influences further impacting the frequency of occurrence (Pooggin, Hohn and Fütterer, 2000).

1.2.2.4. Internal Ribosome Entry Site

Prior to discovery of the various cap-independent initiation mechanisms presented in the sections above, the cap-dependent method was thought to be the only means by which translation initiation could occur within eukaryotes. This idea was challenged in the late 1980s, when viral gene expression studies led to the identification of an alternative mechanism of initiation independent of the 5' cap structure (Pelletier and Sonenberg, 1988). This mechanism, termed Internal Ribosome Entry Site (IRES), allowed for a complex RNA secondary structure, commonly in the 5'UTR, to directly recruit ribosomes within vicinity of the start codon (Komar and Hatzoglou, 2011a). While this mechanism was originally characterized in picornaviruses, additional viral families have been observed to utilize this form of internal initiation, and it has also been recognized in a variety of cellular mRNAs (López-Lastra, Rivas and Barría, 2005).

IRES-mediated translation is a robust process that is commonly exploited by viruses to ensure that the viral mRNA is efficiently translated while the translation of host mRNAs is restricted (López-Lastra *et al.*, 2010). While this primary goal is the same amongst all viruses, the activation of these viral IRESs is diverse. Variances include the structural organization of the IRES element, the requirements of canonical initiation

factors, the requirements of additional, non-canonical proteins (dubbed IRES trans-acting factors, or ITAFS) and the recruitment strategy of ribosomal subunits to the IRES (Kieft, 2008; López-Lastra *et al.*, 2010; Mailliot and Martin, 2018). Because of the significance these IRESs have had in viral infection propagation, substantial effort has been made to understand the intricacies of this process. This has ultimately led to classification of the viral IRESs into four different classes, constructed based on recognized structural and functional characteristics (Kieft, 2008; Mailliot & Martin, 2018). There is a trend whereby IRES elements with more sophisticated structural configurations tend to require fewer additional factors to initiate translation, and as the IRES becomes less structured, more requirements (such as eIFs and ITAFs) become requisite (**Figure 3**) (Kieft, 2008).

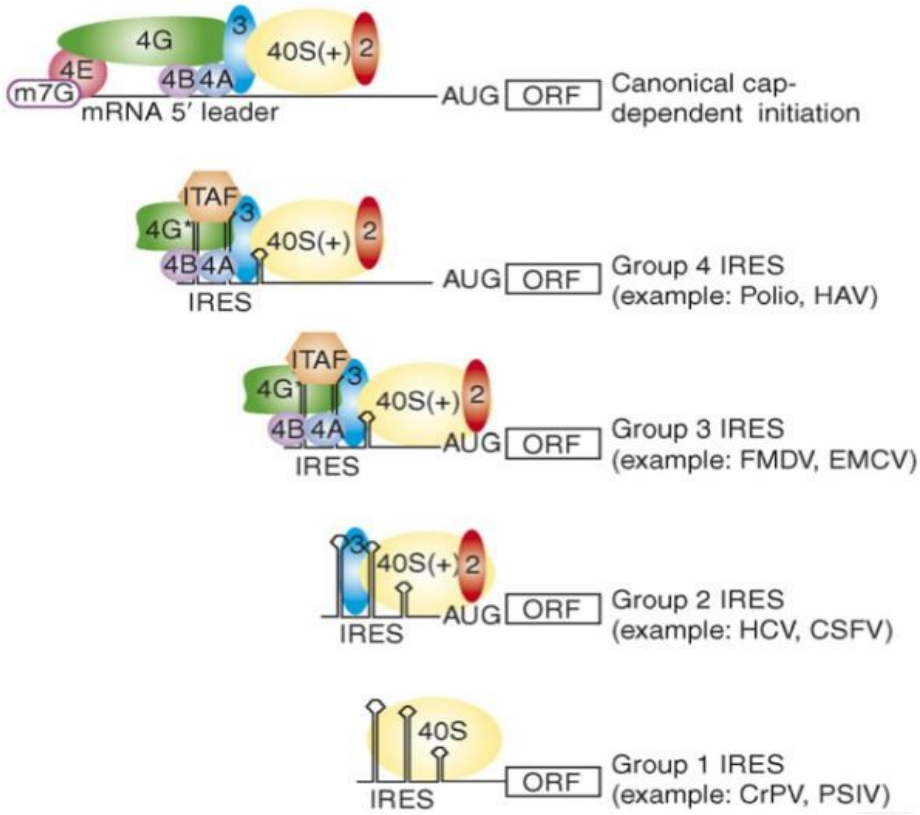

Figure 3. Visual depiction of viral IRES groupings and their requirements to initiate translation, compared with canonical cap-dependent translation initiation. As IRES elements become less structured, there is an increase in the requirements for eIFs and ITAFs to initiate translation. Reproduced with permission from (Kieft, 2008).

The first viral IRES grouping, Class 4, refers to IRESs that form highly intricate secondary and tertiary RNA structures that are capable of directly binding to the ribosome and assembling an 80S complex independent of the scanning process, protein factors or Met-tRNA$_i^{Met}$ (Kieft, 2008; Lozano and Martínez-Salas, 2015; Martinez-Salas *et al.*, 2018; Jaafar and Kieft, 2019). This group is composed of conserved IRESs located in the intergenic region (IGR) between open-reading frames (ORFs) in the viral RNA of single-stranded positive-sense *Dicistroviridae* viruses, such as *Isreali acute paralysis virus* (IAPV), *Plautia stali intestinal virus* (PSIV) and *Cricket paralysis virus* (CrPV) (Mailliot and Martin, 2018). These IGR IRESs adopt a highly compact 3-D structure composed of three pseudoknots (PK) (PKI, II and III), two stem-loops (SL) (SL IV and V), and two independently folding domains (the larger domain possessing regions I and II, and the separate smaller domain III) (Kieft, 2008). The larger domain possesses affinity for the ribosome, as its two stem loops directly contact the small subunit and induce a conformation change of the 40S subunit. The smaller Domain III possesses PKI, which structurally mimics Met-tRNAi and allows for an imitation tRNA-mRNA interaction to occur within the ribosomal A-site (Jaafar & Kieft, 2019; Kieft, 2008). Thus, PKI docks into the A-site, sets the appropriate ORF and initiates translation via pseudotranslocation of PKI into the P-site and a non-AUG start codon into the A-site of the 40S ribosomal subunit (Jan, Kinzy and Sarnow, 2003; Kieft, 2008; Johnson *et al.*, 2017; Jaafar and Kieft, 2019).

The second viral IRES grouping, Class 3, refers to IRES elements present at the 5' end of the viral transcript that directly bind to the 40S subunit and only utilize a limited subset of canonical initiation factors, as well as Met-tRNAi, to initiate translation (Pestova

et al., 2001). This IRES grouping possesses a large, extended general structure, with subdomains of rigid segments interposed by adaptable bridges (Filbin and Kieft, 2009; Mailliot and Martin, 2018). This group consists of members from the family *Flaviviridae*, including the pestiviruses *Classical swine fever virus* (CSFV) and *Bovine viral diarrhea virus* (BVDV), and hepaciviruses such as *Hepatitis C virus* (HCV). The IRES of HCV is exceptionally well documented and its mechanism is reasonably elucidated, and thus will be expanded on.

The 5' UTR of HCV contains four domains (I–IV). Domains I and II are involved in viral replication and the IRES element bridges domains II and III. Domain IV, which is present in HCV but absent from other members of this Class (such as CSFV), contains the AUG start codon within a hairpin structure that is protracted upon IRES presence within the mRNA conduit (Mailliot and Martin, 2018). Mechanistically, the HCV IRES binds to the 40S ribosome in a multi-domain interaction, altering the ribosome conformation and creating a 40S-IRES PIC (Kieft *et al.*, 2001). After 40S ribosomal subunit recruitment, the HCV IRES domain III stably interacts with eIF3, which promotes the recruitment of the ternary complex to the 40S-IRES PIC, and results in the formation of a 48S PIC with the AUG codon present in the ribosomal P-site, without scanning (Pestova *et al.*, 1998; Mailliot and Martin, 2018; Jaafar and Kieft, 2019). The 48S PIC then recruits the 60S subunit, and via subsequent GTP hydrolysis and eIF2 release, yields an 80S ribosome properly positioned at the start codon (Lozano and Martínez-Salas, 2015). This mechanism of initiation, in which 40S recruitment occurs without the requirement of initiation factors, has been cited to be analogous to prokaryotic translation, with the IRES acting as a

functional equivalent to the Shine-Dalgarno sequence (SD) (Gray and Wickens, 1998; Jackson, 2005).

The third and fourth viral IRES groupings, Class 1 and 2, refer to IRESs that require the use of some canonical initiation factors, ITAFs and Met-tRNAi, and are used by members of the *Picornaviridae* family. Class 1 members include the enterovirus *Poliovirus* (PV) and the apthovirus *Rhinovirus* (RV), whereas Class 2 include members such as the cardiovirus *Encephalomyocarditis virus* (EMCV) and apthovirus *Foot-and-Mouth disease virus* (FMDV) (Mailliot and Martin, 2018). These IRESs do not form compact structures, but rather maintain extended and flexible conformations, which consequently makes it problematic to produce high-resolution 3-D models of the elements (Filbin and Kieft, 2009). These two groupings of viral IRESs are similar to one another based on their: presence in the 5' UTR of the viral transcript, sequence length, inability to directly bind to the 40S subunit, requirement of the near-complete set of canonical initiation factors (which does not include eIF4E), presence of a pyrimidine-rich tract located at the 3' end of the IRES and recruitment of ribosomes following the repression of host cap-dependent translation (Jackson, 2005; Kieft, 2008; Mailliot & Martin, 2018). The IRES of PV for Class 1, and IRESs of EMCV and FMDV for Class 2 will be used as examples.

The 5' UTR of PV contains six domains (I–VI), with domain I essential for viral replication and the IRES composed of domains II-VI (Mailliot and Martin, 2018). The PV IRES possesses a cryptic AUG necessary for ribosome recruitment, but the ribosome will scan 160 nucleotides further downstream to initiate translation at a secondary AUG codon (Jackson, 2005; Lozano and Martínez-Salas, 2015; Martinez-Salas *et al.*, 2018). It is known that translation initiation via Class 1 IRESs require all eIFs (minus eIF4E) and the presence

of ITAFs, such as poly(C)-binding protein 2 (PCBP-2) and polypyrimidine-tract binding protein (PTB), but the mechanism has yet to be fully elucidated, with little identified regarding its requirements for canonical factors or their responsibilities (De Breyne *et al.*, 2009; Mailliot and Martin, 2018).

The 5' UTR of EMCV and FMDV contain five domains (I–V), with domain I required for viral replication, and the IRES composed of domains II-V (Mailliot and Martin, 2018). Both EMCV and FMDV possess two in-frame AUG codons and the ribosomes bind to them directly, without scanning (Pestova *et al.*, 2001; Jackson, 2005). With the EMCV IRES, only the second AUG is used as a start codon, present downstream of the beginning of the pyrimidine tract, whereas both AUGs are utilized to initiate translation with the FMDV IRES (Jackson, 2005; Mailliot and Martin, 2018). Also incongruent, EMCV and FMDV have dissimilar ITAF requirements, despite their relatedness (Pestova *et al.*, 2001; Jackson, 2005; Martinez-Salas *et al.*, 2018). In the examples presented of both Classes, a domain is present (domain V and domain IV, for the PV and EMCV/FMDV IRESs, respectively) that provides a binding site for eIF4G. The binding of eIF4G causes the recruitment of eIF4A, which further enhances the eIF4G-IRES interaction and alters the IRES conformation as to increase its binding affinity to the 43S complex (De Breyne *et al.*, 2009). Thus, a shared fundamental mechanism underlies initiation for the unrelated Class 1 and Class 2 IRESs, despite lack of sequence and structural homology (De Breyne *et al.*, 2009).

Unlike picornavirus RNA genomes, which are uncapped and necessitate the use of IRESs, it is assumed that all cellular mRNAs are capped and capable of undergoing cap-dependent initiation (Pestova *et al.*, 2001; Komar and Hatzoglou, 2011b). Despite this

detail, there are many cellular mRNAs that have been identified to possess IRES elements, with both mechanisms operating on the same message (Komar and Hatzoglou, 2011b). It is thought that these cellular IRESs possess two major physiological functions; 1) allow for the translation of mRNAs with complex 5' UTRs under normal physiological conditions, and 2) maintain translation of mRNAs under various (patho)physiological stress conditions (Komar and Hatzoglou, 2011b; Pichon *et al.*, 2012; Moteshareie *et al.*, 2018). As expected, many mRNAs identified to possess an IRES element function to maintain cell homeostasis and play a regulatory role in cell-fate decisions, including transcription factors (HIF-1α), oncogenes (c-*myc*), cell cycle control genes (Cyclin D1) and apoptotic genes (Apaf-1 and Bcl-2) (Stoneley *et al.*, 1998; Coldwell *et al.*, 2000; Lang, Kappel and Goodall, 2002; Sherrill *et al.*, 2004; Shi *et al.*, 2005; Pichon *et al.*, 2012).

While the structure and mechanism of action is well characterized regarding many viral IRESs, the specifics surrounding cellular IRESs are difficult to elucidate. This is attributable to the fact that cellular IRESs are highly diverse, with no conserved sequence or secondary structure unifying them (Komar and Hatzoglou, 2011a; Jackson, 2013). It is important to note that IRES elements are completely dependent on their structural integrity, with the presence of any alterations (such as small indels or point mutations) within the IRES element significantly altering its activity. It has also been suggested that the folding of IRES into its tertiary structure is a dynamic process, implying that specific environmental conditions can cause modification of the IRES structure and consequently alter its translational capabilities (López-Lastra, Rivas and Barría, 2005).

1.2.3. Prokaryotic Mechanism of Translation Initiation

In eukaryotes, translation initiation is a highly complex process relying on mRNA-protein and protein-protein interactions (López-Lastra, Rivas and Barría, 2005). However, translation initiation in prokaryotes is a much simpler process, one that consists of the recruitment of a complex of ribosome-initiator tRNA to the initiation codon of a nascent mRNA strand via direct interaction of the mRNA with the rRNA. Prokaryotic translation initiation only requires the large (50S) and small (30S) ribosomal subunits, mRNA, initiator tRNA, three initiation factors (IF1, IF2 and IF3), and GTP as the energy source (**Figure 4**) (Laursen *et al.*, 2005; Gualerzi and Pon, 2015). The three IFs determine the fidelity and kinetics of the initiation process, each bound as a single copy to critical sites on the 30S subunit (Milón and Rodnina, 2012; Gualerzi and Pon, 2015).

Comparable to the recycling step within eukaryotic initiation, bacterial initiation is activated when subunits from the previous rounds of translation are dissociated and recycled with help from initiation factors. The dissociation process is promoted by IF3 binding to the 30S subunit, in tandem with IF1. IF1 binds to the A-site of the 30S subunit to direct the initiator tRNA into the P-site, which will further stimulate IF3 activity and further promote the dissociation of subunits (Laursen *et al.*, 2005). Marking the onset of the initiation pathway, all three IFs, mRNA, and fMet-tRNA^{fMet} associate with the 30S ribosomal subunit, leading to the formation of the 30S PIC (Simonetti *et al.*, 2009). The 30S PIC associates with the mRNA via a ribosome binding site (RBS), the SD sequence, and causes the initiation codon and the initiator tRNA to be adjusted into the P-site of the ribosome by the IFs (Shine and Dalgarno, 1974; Laursen *et al.*, 2005). Once the decoding process has occurred, the 30S PIC is joined by the large subunit to yield the 70S initiation

complex. This action triggers GTP hydrolysis of IF2, which confirms the start site and forces the dissociation of superfluous complexes (Simonetti *et al.*, 2009). This event is thought to mark the conclusion of translation initiation and allows the translation machinery to proceed to the elongation phase.

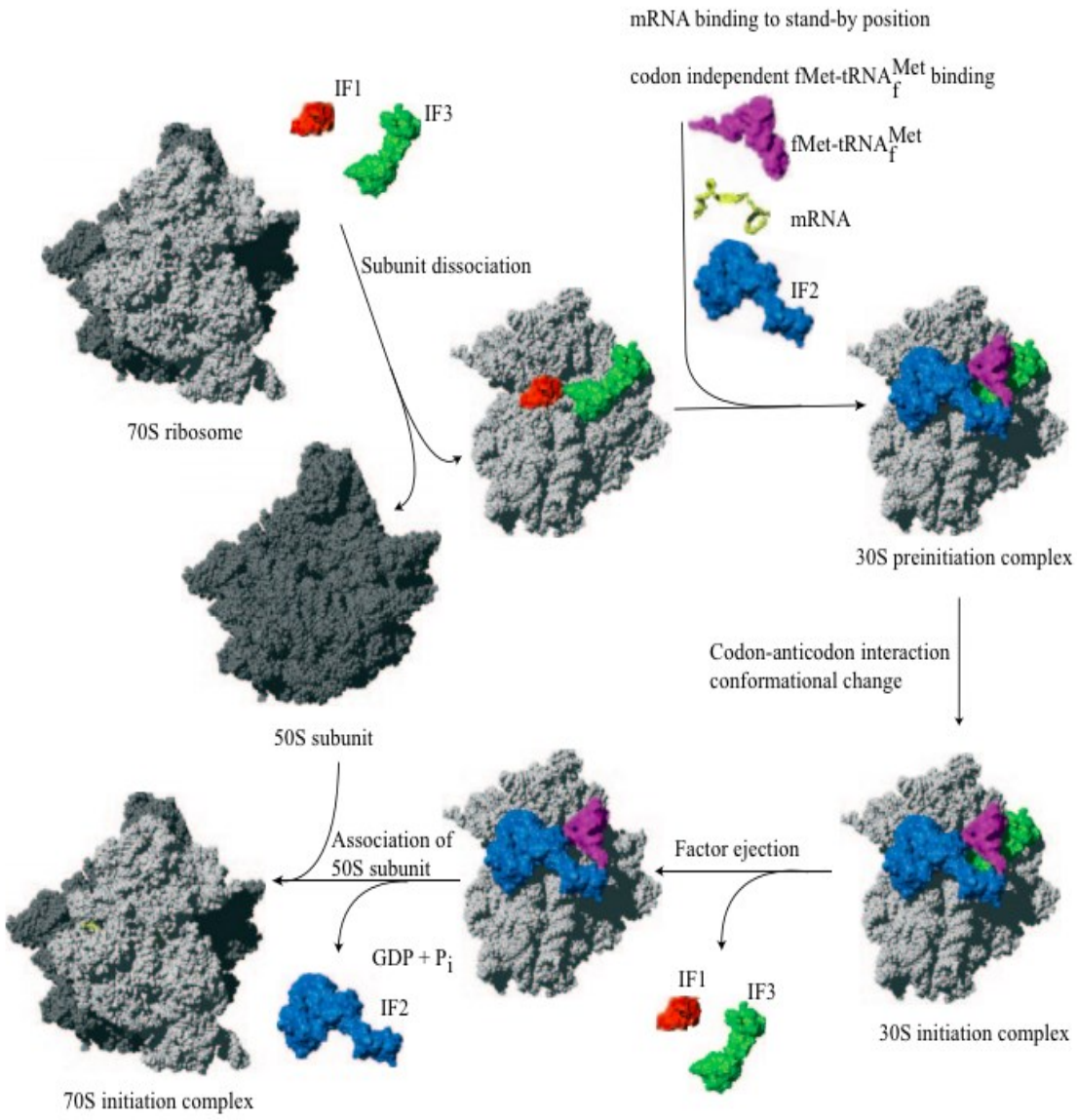

Figure 4. Schematic diagram of prokaryotic translation initiation. The colour coding is as follows: 30S ribosomal subunits (light grey), 50S ribosomal subunits (dark grey), IF1 (red), IF2 (blue), and IF3 (green), mRNA (yellow), and fMet-tRNA^{fMet} (magenta). Details of the process described in the text. Reproduced with permission from (Laursen *et al.*, 2005).

1.2.3.1. Shine-Dalgarno and Enhancer Sequences

The SD sequence is a ribosome binding sequence present in bacterial and archaeal mRNA. This is a purine-rich consensus sequence, AGGAGGU (in *E. coli*), that is present several nucleotides upstream of the AUG start codon and forms base-pairing interactions with a pyrimidine-rich tract, ACCUCCUUA, present at the 3' end of the 16S rRNA (known as the anti-SD) in the 30S subunit (Shine and Dalgarno, 1974; Steitz and Jakest, 1975). The SD helix duplexes with the anti-SD and correspondingly fits within a cleft between the subunit head and the back of the platform of the 16S rRNA. This interaction helps to recruit the 30S ribosome to the mRNA, and along with a spacer region (optimally 4-9 nucleotides in *E. coli*), sets the reading frame and allows for correct placement of the start codon within the P-site of the 30S subunit (Yusupova *et al.*, 2001; Kaminishi *et al.*, 2007; Gualerzi and Pon, 2015). It is important to note that the SD sequence is loosely conserved and can vary in length, along with the spacing between the TIR and itself, with deviations having a discernible effect on translation efficiency (Shultzaberger *et al.*, 2001; Vimberg *et al.*, 2007).

While the binding of the SD consensus sequence is a critical step in the initiation of prokaryotic translation, there are many bacterial mRNAs in which SD sequences are absent and yet, are still efficiently translated (Gualerzi and Pon, 2015). Other sequence elements have been identified that are significant for regulating translation initiation efficiency, one of which is epsilon. The epsilon sequence element (UUAACUUUA) is a translational enhancer that exhibits complementarity to a portion of helix 17 (nucleotides 458-466) of the 16S rRNA (A Golshani *et al.*, 2000; Ashkan Golshani *et al.*, 2000; O'Connor, 2001). While originally identified in the 5'UTR of the phage T7 gene 10

mRNA, it was observed that upstream placement of the element enhanced the translational activity of several genes, including a *lacZ* reporter construct, proposed by base-pairing to its complementary region (Olins and Rangwala, 1989). It was identified by Golshani *et al.* that not only did the initial epsilon element enhance translation of mRNAs containing an SD sequence, but constructs containing an extended epsilon element could also efficiently direct translation initiation in the absence of an SD sequence (Golshani *et al.*, 1997, 2004).

1.3. Proposed Similarities between Prokaryotic & Eukaryotic Translation Initiation

In 2015, Colussi *et al.* published an article in Nature which reported that a eukaryotic IRES, the PSIV IGR IRES, can initiate *in vivo* translation in prokaryotes, with its activity appearing to be independent of an SD-based method of translation initiation (Colussi *et al.*, 2015). To come to this conclusion, they had employed a dual-luciferase expression vector in which the structured PSIV IGR IRES was used to drive expression of Firefly luciferase (FLuc), the experimental reporter, while an SD and enhancer sequence expressed Renilla luciferase (RLuc), an internal control. Their resulting normalized expression data confirmed that the PSIV IGR IRES was functional and can indeed initiate translation within *E. coli*. Colussi *et al.* had also observed that when the structural integrity of the IRES was compromised, via disruption of the PKs within the structure or use of the isolated Domain III, IRES functionality was reduced. This indicated that initiation via the PSIV IGR IRES in bacteria depends on the structural integrity of the compact RNA element, like that in eukaryotes.

They had also elucidated the crystal structure of the IRES element associated with the 70S ribosome complex to a 3.8 Å resolution. The results indicate that the IRES can bind within the tRNA binding sites of bacterial ribosomes, albeit differently than that of eukaryotes. The IRES element appears to form weak and temporary interactions with the bacterial ribosomes, but still remains functional and allows for internal initiation to occur. The proposed mechanism of prokaryotic translation initiation is also dissimilar relative to that observed in eukaryotes. In the prokaryotic mechanism, the ribosome repositions from the IRES to the downstream FLuc AUG to initiate bacterial translation, as opposed to the direct initiation at a non-AUG start codon that is typical of eukaryotic functioning. Importantly, Colussi *et al.* similarly attempted to initiate prokaryotic translation using the IRES of CSFV. This IRES element produced negligible expression levels of FLuc, which illustrated the specificity towards the PSIV IGR IRES. These results, when taken together, appear to suggest that there may be a conserved approach of some features of initiation, based on the structure of this initiation element, between prokaryotes and eukaryotes.

1.4. Focus and Objectives

In both prokaryotes and eukaryotes, translation is initiated via the recruitment of ribosomal subunits to an mRNA transcript in a sequential and multi-step procedure. In general, eukaryotes recruit the small ribosomal subunit to the 5' cap via protein–protein and protein–RNA interactions, followed by scanning of the mRNA sequence. The initiation codon in a favorable context is identified and the 60S subunit joins and completes the assemblage, which is then ready to begin polypeptide synthesis. In prokaryotes, the small ribosomal subunit is directly recruited to the vicinity of the initiation codon. Via a base-

pairing interaction between the 16S rRNA and the SD sequence on the mRNA, the initiation codon is precisely positioned into the ribosomal P-site, equipping the complex for elongation. These pathways have become so divergent over the millennia, eukaryotes are unable to utilize the SD sequence and prokaryotes are incapable of recognizing the 5' cap structure of eukaryotic mRNAs (Colussi *et al.*, 2015). Even with the existence of alternative initiation mechanisms in both domains of life, no universal mechanism had been observed to exist that can overlap between the two, on similar message types.

Conversely, the 2015 Colussi *et al.* article suggested that the PSIV IGR IRES can initiate translation within *E. coli* independent of an SD-based mechanism. It appears that to be functional in a prokaryote, the IRES relies on its 3-D structure and the preservation of its structural integrity to initiate translation, like that in eukaryotes. This would insinuate that there is a conserved approach of initiation based on the structure of this element present between prokaryotes and eukaryotes.

The objective of this thesis is two-fold; 1) to investigate whether the initiation capabilities of the PSIV IGR IRES in prokaryotes is a unique characteristic, or whether other IRES elements, or other complex mRNA structures, can also mediate prokaryotic translation initiation, and 2) to elucidate any potential distinguishing characteristics that allow the PSIV IGR IRES to initiate translation in prokaryotes. To address these objectives, the methodology included the following two key features, all of which were carried out using the same dual-luciferase reporter construct as that of Colussi *et al.* (2015).

Firstly, to deduce whether the prokaryotic initiation capabilities of the PSIV IGR IRES is a unique characteristic, or whether it is a common property amongst IRES elements, a variety of IRES (viral and cellular) and structured RNA elements, were cloned

between the two luciferases of the dicistronic construct. The luciferase expression data presented by Colussi *et al.* was confirmed via experimental replication (using the constructs received from Colussi *et al.*) for PSIV, used as a positive control, and CSFV, used as a negative control. Following protein expression mediated by the various IRES and structured RNA elements, it appears that the ability of the PSIV IGR IRES to initiate prokaryotic translation is a distinct property that is not shared among the other IRES elements tested.

Secondly, to determine whether the PSIV IGR IRES possesses any specific differentiating characteristics, the sequence of the IRES element was mutated (several times) and their ability to initiate translation was evaluated by expression analysis. During these experiments, focus was placed on the initiating capabilities of the PSIV Domain III, as it was previously reported that this domain can bind to the 70S ribosomes through its proper positioning in the decoding groove of the small ribosomal subunit (Zhu *et al.*, 2011; Colussi *et al.*, 2015). Additionally, it was observed that relative to the luciferase expression data for the Full-Length Domain III presented by Colussi *et al.,* our experimentally replicated Full-Length Domain III construct expressed luciferase at considerably lower rates.

The first part of the second objective involved alterations to the Domain III sequence by truncation of the structure via 5' end nucleotide deletions. The truncation of Domain III altered the structural integrity of the IRES and destroyed a PK necessary for eukaryotic translation initiation. The second part of the second objective involved the introduction of four-point mutations within two regions of interest in the truncatedelement. The two specified regions of interest had been identified to possess complementarity

between itself and the 16S rRNA, and the point mutations were introduced as to disrupt any base-pairing capabilities. The Truncated Domain III construct was observed to have an increased ability to initiate prokaryotic translation relative to the Full-Length Domain III, whereas the mutated elements were observed to have a severely reduced ability to initiate prokaryotic translation relative to the Truncated Domain III. Together, these results indicate that the 3-D structure of the IRES appears to have less significance in prokaryotes than previously suggested and that the complementary regions within Domain III might bear some responsibility in the ability of this IRES to initiate prokaryotic translation.

2. Journal Article

2.1. Introduction

Translation is a universally conserved step within the process of gene expression. Of the steps that comprise translation, translation initiation is the most phylogenetically diverse (Laursen *et al.*, 2005). While both prokaryotes and eukaryotes initiate translation via the recruitment of ribosomal subunits to an mRNA transcript in a sequential and multi-step procedure, the details of their execution are dissimilar (Laursen *et al.*, 2005; Benelli and Londei, 2011). In prokaryotes, the small ribosomal subunit is directly recruited to the vicinity of the initiation codon, using only three initiation factors (Gualerzi and Pon, 2015). Through a RNA base-pairing interaction between the 16S rRNA and the Shine-Dalgarno (SD) sequence on the mRNA, the initiation codon is precisely positioned into the ribosomal P-site, equipping the complex for elongation (Laursen *et al.*, 2005). Alternatively, eukaryotes predominantly recruit the small ribosomal subunit to the 5' cap via protein–protein and protein–RNA interactions, with the assistance of over a dozen initiation factors (López-Lastra, Rivas and Barría, 2005). The directional scanning of the mRNA sequence occurs to identify the initiation codon in a favorable context and the 60S subunit joins to complete the assemblage, which is then ready to begin polypeptide synthesis (Sonenberg and Hinnebusch, 2009). Prokaryotic and eukaryotic translation initiation pathways have become so divergent over the millennia, that bacteria and eukaryotes are unable to recognize or utilize the translational signals present in the other's mRNA sequence (Colussi *et al.*, 2015). Thus, despite the existence of alternative initiation mechanisms in both domains of life, no universal mechanism had been observed to exist that can overlap between the two, on similar message types.

Importantly, eukaryotes can also initiate translation through alternative initiation mechanisms, exemplified by the use of an Internal Ribosome Entry Site (IRES). IRES allows for a complex RNA secondary structure to directly recruit ribosomes within vicinity of the start codon, operating independently of the 5' cap structure (Komar and Hatzoglou, 2011a). This mechanism has been found to be used by both cellular mRNAs, to maintain cell homeostasis and play a regulatory role in cell-fate decisions, as well as viruses, to ensure that the viral mRNA is efficiently translated while the translation of host mRNAs is restricted (López-Lastra *et al.*, 2010; Komar and Hatzoglou, 2011a). The structure and mechanisms of viral IRESs are reasonably well-characterized compared to cellular IRESs.

The primary goal of IRES functioning is the same amongst all viruses, but the activation and presentation of these viral IRESs is quite diverse. Viral IRESs can differ with regards to their structural organization, requirements of canonical initiation factors, requirements of additional, non-canonical proteins (dubbed IRES trans-acting factors, or ITAFS) and recruitment strategy of ribosomal subunits (Kieft, 2008; López-Lastra *et al.*, 2010; Mailliot and Martin, 2018). Nevertheless, because of the significance these IRESs have had in viral infection propagation, substantial effort has been made to understand the intricacies of this process. This has ultimately led to classification of the viral IRESs, constructed based on recognized structural and functional characteristics (Kieft, 2008; Mailliot & Martin, 2018).

Class 1 of the viral IRESs is of particular importance, as it possesses the most streamlined mechanism of ribosomal manipulation known. This Class is composed of IRESs found in the intergenic region (IGR) between open-reading frames (ORFs) in the viral RNA of single-stranded positive-sense *Dicistroviridae* viruses, and include those

found in *Plautia stali intestinal virus* (PSIV) and *Cricket paralysis virus* (CrPV) (Mailliot and Martin, 2018). These IGR IRESs adopt a highly compact 3-D RNA structure that is capable of directly binding to the ribosome and assembling an elongation-competent 80S ribosome independent of the scanning process, protein factors or Met-$tRNA_i^{Met}$, by mimicking the tRNA-mRNA interaction within the decoding site of the small ribosomal subunit and inducing a pseudotranslocation event (Kieft, 2008; Lozano and Martínez-Salas, 2015; Martinez-Salas *et al.*, 2018; Jaafar and Kieft, 2019).

It appears that one member of this Class is not only capable of controlling eukaryotic ribosomes, but also prokaryotic ribosomes. In 2015, Colussi *et al*. reported that the PSIV IGR IRES can initiate *in vivo* prokaryotic translation (Colussi *et al.*, 2015). They concluded that the activity of the PSIV IGR IRES is independent of an SD-based method of translation initiation. Rather, the IRES can directly manipulate the phylogenetically conserved core of the ribosome by binding to the ribosomal P-site using tRNA mimicry, like that in eukaryotes, to initiate prokaryotic translation. Importantly, this IRES was deduced to rely on its 3-D structure and the preservation of its structural integrity to initiate translation, also as in eukaryotes. These results would provide evidence towards a conserved approach of initiation based on the structure of this element in both prokaryotes and eukaryotes.

This discovery led us to study different aspects surrounding the ability of the PSIV IGR IRES to initiate prokaryotic translation. Firstly, we sought to deduce whether the initiation capabilities of the PSIV IGR IRES in prokaryotes is a unique characteristic, or whether other IRESs, or other complex RNA structures, can also mediate prokaryotic translation initiation. Secondly, we attempted to elucidate any potential distinguishing

characteristics that allow the PSIV IGR IRES to initiate translation in prokaryotes. To address these objectives all the methodology was carried out using the dual-luciferase reporter construct of Colussi *et al.* (2015). Our results indicate that the ability of the PSIV IGR IRES to initiate prokaryotic translation appears to be a distinct property that is not shared between the other IRES elements studied. It also seems that the 3-D structure of this IRES has less significance than previously suggested and that the presence of complementary regions within the Domain III of the IRES might bear some responsibility in its ability to initiate prokaryotic translation.

2.2. Methods

2.2.1. Organization of Dicistronic Reporter Plasmid

The inducible expression vector utilized was a modified pET30a (Novagen) plasmid, created and provided by the lab of Jeffery Kieft. This plasmid possesses DNA containing the PSIV IGR IRES (nucleotides 6000-6195) between the *Renilla reniformis* luciferase (Renilla, RLuc) and *Photinus pyralis* luciferase (Firefly, FLuc) genes (**Figure 5)**. The RLuc gene is under the translational control of an SD and enhancer sequence, while the FLuc gene is expressed by the upstream IRES. The dicistronic genes are under the transcriptional control of a T7 promoter and the plasmid possesses a kanamycin resistance gene.

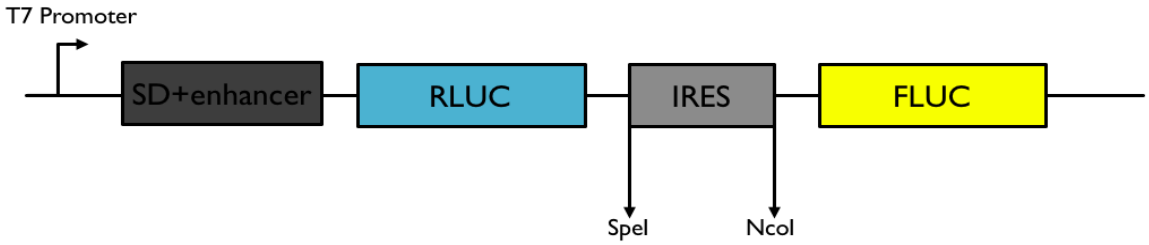

Figure 5. The dual-luciferase reporter construct within the modified pET30a vector utilized. This construct is under the transcriptional control of a T7 promoter. The internal control, RLuc, is under translational control of an SD and enhancer sequence. The experimental reporter, FLuc, is under the translational control of an IRES element. The IRES element is flanked by two restriction sites, *SpeI* and *NcoI*, for which cloning of various inserts was used.

2.2.2. Cloning of Dual-Luciferase Constructs

2.2.2.1 Viral IRES, Cellular IRES, and Structured RNA Elements

Each of the cellular IRES element inserts, from *FLO8*, *GPR1*, *HAP4* and *NCE102* were created via PCR amplification from the genomic DNA of *Saccharomyces cerevisiae* mating type α. Each of the structured RNA element inserts were IRES elements from *FLO8* and *GPR1*, which were PCR amplified from the genomic DNA of *S. cerevisiae* matingtype α and inserted into the restriction sites in reverse. Each of the additional two viral IRES element inserts, from FMDV and the IGR of CrPV, were created via PCR amplification from plasmids possessing the targeted IRES (FMDV IRES amplified from the pVITRO-HPV8 L1L2 plasmid, which was a gift from Richard Roden (Addgene plasmid # 52598) (Kwak *et al.*, 2014) and CrPV IRES amplified from pBgal/CAT plasmid provided by Dr. Martin Holcik). The PSIV IGR IRES and CSFV dual-luciferase constructs were kindly provided by the lab of Jeffery Kieft. The forward and reverse oligonucleotide primers (IDT) (**Appendix 1. Table 2**) possessed intact *SpeI* and *NcoI* restriction sites flanking the hybridization sequence. Amplified inserts were ligated into the Dual-LUC pET30a vector between the *SpeI* and *NcoI* restriction sites using T4 DNA Ligase (New England Biolabs).

2.2.2.2. Mutated Domain III Elements

Each of the PSIV Domain III inserts (**Table 1)** were synthesized by hybridizing complimentary forward and reverse oligonucleotide primers (IDT) possessing the desired sequence with specified mutations. The forward and reverse oligonucleotide primers (IDT) possessed destroyed *SpeI* and *NcoI* restriction sites flanking the hybridization sequence to be used as a confirmation method for cloning. These inserts were ligated into the Dual-

LUC pET30a vector between the *SpeI* and *NcoI* restriction sites using T4 DNA Ligase (New England Biolabs).

Table 1. Sequence of mutated Domain III inserts. Strikethrough represents deleted nucleotides. Underline represents mutated bases. Bolded nucleotides represent complementary regions to the 16S rRNA. Respective destroyed restriction sites flank the 5' and 3' ends of the presented sequences.

Construct Name	Sequence
Full-Length Domain III	TGCTCGCTCAAACATT**AAGTGGTGTTGTGCG**AAAAGAATCTCACTTCAAGAAAA
Truncated Domain III	~~TGCTCGCTCAA~~ACATT**AAGTGGTGTTGTGCG**AAAAGAATCTCACTTCAAGAAAA
ACTT	~~TGCTCGCTCAA~~ACATT**AAGTGG**TGTTACTTGAAAAGAATCTCACTTCAAGAAAA
CACT	~~TGCTCGCTCAA~~ACATTAACACTTGT**TTGTGCG**AAAAGAATCTCACTTCAAGAAAA

2.2.3. Bacterial Cell Cultures

Each of the cloned constructs were transformed into *E. coli* BL21 (DE3) competent cells (Invitrogen) using the heat shock method (Inoue, Nojima and Okayama, 1990). Positive transformants were confirmed via Sanger sequencing at the Ottawa Hospital Research Institute (OHRI).

2.2.4. Luciferase Induction and Measurements

An overnight starter culture was generated using transformed BL21 (DE3) cells grown in Luria Broth (LB) with kanamycin at 37°C with constant agitation. To begin the protein induction process, LB containing kanamycin was inoculated with the overnight starter culture and grown with agitation to an absorbance of 0.4-0.5 at 600 nm (measured

on a Pharmacia Biotech Ultrospec 3000 spectrophotometer). The cultures were then induced with 1 mM IPTG (BioShop) and allowed to grow for an additional 24 hours at 16°C prior to sample acquisition.

Extracted cells were pelleted by centrifugation and washed (1x PBS) prior to being re-pelleted and re-suspended in 250 µL 1X Passive Lysis Buffer (PLB, Promega). Luciferase activity was monitored using the Dual-Luciferase® Reporter Assay System (Promega), per manufacturer's instructions. Luciferase activity was measured in an opaque white 96-well microplate (Nunc Plate, Thermo Fisher) using the Cytation 5 Cell Imaging Multi-Mode Reader (BioTek Instruments) performed at 23±2°C. Experiments used the automatic dual-injection mode, which involves the automatic injection of assay buffers into the wells containing cell lysates, and measured the resulting light intensity emitted over the specified time-period (expressed as Relative Light Units; RLU's). PMT gain was set to a sensitivity of 80 and the optimal well reading height was determined to be 1 mm for the measurement of RLuc expression. PMT gain was set to a sensitivity of 150 and the optimal well reading height was determined to be 5 mm for the measurement of FLuc expression.

2.2.5. Statistical Analyses

Figures of the cellular IRES and RNA element secondary structures were predicted and generated by RNAFold Server, in which the linear sequence of the element was entered and the minimum free energy (MFE) structure was generated using the algorithm of Zuker and Stiegler (Zuker and Stiegler, 1981; Hofacker, 2003).

IRES activity was deduced from the calculated and graphed normalized luciferase expression data at the 24-hour time point. Normalization of the dual-luciferase data,

calculated as FLuc (experimental reporter) activity divided by RLuc (control reporter) activity, was done to correct for variability in growth patterns, induction, and potential differences in protein stability between individual cultures. Normalized Δ Fold Activity represents the fold change in normalized luciferase activity between the PSIV construct and other IRES constructs, set between 0 and 1, calculated as the average normalized luciferase activity of the specified IRES construct divided by the average normalized luciferase activity of the PSIV IRES construct. Averages for each IRES construct were generated from a minimum of four biological and three technical replicates. Error bars present within the graphs represent the standard errors of the average values for each construct.

To assess differences between the IRES activities of constructs, one-way analysis of variance (ANOVA) was applied followed by a *post hoc* Tukey Test. Statistical significance was assumed for values of $p < 0.05$.

2.3. Results and Discussion

Despite having a shared objective, the process of translation initiation is highly disparate between eukaryotes and prokaryotes, both mechanistically and with the initiation factors used. Both domains initiate translation via the recruitment of ribosomal subunits to an mRNA transcript in a sequential and multi-step procedure. Although, eukaryotes generally utilize significantly more initiation factors than prokaryotes, and the process is more complex. Eukaryotic mRNAs possess a 5' cap structure to which the small ribosomal subunit is recruited to through protein–protein and protein–RNA interactions, followed by scanning of the 5'UTR as to identify an initiation codon in a favorable context. Once identified, the 70S subunit joins and completes the assemblage, which is then ready to begin polypeptide synthesis. However, prokaryotic mRNAs possess an SD sequence upstream of the initiation codon, to which the small ribosomal subunit is directly recruited to. Through an RNA base-pairing interaction between the 16S rRNA and the SD sequence on the mRNA, the initiation codon is placed precisely into the ribosomal P-site and the complex is ready to begin elongation.

The 2015 Colussi *et al.* article suggested that the eukaryotic PSIV IGR IRES can initiate translation within a prokaryotic organism (*E. coli*). The activity of this IRES element appears to be independent of an SD-like mechanism of initiation but is rather attributable to its manipulation of the conserved sites of the ribosomal subunits. Like that in eukaryotes, alterations of the 3-D structure of the IRES reduced its capability to initiate translation, suggesting a conserved approach of initiation, based on the structure of this initiation element, between prokaryotes and eukaryotes.

In the current work, we used the same dual-luciferase reporter construct as that of Colussi *et al.* (2015) to monitor the levels of translation that each experimental construct initiated. With use of the dual-luciferase construct, the RLuc acts as an internal control, under the translational control of an SD and enhancer sequence and the FLuc is the experimental reporter, whose expression is driven by the IRES element inserted upstream. Normalized ratios of luciferase expression are calculated to remove between-sample variability caused by external influences.

2.3.1. Capabilities of Eukaryotic IRESs to Initiate Prokaryotic Translation

We determined whether other IRESs are capable of initiating translation in *E. coli* by cloning a variety of viral and cellular IRES elements, as well as several other highly structured RNA sequences, between the luciferase genes of the dual-luciferase reporter construct (Colussi *et al.*, 2015) . The yeast cellular IRESs utilized were those from *FLO8*, *GPR1*, *HAP4* and *NCE102* (**Figure 6**). *FLO8*, *GPR1* and *NCE102* have all been implicated as invasive growth genes, whereas *HAP4* is involved in respiratory gene expression, with each respective IRES elements necessary for these functions (Seino *et al.*, 2005; Gilbert *et al.*, 2007). The viral IRESs utilized were those from the IGR of PSIV, the IGR of CrPV, CSFV and FMDV (**Figure 7**). The PSIV IGR IRES was used as a positive control and the CSFV IRES was used as a negative control, both of which produced results that corroborated Colussi *et al.* findings. CrPV is a Class 4 IRES, which folds compactly and initiates eukaryotic translation independent of any canonical initiation factors or ITAFs (Kieft, 2008). FMDV is a Class 2 IRES, which forms an extended and flexible

conformation and initiates eukaryotic translation with the help of a significant amount of canonical initiation factors and IT AFs (Filbin and Kieft, 2009; Mailliot and Martin, 2018).

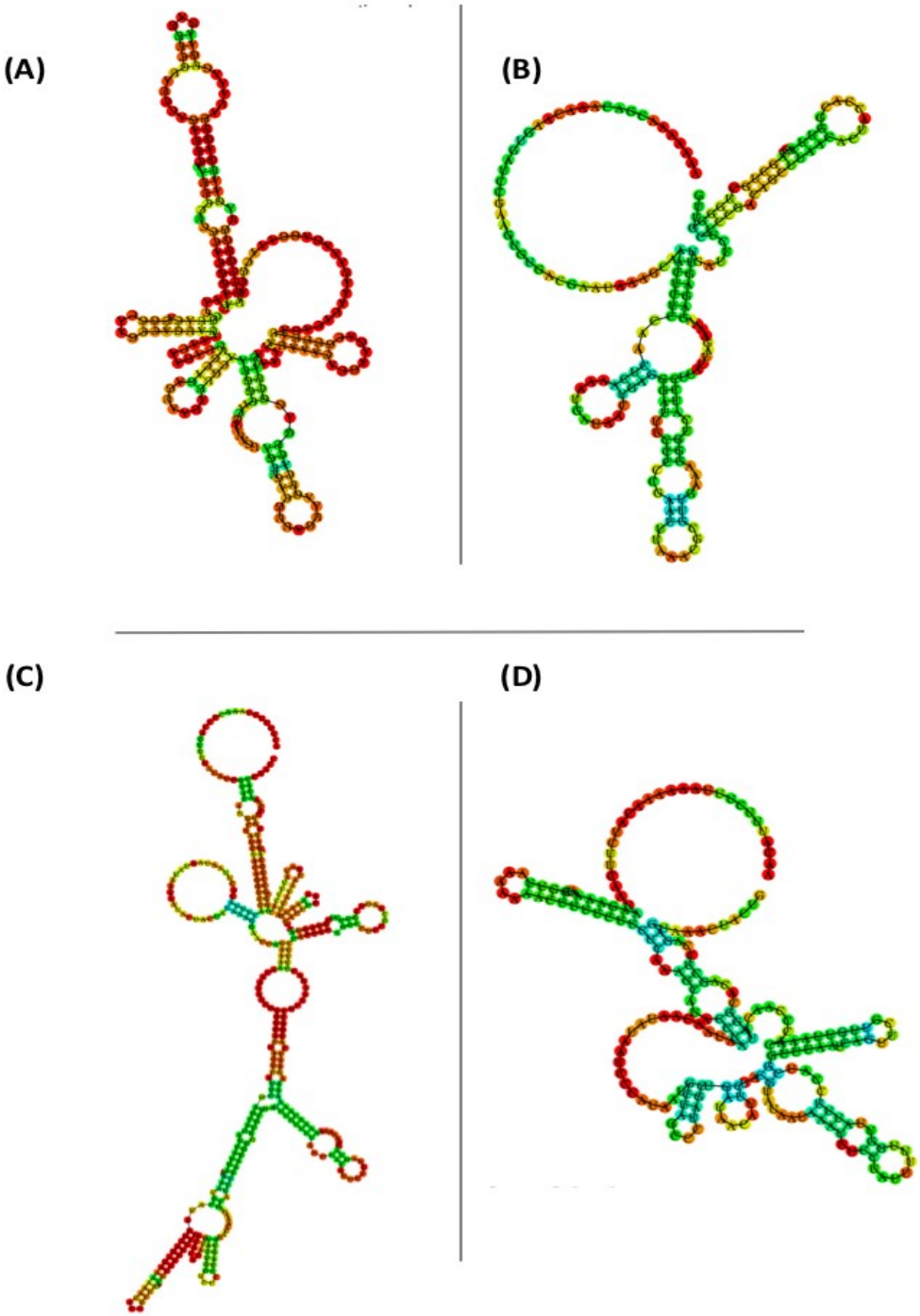

Figure 6. RNAfold predicted MFE secondary structures of the yeast cellular IRES elements used. (**A**) *FLO8* IRES, (**B**) *GPR1* IRES, (**C**) *HAP4* IRES, and (**D**) *NCE102* IRES.

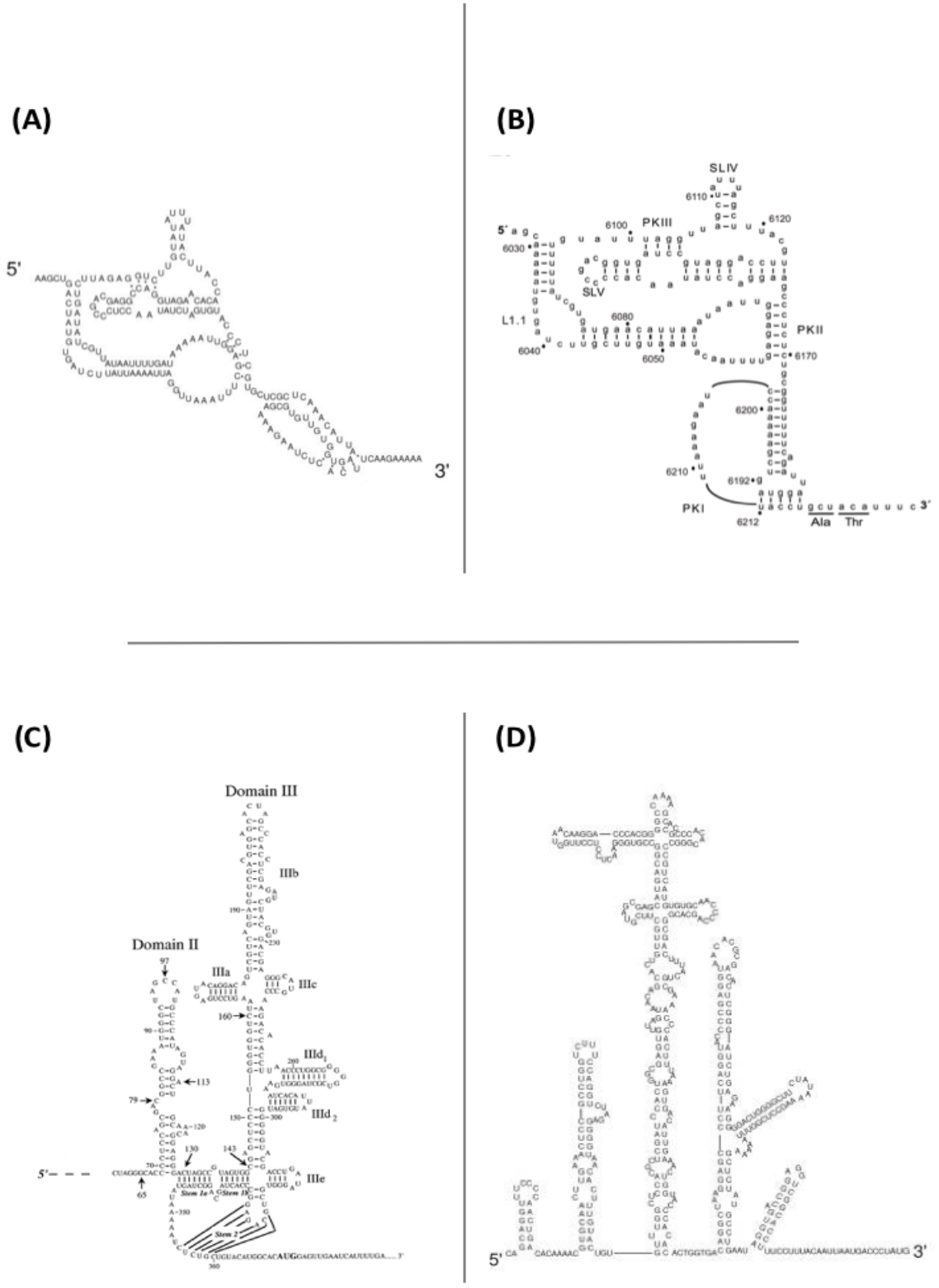

Figure 7. Secondary structures of viral IRES elements used. (**A**) PSIV IGR IRES, reproduced with permission from (Filbin and Kieft, 2009). (**B**) CrPV IGR IRES, reproduced with permission from (Wang and Jan, 2014). (**C**) CSFV IRES, reproduced with permission from (Fletcher and Jackson, 2002). (**D**) FMDV IRES, reproduced with permission from (Filbin and Kieft, 2009).

Finally, to guarantee that not just any structured RNA element can initiate prokaryotic translation, IRES elements from *FLO8* and *GPR1* were inserted in-between the *SpeI* and *NcoI* restrictions sites of the dual-luciferase construct in reverse (**Figure 8**). The reversed cellular IRES elements possess an absence of IRES functionality, but maintain a stable secondary structure (Xia, 2007; Xia and Holcik, 2009).

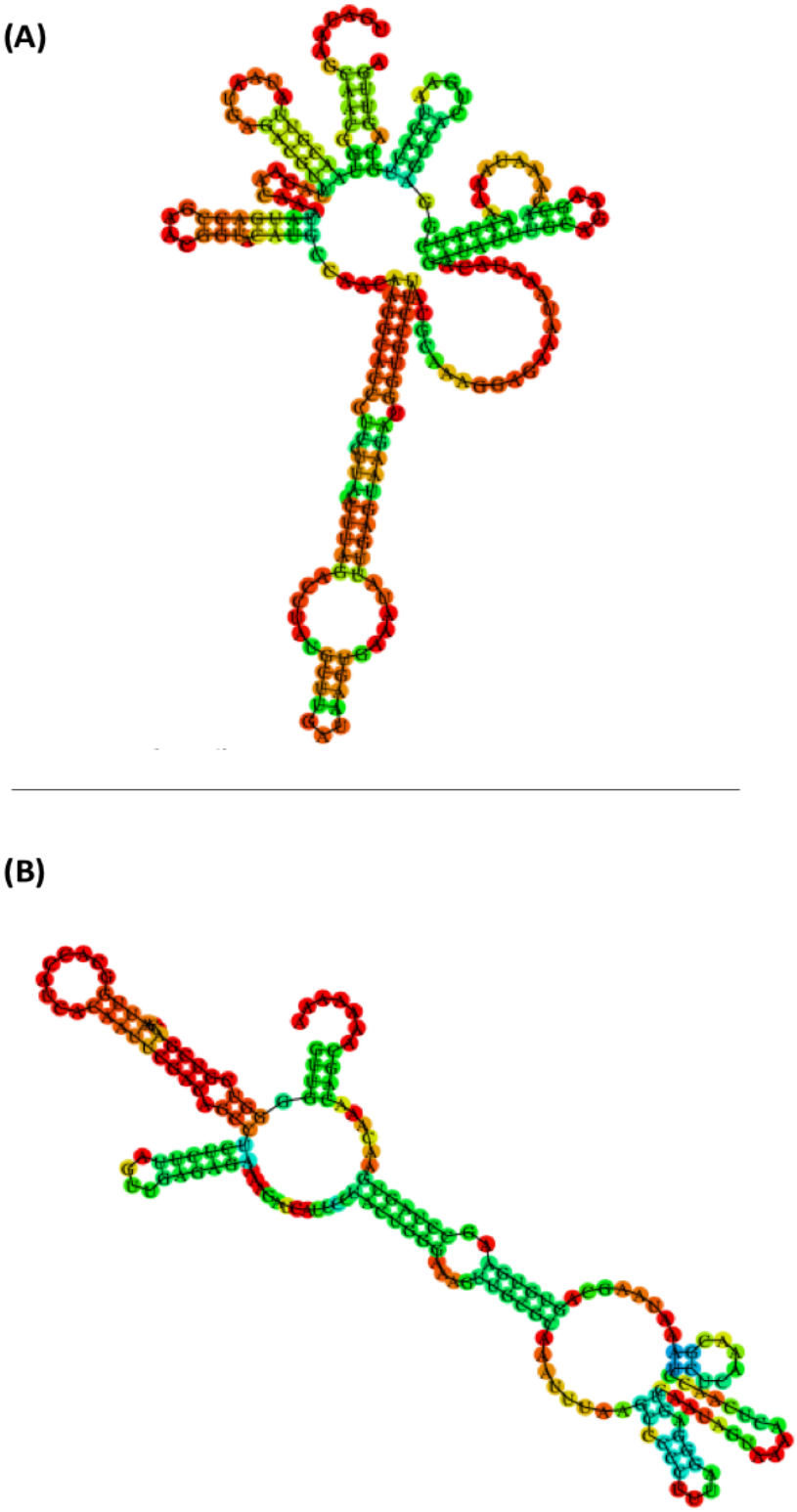

Figure 8. RNAfold predicted *MFE secondary structures of* reverse inserted yeast cellular IRES elements used. (**A**) *FLO8* reverse structured RNA element, and (**B**) *GPR1* reverse structured RNA element.

Based on the normalized luciferase expression data obtained, it appears that the ability of the PSIV IGR IRES to initiate prokaryotic translation is a distinct property that is not common between the other IRES elements employed (**Figure 9**). The cellular IRES elements, such as those from *NCE102*, *HAP4* and *FLO8*, show the highest levels of IRES activity via normalized luciferase expression, at 25%, 23% and 12% of that produced by PSIV, respectively. While not negligible results, these values are not significant enough to provide definitive evidence that these IRES elements can drive initiation of prokaryotic translation. At the lower end, other IRES elements, such as those from *GPR1,* FMDV and CrPV, exhibited negligible results, at just 1%, 0.2% and 1% of that produced by PSIV. Taken into consideration with the inconsequential expression data generated from the insertion of highly structured RNA elements, all approximately 2% of PSIV expression, these results indicate that prokaryotic initiation is not strictly dependent on just any complex secondary structure attempting to sequester ribosomal subunits.

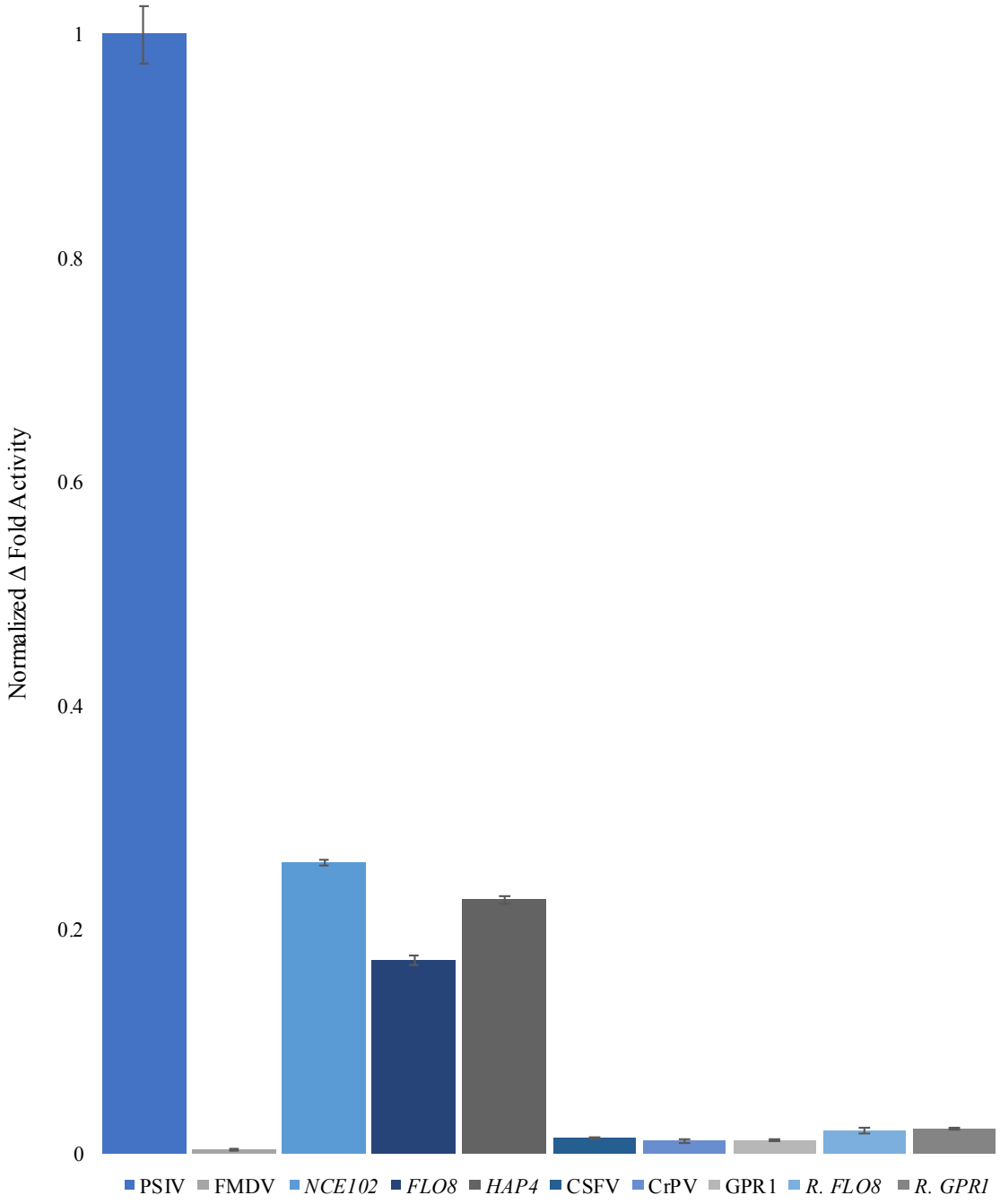

Figure 9. Normalized fold change of luciferase expression illustrating the difference in ability of various IRES and RNA elements to express FLuc relative to the SD and enhancer controlled RLuc, compared to the PSIV construct, within E. coli. Each experimental construct consists of a normalized average generated from a minimum of four biological and three technical replicates. The error bars represent the standard error of the average values. ANOVA ($F=316.7$, $p<0.001$) followed by post hoc Tukey test showed that PSIV IGR IRES activity differed significantly from each of the remaining constructs ($p<0.0001$).

2.3.2. Truncation of Domain III

Colussi *et al.* had indicated that the structural characteristics of the PSIV IGR IRES element enabled it to initiate prokaryotic translation, and that its structural integrity is paramount to its functionality. To assess whether structural integrity has a fundamental purpose in the functionality of the PSIV IGR IRES, focus was placed on the Domain III of the IRES, following previous evidence that it is able to independently bind to the decoding groove of the 16S rRNA, as well as initiate prokaryotic translation (Zhu *et al.*, 2011; Colussi *et al.*, 2015).

To determine the relative importance of the PSIV Domain III structural integrity to its own functionality, its tertiary structure was disrupted. This disruption was achieved through truncation of the PSIV Domain III by the deletion of upstream nucleotides, which induced the destruction of PKI (**Figure 10**). Theoretically, this disruption should have rendered the Truncated Domain III non-functional relative to the Full-Length Domain III. The resulting expression data indicates otherwise, with the expression values from the Truncated Domain III significantly increased, at 76%, from the 16% exhibited by the Full-Length Domain III (**Figure 11**). Should the structural integrity have played a commanding role, deletion of any nucleotides from the Domain III structure would result in expression approaching zero percent. This leads us to believe that the secondary structure plays a reduced role in the IRES mechanism than previously thought.

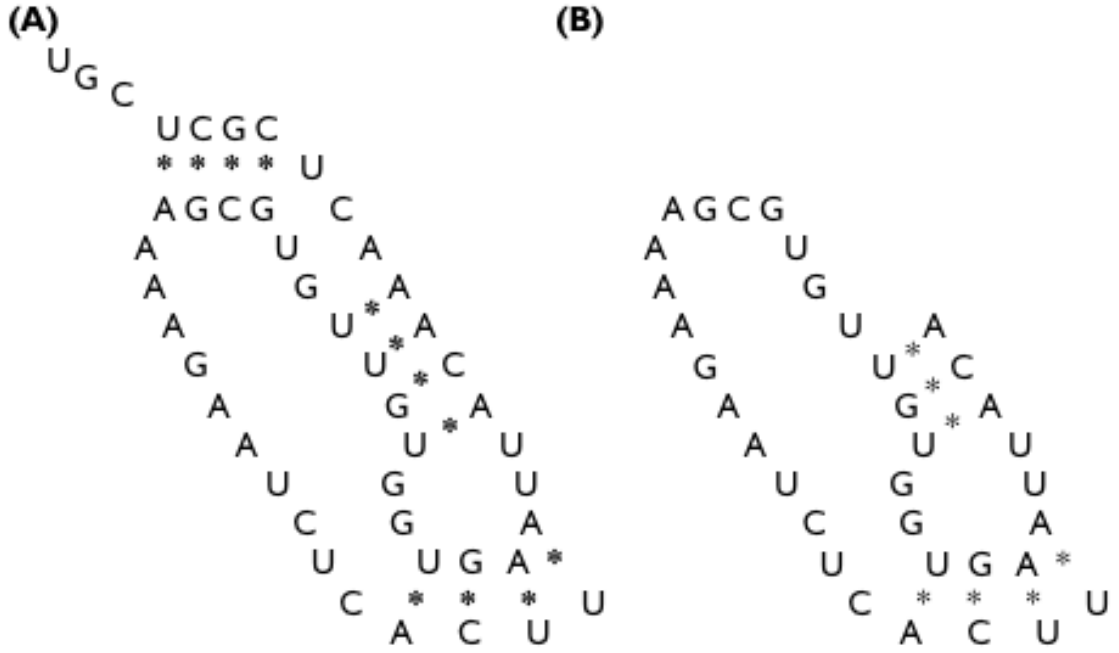

Figure 10. (**A**) Visual representation of the PSIV Full-Length Domain III. Sequences forming Domains I and II were deleted. (**B**) Visual representation of the PSIV Truncated Domain III, whereby the first 11 nucleotides were removed from the 5' end of the Full-Length Domain III sequence.

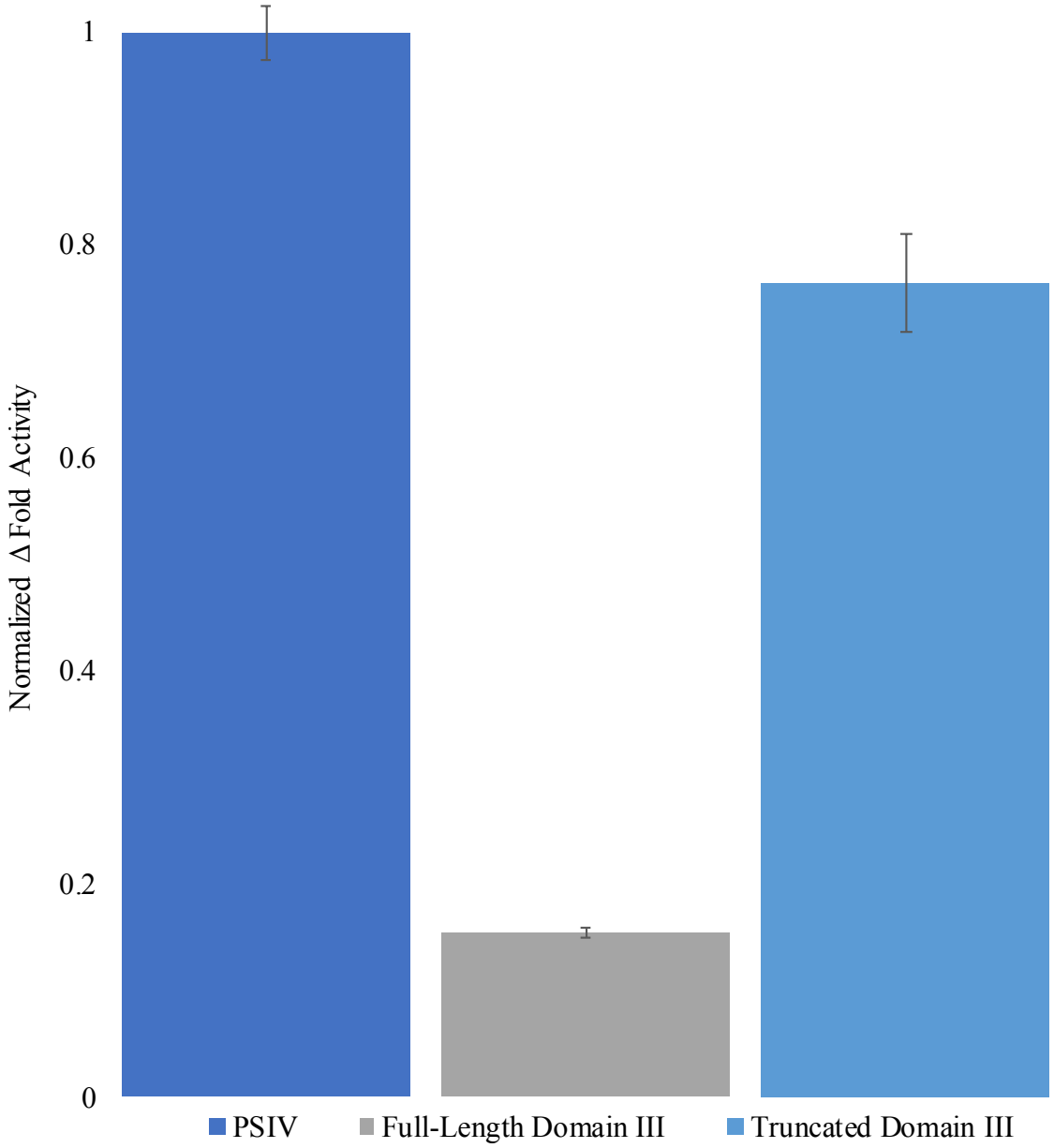

Figure 11. Normalized fold change of luciferase expression illustrating the difference in ability of the Full-Length and Truncated Domain III elements to express FLuc relative to the SD and enhancer controlled RLuc, compared to the PSIV construct, within *E. coli*. Each experimental construct consists of a normalized average generated from a minimum of four biological and three technical replicates. The error bars represent the standard error of the average values. ANOVA ($F=55.37$, $p<0.0001$) followed by *post hoc* Tukey test showed that IRES activity of each construct differed significantly from each of other constructs ($p=0.0062$ and $p<0.0001$(x2)).

Importantly, while the sequence of the Full-Length Domain III we had created matches relative to the sequence of the Full-Length Domain III published by Colussi *et al.*, we were unable to replicate the normalized luciferase expression results put forth in the paper by Colussi *et al* (2015). The normalized luciferase data of the Full-Length Domain III we expressed was 16% of that of the PSIV IGR IRES, which is considerably lower than what they had reported. Although, this finding is in agreement with Colussi *et al.,* whereby the independent Domain III is capable of initiating prokaryotic translation, yet at a significantly lower level than the entire PSIV IGR IRES.

2.3.3. Mutating Complementary Regions

As the other IRES elements were unable to effectively initiate prokaryotic translation and the structural integrity of Domain III appears less important than originally thought, we attempted to further elucidate the mode by which this IRES initiates prokaryotic translation. We considered whether complementarity between Domain III and the 16S rRNA plays a role in the ability of Domain III to mediate bacterial translation initiation. In eukaryotes, some IRES elements initiate translation through ribosomal recruitment via base-pair complementarity with the 18S rRNA, such as with HCV and the insulin-like growth factor 1 (IGF1R) (Martinez-Salas *et al.*, 2018). In bacteria, translation initiation is primarily achieved through the recruitment of ribosomal subunits to the mRNA via base pairing to the 16S rRNA, such as with the SD and anti-SD interaction (Shine and Dalgarno, 1974; Zhu *et al.*, 2011). Although, non-SD base-pairing interactions have also been reported to initiate prokaryotic translation, such as with the epsilon sequence (Olins and Rangwala, 1989; Golshani *et al.*, 1997, 2004; O'Connor, 2001).

Two regions of interest were identified within the PSIV Domain III, designated Box 1 and Box 2, that possess complementarity to 16S rRNA regions, nucleotides 932-938 and nucleotides 1465-1471, respectively (**Figure 12**). These regions of complementarity present on the 16S rRNA, Anti-Box 1 and Anti-Box 2, are plausible for two reasons: 1) their location within the subunit, and 2) the function associated with their location within the subunit. The nucleotides of Anti-Box 1 are a segment of the 3' minor domain and portion of the penultimate stem, which has been identified to contribute to the decoding process and translational fidelity (Allen & Noller, 1991; Firpo & Dahlberg, 1998). The nucleotides of Anti-Box 2 are a segment the 5' major domain and part of helix 28, which constitutes the neck of the small ribosomal subunit and creates a portion of the ribosomal P-site. Helix 28 has been observed to contact the SD helix during the SD-anti-SD interaction (Gualerzi and Pon, 2015). Thus, these complimentary regions on the 16SrRNA are located within vicinity of the decoding region of the subunit, and are in spatial proximity of where Domain III (and the corresponding boxes) interact within the ribosome and allow for Watson-Crick base-pairing to occur.

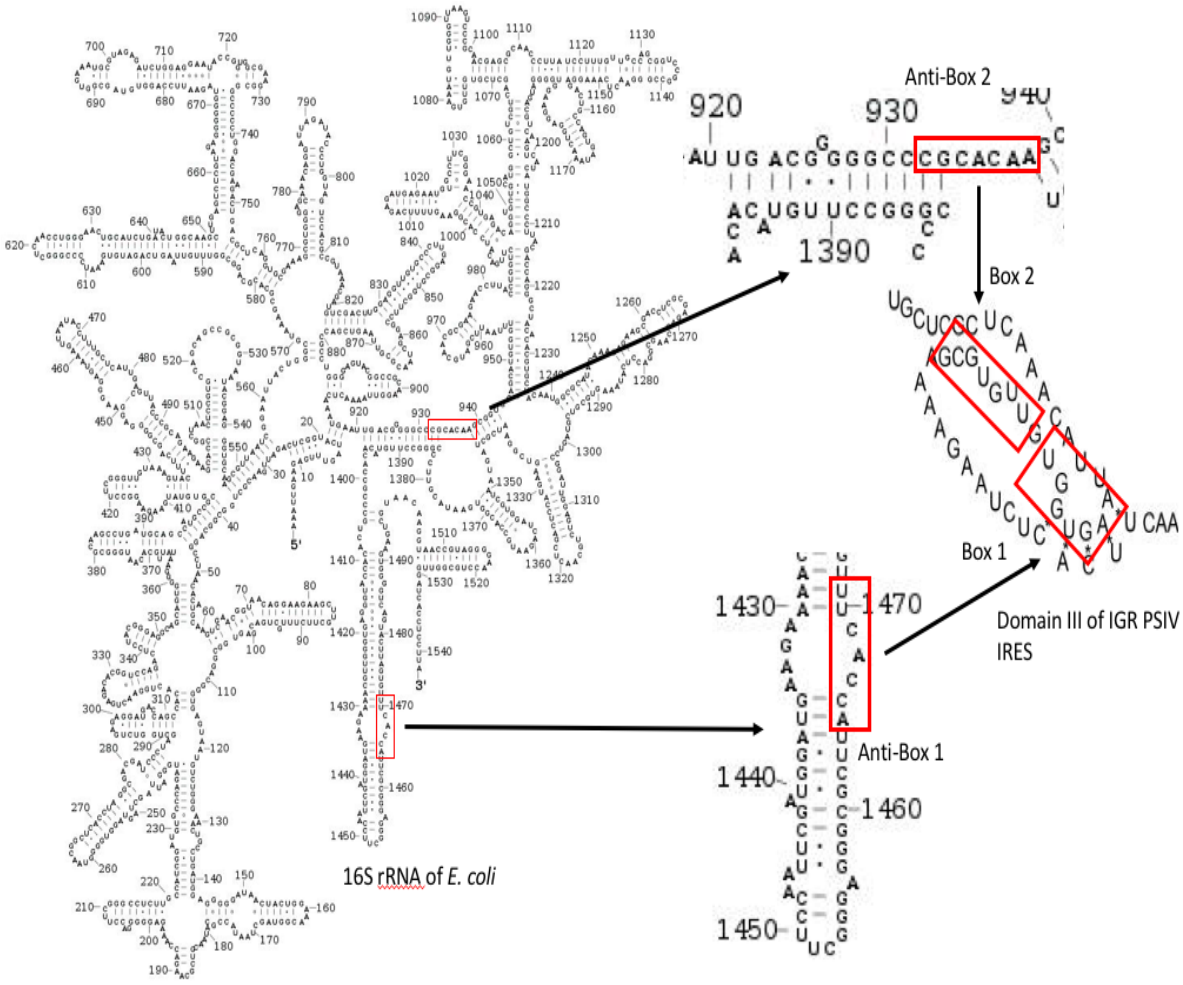

Figure 12. Complementarity illustrated between nucleotides 932-938 and 1465-1471 of the 16S rRNA of *E. coli* and Domain III of the PSIV IGR IRES.

Should complementarity to the 16S rRNA contribute to the ability of Domain III to mediate bacterial translation initiation, then destroying that complementarity would abrogate the increased initiation efficiency presented by the Truncated Domain III. As anticipated, mutational analysis of the independent Truncated Domain III illustrate that the introduced mutations (designated CACT and ACTT) (**Figure 13**) destroying complementarity to the 16S subunit within the Boxes reduce the ability of the mutant IRES elements to express FLuc (**Figure 14)**. These findings suggest that the complementary regions may assume some responsibility in the ability of the PSIV IGR IRES to mediate prokaryotic translation.

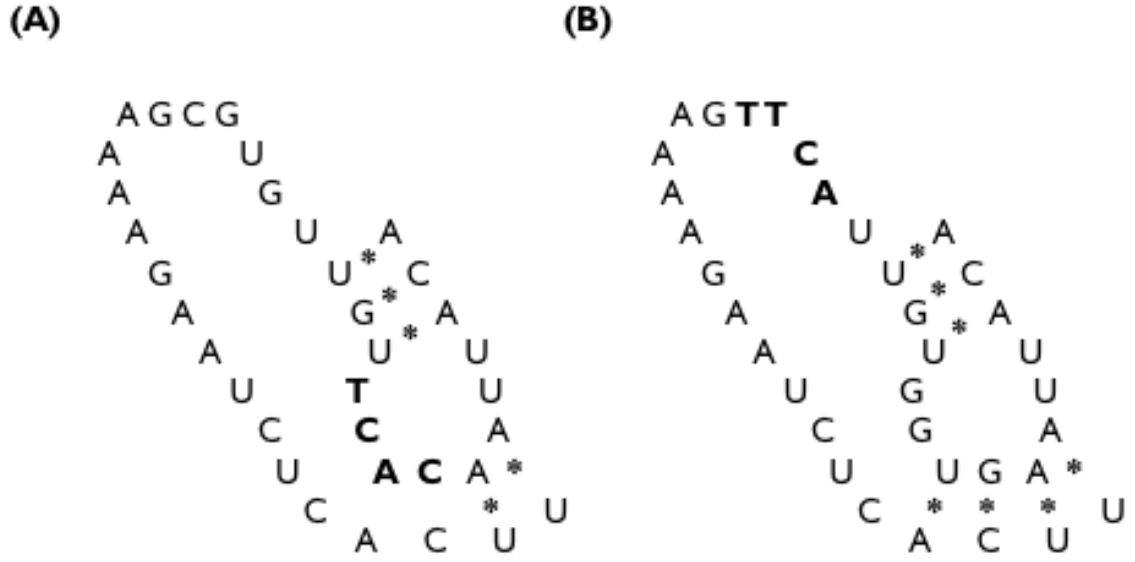

Figure 13. (**A**) Visual representation of the Truncated Domain III bearing the introduced mutations CACT within Box I. (**B**) Visual representation of the Truncated Domain III bearing the introduced mutations ACTT within Box 2.

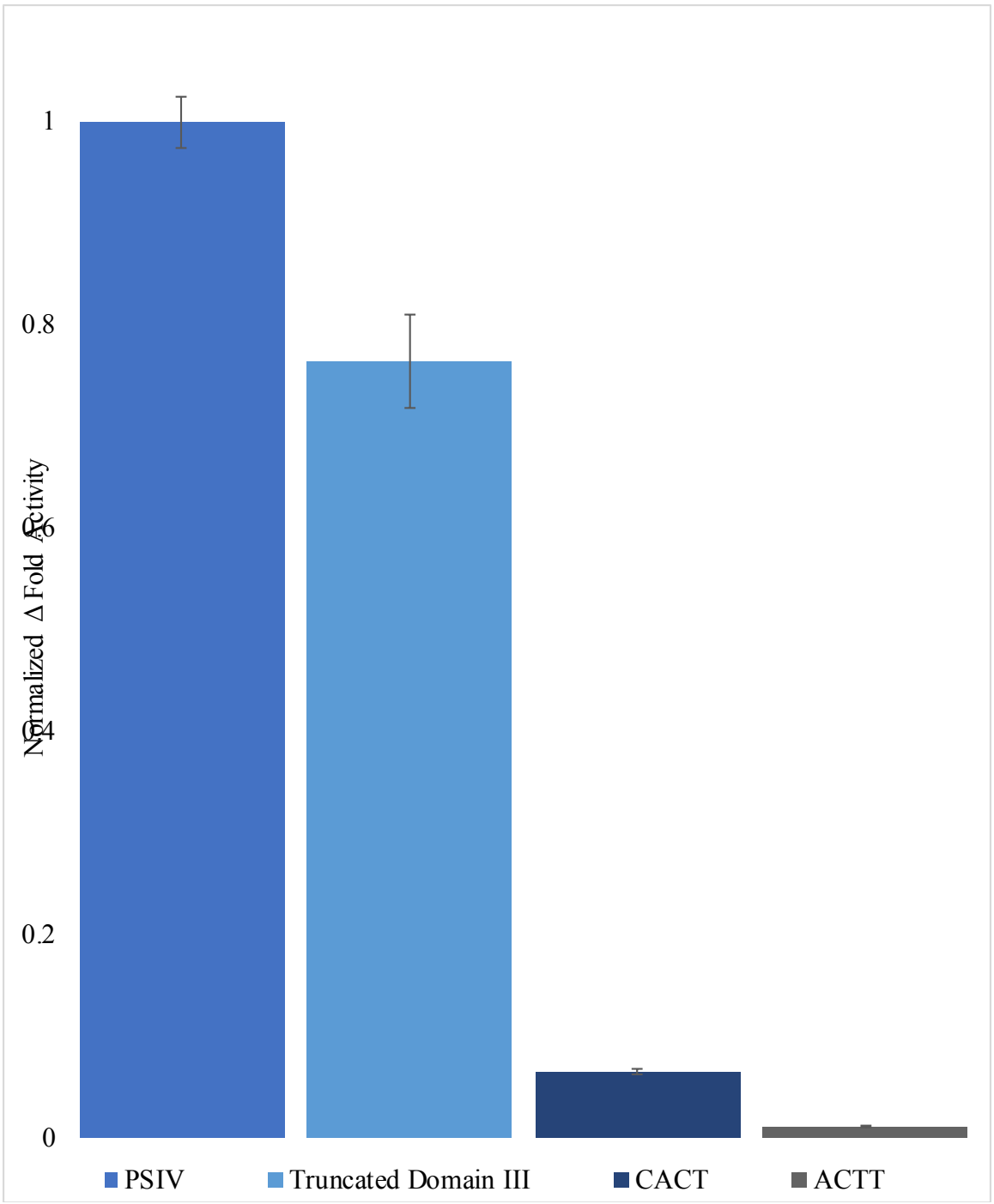

Figure 14. Normalized fold change of luciferase expression illustrating the difference in ability of mutated Truncated Domain III elements, CACT and ACTT, compared to the Truncated Domain III, to express FLuc relative to the SD and enhancer controlled RLuc, within *E. coli*. Each experimental construct consists of a normalized average generated from a minimum of four biological and three technical replicates. The error bars represent the standard error of the average values. ANOVA ($F=113.8$, $p<0.0001$) followed by *post hoc* Tukey test showed that IRES activity of each construct differed significantly from each of other constructs ($p=0.0026$ and $p<0.0001$(x3)), with exception of CACT compared to ACTT, which was not-significant ($p=0.8931$).

2.4. Conclusion

While the fundamental goals of translation initiation are the same between all cells of life, it is the most phylogenetically diverse step within the process of translation. This is highlighted by the minor correspondence between the mechanisms and apparatus utilized. Until recently, there has been no evidence of a molecular mechanism that can initiate translation in both prokaryotes and eukaryotes. In 2015, Colussi *et al.* reported that the eukaryotic PSIV IGR IRES can successfully initiate translation in the prokaryotic organism *E. coli*. IRES is a complex RNA secondary structure present in eukaryotic viruses and cellular mRNAs that can directly recruit ribosomes within the vicinity of the start codon. The findings of Colussi *et al.* indicate that this structured RNA molecule, that was believed to work exclusively in eukaryotes, can effectively circumvent domain-specific divergences. Their work states that this IRES can capitalize on conserved ribosomal features and initiate protein synthesis in both bacteria and eukaryotes. While an interesting discovery, it was not yet established if this was a unique property to the PSIV IGR IRES, or whether other IRES elements were also capable of initiating prokaryotic translation. With the results presented here, it appears that this IRES, relative to the other IRES elements employed in this study, is unique in its ability to initiate prokaryotic translation, as the other IRES elements were unable to initiate prokaryotic translation as efficiently as the PSIV IGR IRES. Our results also suggest the possibility that the structural integrity requirements for this IRES in prokaryotes differ from its functioning in eukaryotic organisms. Lastly, the break-down of complementary interactions identified by this study may imply the importance of base-pairing between Domain III and the 16S rRNA for this IRES ability to initiate prokaryotic translation.

3. General Discussion

3.1. Capabilities of Eukaryotic IRESs to Initiate Prokaryotic Translation

Our protein expression data generated using the various eukaryotic IRES and structured RNA elements suggests that the ability of the PSIV IGR IRES to initiate prokaryotic translation is a distinct property that is not common between the other IRES elements employed in this study (**Figure 9**). The cellular IRES elements, such as those from *NCE102*, *HAP4* and *FLO8*, show the highest levels of IRES activity via normalized luciferase expression, at 25%, 23% and 12% of that produced by PSIV, respectively. While not negligible results, these values are not significant enough to provide definitive evidence that these IRES elements can drive initiation of prokaryotic translation. At the lower end, other IRES elements, such as those from *GPR1,* FMDV and CrPV, exhibited negligible results, at just 1%, 0.2% and 1% of that produced by PSIV. Taken into consideration with the inconsequential expression data generated from the insertion of highly structured RNA elements, all approximately 2% of PSIV expression, indicate that prokaryotic initiation is not strictly dependent on just any complex secondary structure attempting to sequester ribosomal subunits. This is not an entirely surprising finding based on the known (and other potentially unknown) functional requirements of these elements within their eukaryotic host (Filbin & Kieft, 2009; Kieft, 2008; Mailliot & Martin, 2018). Consequently, this statement has limited significance. As this is fundamental and novel work, it is not currently known whether IRES elements, when forced to function within a different type of host (from eukaryotic to prokaryotic), require the same types and quantities of factors to function optimally or are distinctly altered. For the purpose of this discussion, the presumption is that they remain the same.

The relatively low level of functionality exhibited by the yeast cellular IRESs used in this thesis, *NCE102*, *HAP4, GPR1* and *FLO8* (**Figure 6**), were to be expected. This primarily corresponds to the lack of knowledge surrounding their functionality within their native host, such as the specifics of their secondary structure, ITAF and canonical initiation factor requirements, and recruitment mechanisms. This background knowledge is necessary to determine the likelihood that any of the IRESs selected could conceivably function in a prokaryotic organism. Thus, until more data regarding the requirements of these IRES elements comes to light, I am unable to provide any educated suppositions currently.

Contrary to that of the cellular IRESs, the specifics regarding the structural architecture and requirements of the non-PSIV viral IRESs used, CSFV, FMDV and CrPV (**Figure 7**) are reasonably well elucidated (Pilipenko *et al.*, 2000; Pestova and Hellen, 2003; Pestova *et al.*, 2008). The IRES element of CSFV was already determined by Colussi *et al.* (2015), and confirmed with this project, to be unable to initiate prokaryotic translation. As a member of the Class 3 IRESs, the CSFV IRES can directly and stably bind to the 40S subunit and initiate eukaryotic translation using only a limited subset of canonical initiation factors, as well as Met-tRNAi (Pestova *et al.*, 2001). While this IRES does not require the use of any ITAFs, it does require the addition of eIF2, GTP and initiator-tRNA to properly assemble into a 48S complex at the initiation codon. This interaction is promoted by the requisite eIF3, which then further necessitated for the formation of the 80S complex on the IRES (Pestova *et al.*, 1998; Sizova *et al.*, 1998; Mailliot and Martin, 2018; Jaafar and Kieft, 2019). As the 40S and 30S subunits possess conserved areas of binding between the two, the problem likely does not lie in the binding

aspect of the mechanism (Spahn *et al.*, 2001). Where the issue might lie is the lack of requisite factors to assemble an elongation competent 70S ribosome. While eIF2, GTP and initiator tRNA all either exist in bacteria or possess a bacterial equivalent, eIF3 or any respective counterpart is absent (Kyrpides and Woese, 1998; Roll-Mecak *et al.*, 2000; Wilson and Cate, 2012). Thus, being an indispensable part of the mechanism, the lack of eIF3 could explain the inability of the IRES to initiate prokaryotic translation. It is possible that the IRES can bind to the 30S subunit but is unable to further assemble into a complete ribosomal complex.

Similarly, based on its functional requirements in eukaryotes, the FMDV IRES was an unlikely candidate to initiate prokaryotic translation. The FMDV IRES was determined to be an unlikely candidate partially based on its more extended structure, which makes it unable to directly bind to the 40S ribosome in eukaryotes, but also based on its ITAF and canonical initiation factor requirements. The FMDV IRES requires various ITAFS, such as PTB, ITAF45 and PAG2G4, as well as several canonical initiation factors, including eIF2, eIF3, eIF4A and the C-terminus of eIF4G, to properly form the 48S complex in its eukaryotic host (Andreev *et al.*, 2007; Komar, Mazumder and Merrick, 2012; Mailliot and Martin, 2018; Martinez-Salas *et al.*, 2018). The majority of these requirements do not possess a bacterial counterpart and thus, the lack of resources available to the IRES within *E. coli* make it unlikely that the IRES could function properly.

Mildly surprising of a result was the inability of the PSIV IGR IRES' conserved relative, the CrPV IGR IRES, to initiate prokaryotic translation. The *Dicitroviridae* IGR IRESs, like PSIV and CrPV, are highly conserved relative to one another. These elements fold into similar compact architectures and maintain the same manipulative mechanism of

eukaryotic ribosome recruitment. These IRESs can directly and independently bind to eukaryotic ribosomal subunits and initiate translation (Mailliot and Martin, 2018; Martinez-Salas *et al.*, 2018). Along with the fact that the phylogenetically conserved core of the ribosome includes the tRNA binding sites, the interactions of the two IRESs should be similar between the prokaryotic and eukaryotic ribosomes (Spahn *et al.*, 2001; Zhu *et al.*, 2011; Melnikov *et al.*, 2012). Corresponding to this knowledge, it has been previously reported that the independent Domain III of the CrPV IGR IRES is in fact capable of binding to the 70S ribosomal subunit, just as the PSIV IGR IRES (Zhu *et al.*, 2011). It was identified by Zhu *et al.* that the PKI of Domain III in both IRESs act as a tRNA-mRNA mimic within the ribosomal P-site, with contacts established similarly in the 30S subunit like those observed within the 40S subunits. So, both PSIV and CrPV IGR IRESs possess similar architectures and both Domain IIIs can independently bind the 30S ribosomes, but the PSIV IGR IRES can initiate prokaryotic translation and the CrPV IGR IRES cannot. This indicates that there are additional factors responsible for the functionality of one IRES vs. the other that have not yet been addressed and cannot be easily resolved with basic commentary on the structural aspects of the IRES elements.

A potential factor playing a role in the lack of IRES functioning, other than protein requirements, could involve external or environmental influences. For example, many cellular IRESs are "activated" or predominantly utilized under a variety of (patho)physiological stress conditions, such as nutrient deprivation, hypoxia, oxidative stress and mitosis, when cap-dependent translation is limited or inhibited (Spriggs *et al.*, 2008; Komar and Hatzoglou, 2011a; Godet *et al.*, 2019). Even viruses, such as FMDV and PV, inhibit canonical translation initiation by encoding proteases that cleave eIF4G to

increase ribosomal affinity of the viral IRES and decrease competition (Belsham, McInerney and Ross-Smith, 2000; Mailliot and Martin, 2018). As such, it is possible that the appropriate conditions surrounding the IRES elements in question have not been met as to activate their proper functioning within the prokaryotic host.

An additional distinction, albeit minor, regards the use of the AUG initiation codon by the PSIV IGR IRES. The PSIV IGR IRES is known to initiate at a non-AUG start codon, CAA encoding for glutamine, directly at the IRES in eukaryotes (Sasaki and Nakashima, 2000). Contrary, the PSIV IGR IRES initiates prokaryotic translation at a canonical AUG start codon that was separated from the IRES by a spacer region. Prokaryotes have been known to utilize non-AUG codons to initiate translation previously, such as GUG, UUG and AUA; with the midmost U being conserved and the 3'G deemed as canonical (Gualerzi and Pon, 2015). That being said, neither triplet-codon encoding glutamine, CAA and CAG, possesses a central U, which could explain the use of a canonical start codon. It does lead to the speculation of whether another initiation codon could be recognized and utilized to any degree, whether by this IRES or others.

3.2. Truncation of Independent Domain III

Colussi *et al.* had indicated that the structural characteristics of the PSIV IGR IRES element enabled it to initiate prokaryotic translation, and that its structural integrity is paramount to its functionality. To assess whether structural integrity has a fundamental purpose in the functionality of PSIV IGR IRES, focus was placed on the Domain III of the IRES, following previous evidence that it is able to independently bind to the decoding

groove of the 16S rRNA, as well as initiate prokaryotic translation (Zhu *et al.*, 2011; Colussi *et al.*, 2015).

Based on the normalized expression data generated from the Truncated Domain III relative to the Full-Length Domain III, it appears that the 3-D structure and structural integrity of the IRES is less significant in prokaryotes than eukaryotes. The disruption of the independent Domain III involved its truncation by the deletion of few upstream nucleotides (**Figure 10**). This subsequently destroyed the PKI of Domain III, reducing its 3-D structure and extending its 2-D structure. Theoretically, this disruption should have rendered the Truncated Domain III non-functional, as the PKI was of highlighted importance due to its apparent ability to mimic a mRNA-tRNA interaction within the ribosomal P-site, and this interaction would consequently be spatially modified. The resulting expression data indicates otherwise, with the expression values from the Truncated Domain III considerably increased, at 76%, from the 16% exhibited by the Full-Length Domain III (**Figure 11**). Should the structural integrity have played a commanding role, deletion of any nucleotides from the Domain III structure would result in expression approaching zero percent.

Importantly, this truncation was already on top the removal of the first two domains (which house the stem loops and PKII and III, which were cited as being important), thus the structural integrity of the IRES was completely abolished and it would be assumed the luciferase data would reflect this. The expression data actually indicates otherwise based on the increased ratio relative to the Full-Length Domain III.

3.3. Induced Mutations of Complementary Regions

Luciferase expression data presented of the mutated Truncated Domain III elements illustrate that mutations destroying complementarity to the 16S subunit within the Boxes (designated CACT and ACTT) (**Figure 12** and **Figure 13**) reduce the ability of the mutant IRES elements to initiate translation (**Figure 14)**. These findings suggest that the identified complementary regions, Box 1 and Box 2, may assume some responsibility in the ability of the PSIV IGR IRES to mediate prokaryotic translation.

Box 1 and Box 2 are regions present in the PSIV Domain III that possesses complementarity to 16S rRNA regions, nucleotides 932-938 and nucleotides 1465-1471, respectively (**Figure 12**). These regions of complementarity present on the 16S rRNA, Anti-Box 1 and Anti-Box 2, are plausible for two reasons: 1) their location within the subunit, and 2) the function associated with their location within the subunit. The nucleotides of Anti-Box 1 are a segment of the 3' minor domain and portion of the penultimate stem, which has been identified to contribute to the decoding process and translational fidelity (Allen & Noller, 1991; Firpo & Dahlberg, 1998). The nucleotides of Anti-Box 2 are a segment the 5' major domain and part of helix 28, which constitutes the neck of the small ribosomal subunit and creates a portion of the ribosomal P-site. Helix 28 has been observed to contact the SD helix during the SD-aSD interaction (Gualerzi and Pon, 2015). Thus, these complimentary regions on the 16S rRNA are located within vicinity of the decoding region of the subunit, and are in proximity of where Domain III (and the corresponding boxes) interact within the ribosome and can allow for Watson-Crick base-pairing to occur.

Initiation of translation via the help of base-pairing not an uncommon mechanism of action. In eukaryotes, some IRESs initiate translation through ribosomal recruitment via base-pair complementarity with the 18S rRNA, such as HCV and the cellular IRES of the insulin-like growth factor 1 (IGF1R) (Martinez-Salas *et al.*, 2018). In bacteria, translation initiation is primarily achieved through the recruitment of ribosomal subunits to the mRNA via base pairing to the 16S rRNA, such as with the SD and anti-SD interaction (Shine and Dalgarno, 1974; Zhu *et al.*, 2011). Although, non-SD base-pairing interactions have also been reported to initiate prokaryotic translation, such as with the epsilon sequence (Olins and Rangwala, 1989; Golshani *et al.*, 1997, 2004; O'Connor, 2001).

Furthermore, regarding the mechanism of action, the functionality of the PSIVIGR IRES allows it to recruit and bind the ribosome directly, both in prokaryotes and eukaryotes. Following ribosomal subunit binding in *E. coli,* the ribosome was noted to reposition to the downstream AUG of the FLuc mRNA. This is a deviation from the mechanism of the PSIV IGR IRES observed in eukaryotes. Is it possible that the repositioning of the ribosome resembles a scanning mechanism that is present (not of Class 4 IRESs) in eukaryotes? Or is it possible that the PSIV IGR IRES resembles and implements itself like a larger SD sequence to the ribosome? *Dicistroviridae* IRESs are quite small relative to other IRES elements, but are still roughly 200 nucleotides in size, which is much larger than the average SD consensus sequence. The spacer region present following the IRES element is also much longer than is common or optimal in *E. coli*. Importantly though, the identified Anti-Box 2 of the 16S rRNA contacts the SD helix during the SD-anti-SD interaction, and as PKI of Domain III is observed to interact with

the P-site of the 16S rRNA, with proposed base-pairing potentially at play, arguing that this IRES mimics an SD sequence to the ribosome is not implausible.

3.4. Conclusions

While the fundamental goals of translation initiation are the same between all cells of life, it is the most phylogenetically diverse step within the process of translation. This is highlighted by the minor correspondence between the mechanisms and apparatus utilized. Until recently, there has been no evidence of a molecular mechanism that can initiate translation in both prokaryotes and eukaryotes. In 2015, Colussi *et al*. reported that the PSIV IGR IRES can successfully initiate translation in the prokaryotic organism *E. coli*. IRES is a complex RNA secondary structure present in eukaryotic viruses and mRNAs that can directly recruit ribosomes within the vicinity of the start codon. The findings of Colussi *et al.* indicate that this structured RNA molecule, that was believed to work exclusively in eukaryotes, can effectively circumvent domain-specific divergences. Their work states that this IRES can capitalize on conserved ribosomal features and initiate protein synthesis in both bacteria and eukaryotes. While an interesting discovery, it was not yet established if this was a unique property to the PSIV IGR IRES, or whether other IRES elements were also capable of initiating prokaryotic translation. With the results presented here, it appears that this IRES, relative to the other IRES elements employed in this study, is unique in its ability to initiate prokaryotic translation, as the other IRES elements were unable to initiate prokaryotic translation as efficiently as the PSIV IGR IRES. Our results also suggest the possibility that the structural integrity requirements for this IRES in prokaryotes differ from its functioning in eukaryotic organisms. Lastly, the break-down of complementary

interactions identified by this study may imply the importance of base-pairing between Domain III and the 16S rRNA for this IRES' ability to initiate prokaryotic translation.

3.5. Future Directions

The identification of an IRES element that is functional in both prokaryotes and eukaryotes is a significant discovery. While the results of this thesis indicate that the ability of the PSIV IGR IRES to initiate prokaryotic translation is a distinct property, only a very minor subset of IRES elements were tested. Thus, in order to fully elucidate the scope of IRES functionality in a prokaryotic host, many more IRES elements (both viral and cellular) must be tested and verified. Not only that, but trials with these elements must also involve variable features, such as spacer length following the IRES element and start codon choice. Particularly, the external and environmental influences exerting effect on the host organism must be altered as to understand the specifics surrounding differential IRES function and hypothetical activation.

Mechanistically, focus should be placed towards understanding how altering the secondary structure of this IRES can subsequently affect its ability to initiate prokaryotic translation, for the whole PSIV IGR IRES, as well as further work on the specifics surrounding Domain III. More research must also be done regarding the complementary regions present within the Domain III of the PSIV IGR IRES and the influences they may have on the potential IRES mechanism. This is particularly directed towards the speculation that the PSIV IGR IRES functions similarly to an SD sequence within *E. coli*.

The statement of "ability to initiate prokaryotic translation" that was included several times within this thesis is a sizeable, yet currently unsubstantiated statement. *E.*

coli, while a model organism, is just one species of bacteria and the scope this statement refers to wider range of prokaryotic organisms, that has not yet been validated. Thus, if this IRES element does exploit the conserved portion of the ribosome based on its structural architecture, it should be able to do so in other bacterial species; where protein composition, protein functions, gene expression control, and other variables can be dissimilar relative to *E. coli*.

Further regarding the statements made surrounding this discovery, proclamations indicating that the PSIV IGR IRES' ability to function in both eukaryotes and a prokaryote is "conservation of initiation indicative of an early RNA World" is also an unsubstantiated claim that must be investigated. If there is conservation, why have we not encountered IRES elements with their corresponding mechanisms in bacterial mRNAs previously? And why do bacteriophages not use a similar mechanism to infect host cells if it is a functional mechanism, and one that is clearly popular and efficient, as evidenced by eukaryotic virus utilization? It is important to consider that forcing an element or mechanism to function in a non-native host is great in the pursuit of biotechnological advancement but yields nothing organic towards understanding native functionality of the organisms, mechanisms or structure in question.

The work presented within this thesis, and the work done by Colussi *et al*. (2015) represents just an infinitesimal fragment of the knowledge that is yet to be discovered regarding IRES function. And as with all fundamental work, significant trial and error, and exploration must be done to understand the greater picture.

4. References

Allen, G. S. and Frank, J. (2007) 'Structural insights on the translation initiation complex: Ghosts of a universal initiation complex', *Molecular Microbiology*, 63(4), pp. 941–950. doi: 10.1111/j.1365-2958.2006.05574.x.

Allen, P. N. and Noller, H. F. (1991) 'A single base substitution in 16S ribosomal RNA suppresses streptomycin dependence and increases the frequency of translational errors', *Cell*. Elsevier, 66(1), pp. 141–148. doi: 10.1016/0092-8674(91)90146-P.

Andreev, D. E. *et al.* (2007) 'Differential factor requirement to assemble translation initiation complexes at the alternative start codons of foot-and-mouth disease virus RNA', *RNA*. Cold Spring Harbor Laboratory Press, 13(8), pp. 1366–1374. doi: 10.1261/rna.469707.

Belsham, G. J., McInerney, G. M. and Ross-Smith, N. (2000) 'Foot-and-Mouth Disease Virus 3C Protease Induces Cleavage of Translation Initiation Factors eIF4A and eIF4G within Infected Cells', *Journal of Virology*, 74(1), pp. 272–280. doi: 10.1128/jvi.74.1.272-280.2000.

Ben-Shem, A. *et al.* (2011) 'The structure of the eukaryotic ribosome at 3.0 Å resolution', *Science*, 334(6062), pp. 1524–1529. doi: 10.1126/science.1212642.

Benelli, D. and Londei, P. (2011) 'Translation initiation in Archaea: conserved and domain-specific features', *Biochemical Society Transactions*, 39(1), pp. 89–93. doi: 10.1042/bst0390089.

Bergenhem, N. C. H. *et al.* (1992) 'Mutation creates an open reading frame within the 5′ untranslated region of macaque erythrocyte carbonic anhydrase (CA) I mRNA that suppresses CA I expression and supports the scanning model for translation', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 89(18), pp. 8798–8802. doi: 10.1073/pnas.89.18.8798.

De Breyne, S. *et al.* (2009) 'Direct functional interaction of initiation factor eIF4G with type 1 internal ribosomal entry sites', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 106(23), pp. 9197–9202. doi: 10.1073/pnas.0900153106.

Coldwell, M. J. *et al.* (2000) 'Initiation of Apaf-1 translation by internal ribosome entry', *Oncogene*, 19(7), pp. 899–905. doi: 10.1038/sj.onc.1203407.

Colussi, T. M. *et al.* (2015) 'Initiation of translation in bacteria by a structured eukaryotic IRES RNA', *Nature*, 519(7541), pp. 110–113. doi: 10.1038/nature14219.

Filbin, M. E. and Kieft, J. S. (2009) 'Toward a structural understanding of IRES RNA function', *Current Opinion in Structural Biology*. NIH Public Access, 19(3), pp. 267–276. doi: 10.1016/j.sbi.2009.03.005.

Firpo, M. A. and Dahlberg, A. E. (1998) 'The importance of base pairing in the penultimate stem of Escherichia coli 16S rRNA for ribosomal subunit association', *Nucleic Acids Research*, 26(9), pp. 2156–2160. doi: 10.1093/nar/26.9.2156.

Fletcher, S. P. and Jackson, R. J. (2002) 'Pestivirus Internal Ribosome Entry Site (IRES) Structure and Function: Elements in the 5' Untranslated Region Important for IRES

Function', *Journal of Virology*, 76(10), pp. 5024–5033. doi: 10.1128/jvi.76.10.5024-5033.2002.

Gilbert, W. V. *et al.* (2007) 'Cap-independent translation is required for starvation-induced differentiation in yeast', *Science*, 317(5842), pp. 1224–1227. doi: 10.1126/science.1144467.

Godet, A. C. *et al.* (2019) 'IRES Trans-Acting Factors, Key Actors of the Stress Response', *International journal of molecular sciences*. Multidisciplinary Digital Publishing Institute, 20(4), p. 924. doi: 10.3390/ijms20040924.

Golshani, A. *et al.* (1997) 'Does the epsilon sequence of phage T7 function as an initiator for the translation of CAT mRNA in Escherichia coli?', *Biochemical and Biophysical Research Communications*, 236(2), pp. 253–256. doi: 10.1006/bbrc.1997.6842.

Golshani, A *et al.* (2000) 'Enhancing activity of epsilon in Escherichia coli and Agrobacterium tumefaciens cells.', *Biochemical and biophysical research communications*, 269(2), pp. 508–12. doi: 10.1006/bbrc.2000.2327.

Golshani, Ashkan *et al.* (2000) 'Epsilon as an initiator of translation of CAT mRNA in Escherichia coli', *Biochemical and Biophysical Research Communications*. Academic Press Inc., 273(2), pp. 528–531. doi: 10.1006/bbrc.2000.2958.

Golshani, A. *et al.* (2004) 'Escherichia coli mRNAs with strong Shine/Dalgarno sequences also contain 5′ end sequences complementary to domain # 17 on the 16S ribosomal RNA', *Biochemical and Biophysical Research Communications*. Academic Press, 316(4), pp. 978–983. doi: 10.1016/j.bbrc.2004.02.169.

Gray, N. K. and Wickens, M. (1998) 'CONTROL OF TRANSLATION INITIATION IN ANIMALS', *Annual Review of Cell and Developmental Biology*, 14(1), pp. 399–458. doi: 10.1146/annurev.cellbio.14.1.399.

Gualerzi, C. O. and Pon, C. L. (2015) 'Initiation of mRNA translation in bacteria: Structural and dynamic aspects', *Cellular and Molecular Life Sciences*. Springer, 72(22), pp. 4341–4367. doi: 10.1007/s00018-015-2010-3.

Hertz, M. I. *et al.* (2013) 'Ribosomal Protein S25 Dependency Reveals a Common Mechanism for Diverse Internal Ribosome Entry Sites and Ribosome Shunting', *Molecular and Cellular Biology*, 33(5), pp. 1016–1026. doi: 10.1128/mcb.00879-12.

Hofacker, I. L. (2003) 'Vienna RNA secondary structure server', *Nucleic Acids Research*, 31(13), pp. 3429–3431. doi: 10.1093/nar/gkg599.

Inoue, H., Nojima, H. and Okayama, H. (1990) 'High efficiency transformation of Escherichia coli with plasmids', *Gene*. Elsevier, 96(1), pp. 23–28. doi: 10.1016/0378-1119(90)90336-P.

Jaafar, Z. A. and Kieft, J. S. (2019) 'Viral RNA structure-based strategies to manipulate translation', *Nature Reviews Microbiology*. NIH Public Access, 17(2), pp. 110–123. doi: 10.1038/s41579-018-0117-x.

Jackson, R. J. (2005) 'Alternative mechanisms of initiating translation of mammalian mRNAs', *Biochemical Society Transactions*, 33(6), pp. 1231–1241. doi: 10.1042/BST20051231.

Jackson, R. J. (2013) 'The current status of vertebrate cellular mRNA IRESs', *Cold Spring Harbor Perspectives in Biology*, 5(2). doi: 10.1101/cshperspect.a011569.

Jackson, R. J., Hellen, C. U. T. and Pestova, T. V. (2010) 'The mechanism of eukaryotic translation initiation and principles of its regulation', *Nature Reviews Molecular Cell Biology*. Nature Publishing Group, 11(2), pp. 113–127. doi: 10.1038/nrm2838.

Jan, E., Kinzy, T. G. and Sarnow, P. (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*. National Academy of Sciences, 100(26), pp. 15410–15415. doi: 10.1073/pnas.2535183100.

Jenner, L. *et al.* (2012) 'Crystal structure of the 80S yeast ribosome', *Current Opinion in Structural Biology*, 22(6), pp. 759–767. doi: 10.1016/j.sbi.2012.07.013.

Johnson, A. G. *et al.* (2017) 'Dynamics of IRES-mediated translation', *Philosophical Transactions of the Royal Society B: Biological Sciences*. The Royal Society, 372(1716). doi: 10.1098/rstb.2016.0177.

Kaminishi, T. *et al.* (2007) 'A Snapshot of the 30S Ribosomal Subunit Capturing mRNA via the Shine-Dalgarno Interaction', *Structure*. Cell Press, 15(3), pp. 289–297. doi: 10.1016/j.str.2006.12.008.

Kapp, L. D. and Lorsch, J. R. (2004) 'The Molecular Mechanics of Eukaryotic Translation', *Annual Review of Biochemistry*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 73(1), pp. 657–704. doi: 10.1146/annurev.biochem.73.030403.080419.

Kieft, J. S. *et al.* (2001) 'Mechanisms of internal ribosome entry in translation initiation', in *Cold Spring Harbor Symposia on Quantitative Biology*, pp. 277–283.

Kieft, J. S. (2008) 'Viral IRES RNA structures and ribosome interactions', *Trends in Biochemical Sciences*. NIH Public Access, 33(6), pp. 274–283. doi: 10.1016/j.tibs.2008.04.007.

Klinge, S. *et al.* (2011) 'Crystal structure of the eukaryotic 60S ribosomal subunit in complex with initiation factor 6', *Science*, 334(6058), pp. 941–948. doi: 10.1126/science.1211204.

Komar, A. A. and Hatzoglou, M. (2011a) 'Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states', *Cell Cycle*. Taylor & Francis, 10(2), pp. 229–240. doi: 10.4161/cc.10.2.14472.

Komar, A. A. and Hatzoglou, M. (2011b) 'Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states', *Cell Cycle*, 10(2), pp. 229–240. doi: 10.4161/cc.10.2.14472.

Komar, A. A., Mazumder, B. and Merrick, W. C. (2012) 'A new framework for understanding IRES-mediated translation', *Gene*. NIH Public Access, 502(2), pp. 75–86. doi: 10.1016/j.gene.2012.04.039.

Kozak, M. (1987a) 'An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs', *Nucleic Acids Research*, 15(20), pp. 8125–8148. doi: 10.1093/nar/15.20.8125.

Kozak, M. (1987b) 'Effects of intercistronic length on the efficiency of reinitiation by eucaryotic ribosomes.', *Molecular and Cellular Biology*. American Society for Microbiology Journals, 7(10), pp. 3438–3445. doi: 10.1128/mcb.7.10.3438.

Kozak, M. (1991a) 'A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes.', *Gene expression*, 1(2), pp. 111–115. doi: 10.1083/jcb.115.4.887.

Kozak, M. (1991b) 'A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes.', *Gene expression*. Cognizant Communication Corporation, 1(2), pp. 111–115. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1820208 (Accessed: 21 July 2019).

Kozak, M. (2002) 'Pushing the limits of the scanning mechanism for initiation of translation', *Gene*, pp. 1–34. doi: 10.1016/S0378-1119(02)01056-9.

Kwak, K. *et al.* (2014) 'Impact of inhibitors and L2 antibodies upon the infectivity of diverse alpha and beta human papillomavirus types', *PLoS ONE*. Edited by Z.-M. Zheng, 9(5), p. e97232. doi: 10.1371/journal.pone.0097232.

Kyrpides, N. C. and Woese, C. R. (1998) 'Universally conserved translation initiation factors', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 95(1), pp. 224–228. doi: 10.1073/pnas.95.1.224.

de la Cruz, J., Karbstein, K. and Woolford, J. L. (2015) 'Functions of Ribosomal Proteins in Assembly of Eukaryotic Ribosomes In Vivo', *Annual Review of Biochemistry*. Annual Reviews, 84(1), pp. 93–129. doi: 10.1146/annurev-biochem-060614-033917.

Lafontaine, D. L. J. and Tollervey, D. (2001) 'The function and synthesis of ribosomes', *Nature Reviews Molecular Cell Biology*, pp. 514–520. doi: 10.1038/35080045.

Lang, K. J. D., Kappel, A. and Goodall, G. J. (2002) 'Hypoxia-inducible factor-1α mRNA contains an internal ribosome entry site that allows efficient translation during normoxia and hypoxia', *Molecular Biology of the Cell*, 13(5), pp. 1792–1801. doi: 10.1091/mbc.02-02-0017.

Laursen, B. S. *et al.* (2005) 'Initiation of Protein Synthesis in Bacteria', *Microbiology and Molecular Biology Reviews*, 69(1), pp. 101–123. doi: 10.1128/mmbr.69.1.101-123.2005.

López-Lastra, M. *et al.* (2010) 'Translation initiation of viral mRNAs', *Reviews in Medical Virology*, 20(3), pp. 177–195. doi: 10.1002/rmv.649.

López-Lastra, M., Rivas, A. and Barría, M. I. (2005) 'Protein synthesis in eukaryotes: The growing biological relevance of cap-independent translation initiation', *Biological Research*, pp. 121–146. doi: 10.4067/S0716-97602005000200003.

Lozano, G. and Martínez-Salas, E. (2015) 'Structural insights into viral IRES-dependent translation mechanisms', *Current Opinion in Virology*, 12, pp. 113–120. doi: 10.1016/j.coviro.2015.04.008.

Mailliot, J. and Martin, F. (2018) 'Viral internal ribosomal entry sites: four classes for one goal', *Wiley Interdisciplinary Reviews: RNA*, 9(2). doi: 10.1002/wrna.1458.

Martinez-Salas, E. *et al.* (2018) 'Insights into structural and mechanistic features of viral

IRES elements', *Frontiers in Microbiology*. Frontiers, 8(JAN), p. 2629. doi: 10.3389/fmicb.2017.02629.

Melnikov, S. *et al.* (2012) 'One core, two shells: Bacterial and eukaryotic ribosomes', *Nature Structural and Molecular Biology*, 19(6), pp. 560–567. doi: 10.1038/nsmb.2313.

Mével-Ninio, M. *et al.* (1996) 'The three dominant female-sterile mutations of the Drosophila ovo gene are point mutations that create new translation-initiator AUG codons.', *Development (Cambridge, England)*, 122(12), pp. 4131–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9012532 (Accessed: 27 July 2019).

Milón, P. and Rodnina, M. V. (2012) 'Kinetic control of translation initiation in bacteria', *Critical Reviews in Biochemistry and Molecular Biology*, pp. 334–348. doi: 10.3109/10409238.2012.678284.

Moteshareie, H. *et al.* (2018) 'Heavy metal sensitivities of gene deletion strains for ITT1 and RPS1A connect their activities to the expression of URE2, a key gene involved in metal detoxification in yeast', *PLoS ONE*, 13(9), pp. 1–18. doi: 10.1371/journal.pone.0198704.

O'Connor, M. (2001) 'Enhancement of translation by the epsilon element is independent of the sequence of the 460 region of 16S rRNA', *Nucleic Acids Research*. Narnia, 29(7), pp. 1420–1425. doi: 10.1093/nar/29.7.1420.

Olins, P. O. and Rangwala, S. H. (1989) 'A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the lacZ gene in Escherichia coli', *Journal of Biological Chemistry*, 264(29), pp. 16973–16976. Available at: http://www.jbc.org/content/264/29/16973.full.pdf (Accessed: 27 July 2019).

Pelletier, J. and Sonenberg, N. (1988) 'Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA', *Nature*, 334(6180), pp. 320–325. doi: 10.1038/334320a0.

Pestova, T. V *et al.* (1998) 'A prokaryotic-like mode of cytoplasmic eukaryotic ribosome binding to the initiation codon during internal translation initation of hepatitis C and classical swine fever virus RNAs', *Genes and Development*. Cold Spring Harbor Laboratory Press, 12(1), pp. 67–83. doi: 10.1101/gad.12.1.67.

Pestova, T. V *et al.* (2001) 'Molecular mechanisms of translation initiation in eukaryotes', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 98(13), pp. 7029–7036. doi: 10.1073/pnas.111145798.

Pestova, T. V *et al.* (2008) 'eIF2-dependent and eIF2-independent modes of initiation on the CSFV IRES: A common role of domain II', *EMBO Journal*. European Molecular Biology Organization, 27(7), pp. 1060–1072. doi: 10.1038/emboj.2008.49.

Pestova, T. V and Hellen, C. U. T. (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*. Cold Spring Harbor Laboratory Press, 17(2), pp. 181–186. doi: 10.1101/gad.1040803.

Pichon, X. *et al.* (2012) 'RNA Binding Protein/RNA Element Interactions and the Control of Translation', *Current Protein & Peptide Science*. Bentham Science Publishers,

13(4), pp. 294–304. doi: 10.2174/138920312801619475.

Pilipenko, E. V *et al.* (2000) 'A cell cycle-dependent protein serves as a template-specific translation initiation factor.', *Genes & development*. Cold Spring Harbor Laboratory Press, 14(16), pp. 2028–45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10950867 (Accessed: 9 August 2019).

Pooggin, M. M., Hohn, T. and Fütterer, J. (2000) 'Role of a short open reading frame in ribosome shunt on the cauliflower mosaic virus RNA leader', *Journal of Biological Chemistry*. American Society for Biochemistry and Molecular Biology, 275(23), pp. 17288–17296. doi: 10.1074/jbc.M001143200.

Pooggin, M. M. and Ryabova, L. A. (2018) 'Ribosome shunting, polycistronic translation, and evasion of antiviral defenses in plant pararetroviruses and beyond', *Frontiers in Microbiology*. Frontiers, 9(APR), p. 644. doi: 10.3389/fmicb.2018.00644.

Ramakrishnan, V. (2002) 'Ribosome Structure and the Mechanism of Translation', *Cell*. Cell Press, 108(4), pp. 557–572. doi: 10.1016/S0092-8674(02)00619-0.

Ramesh, M. and Woolford, J. L. (2016) 'Eukaryote-specific rRNA expansion segments function in ribosome biogenesis', *RNA*, 22(8), pp. 1153–1162. doi: 10.1261/rna.056705.116.

Roll-Mecak, A. *et al.* (2000) 'X-ray structures of the universal translation initiation factor IF2/eIF5B: Conformational changes on GDP and GTP binding', *Cell*. Cell Press, 103(5), pp. 781–792. doi: 10.1016/S0092-8674(00)00181-1.

Ryabova, L. A., Pooggin, M. M. and Hohn, T. (2006) 'Translation reinitiation and leaky scanning in plant viruses', *Virus Research*. Elsevier, 119(1), pp. 52–62. doi: 10.1016/j.virusres.2005.10.017.

Sasaki, J. and Nakashima, N. (2000) 'Methionine-independent initiation of translation in the capsid protein of an insect RNA virus', *Proceedings of the National Academy of Sciences of the United States of America*, 97(4), pp. 1512–1515. doi: 10.1073/pnas.010426997.

Seino, A. *et al.* (2005) 'Translational control by internal ribosome entry site in Saccharomyces cerevisiae', *Biochimica et Biophysica Acta - Gene Structure and Expression*. Elsevier, 1681(2–3), pp. 166–174. doi: 10.1016/j.bbaexp.2004.11.009.

Sherrill, K. W. *et al.* (2004) 'BCL-2 translation is mediated via internal ribosome entry during cell stress', *Journal of Biological Chemistry*, 279(28), pp. 29066–29074. doi: 10.1074/jbc.M402727200.

Shi, Y. J. *et al.* (2005) 'Cyclin D1 and c-myc internal ribosome entry site (IRES)-dependent translation is regulated by AKT activity and enhanced by rapamycin through a p38 MAPK- and ERK-dependent pathway', *Journal of Biological Chemistry*, 280(12), pp. 10964–10973. doi: 10.1074/jbc.M407874200.

Shine, J. and Dalgarno, L. (1974) 'The 3'-Terminal Sequence of Escherichia coli 16S Ribosomal RNA: Complementarity to Nonsense Triplets and Ribosome Binding Sites', 71(4), pp. 1–5. Available at: https://www.pnas.org/content/pnas/71/4/1342.full.pdf (Accessed: 27 July 2019).

Shultzaberger, R. K. *et al.* (2001) 'Anatomy of Escherichia coli ribosome binding sites',

Journal of Molecular Biology, 313(1), pp. 215–228. doi: 10.1006/jmbi.2001.5040.

Simonetti, A. *et al.* (2009) 'A structural view of translation initiation in bacteria', *Cellular and Molecular Life Sciences*, 66(3), pp. 423–436. doi: 10.1007/s00018-008-8416-4.

Sizova, D. V *et al.* (1998) 'Specific interaction of eukaryotic translation initiation factor 3 with the 5' nontranslated regions of hepatitis C virus and classical swine fever virus RNAs.', *Journal of virology*. American Society for Microbiology (ASM), 72(6), pp. 4775–82. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9573242 (Accessed: 5 August 2019).

Somogyi, P. *et al.* (1993) 'Ribosomal pausing during translation of an RNA pseudoknot.', *Molecular and Cellular Biology*. American Society for Microbiology Journals, 13(11), pp. 6931–6940. doi: 10.1128/mcb.13.11.6931.

Sonenberg, N. and Hinnebusch, A. G. (2009) 'Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets', *Cell*. Cell Press, 136(4), pp. 731–745. doi: 10.1016/j.cell.2009.01.042.

Spahn, C. M. T. *et al.* (2001) 'Structure of the 80S ribosome from Saccharomyces cerevisiae - tRNA-ribosome and subunit-subunit interactions', *Cell*. Elsevier, 107(3), pp. 373–386. doi: 10.1016/S0092-8674(01)00539-6.

Spriggs, K. A. *et al.* (2008) 'Re-programming of translation following cell stress allows IRES-mediated translation to predominate', *Biology of the Cell*. John Wiley & Sons, Ltd (10.1111), 100(1), pp. 27–38. doi: 10.1042/bc20070098.

Stacey, S. N. *et al.* (2000) 'Leaky Scanning Is the Predominant Mechanism for Translation of Human Papillomavirus Type 16 E7 Oncoprotein from E6/E7 Bicistronic mRNA', *Journal of Virology*. American Society for Microbiology (ASM), 74(16), pp. 7284–7297. doi: 10.1128/jvi.74.16.7284-7297.2000.

Steitz, J. A. and Jakest, K. (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 (specificity in polypeptide chain initiation/RNA secondary structure)*, *Biochemistry*. Available at: https://www.pnas.org/content/pnas/72/12/4734.full.pdf (Accessed: 27 July 2019).

Stoneley, M. *et al.* (1998) 'C-Myc 5' untranslated region contains an internal ribosome entry segment', *Oncogene*, 16(3), pp. 423–428. doi: 10.1038/sj.onc.1201763.

Vimberg, V. *et al.* (2007) 'Translation initiation region sequence preferences in Escherichia coli', *BMC Molecular Biology*. BioMed Central, 8(1), p. 100. doi: 10.1186/1471-2199-8-100.

Wang, Q. S. and Jan, E. (2014) 'Switch from cap- to factorless IRES-dependent 0 and +1 frame translation during cellular stress and dicistrovirus infection', *PLoS ONE*. Edited by T. Preiss. Public Library of Science, 9(8), p. e103601. doi: 10.1371/journal.pone.0103601.

Wilson, D. N. and Cate, J. H. D. (2012) 'The structure and function of the eukaryotic ribosome', *Cold Spring Harbor Perspectives in Biology*. Cold Spring Harbor Laboratory Press, 4(5), p. 5. doi: 10.1101/cshperspect.a011536.

Wool, I. G. (1979) 'The Structure and Function of Eukaryotic Ribosomes', *Annual Review of Biochemistry*. Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, 48(1), pp. 719–754. doi: 10.1146/annurev.bi.48.070179.003443.

Xia, X. (2007) 'Internal ribosomal entry site lacks secondary structure', *Nature Precedings*. doi: 10.1038/npre.2007.1248.1.

Xia, X. and Holcik, M. (2009) 'Strong eukaryotic IRESs have weak secondary structure', *PLoS ONE*. Edited by T. Preiss. Public Library of Science, 4(1), p. e4136. doi: 10.1371/journal.pone.0004136.

Yusupova, G. Z. *et al.* (2001) 'The path of messenger RNA through the ribosome', *Cell*. Elsevier, 106(2), pp. 233–241. doi: 10.1016/S0092-8674(01)00435-4.

Zhu, J. *et al.* (2011) 'Crystal structures of complexes containing domains from two viral internal ribosome entry site (IRES) RNAs bound to the 70S ribosome', *Proceedings of the National Academy of Sciences of the United States of America*, 108(5), pp. 1839–1844. doi: 10.1073/pnas.1018582108.

Zinoviev, A., Hellen, C. U. T. and Pestova, T. V (2015) 'Multiple Mechanisms of Reinitiation on Bicistronic Calicivirus mRNAs', *Molecular Cell*. NIH Public Access, 57(6), pp. 1059–1073. doi: 10.1016/j.molcel.2015.01.039.

Zuker, M. and Stiegler, P. (1981) 'Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information', *Nucleic Acids Research*, 9(1), pp. 133–148. doi: 10.1093/nar/9.1.133.

5. Appendix 1

Table 2. Oligonucleotide primers used in the PCR amplification of viral and cellular IRESs and structured RNA elements for their insertion into the dual-luciferase pET30a reporter vector. The first four nucleotides in the forward and reverse primers are leader sequences. ACTAGT is the *SpeI* restriction site and CCATGG is the *NcoI* restriction site.

Construct	Primer Sequence
FMDV	F: GATC ACTAGT AGCAGGTTTCCCCAATGACAC
	R: GATC CCATGG TTGCAAAGGAAAGGTGCCGA
CrPV	F: GATC ACTAGT AAAGCA AAAATGTGAT CTTGCTTGTA AATACAATTTTGA
	R: GATC CCATGG GAAATGTAGCAGGTAAATTTCTTAGG
FLO8	F: GATC ACTAGT AGTTGACGTTAGTAAGTCACTGAGG
	R: GATC CCATGG CACTATTCGTTGCCATT
GPR1	F: GATC ACTAGT AAAAAAACGACAAACAAGTGATCCGAAG
	R: TAAT CCATGG CAACCCCAGCAGCTGGTTAA
HAP4	F: GGTC ACTAGT ACCTCTCTAAACCCCAGTTTTATATC
	R: GATC CCATGG GGTTGGTATTTGGGG
NCE102	F: GGTC ACTAGT AACATTTCTTTAAAAAACATCTTCAATAAGAAAAATCGG
	R: TAAT CCATGG CAGTAGTTTACTCTGGAGCTGTGC
FLO8 reverse	F: TA CCATGG AGTTGACGTTAGTAAGTCA
	R: TA ACTAGT CACTATTCGTTGCCATTGC
GPR1 reverse	F: TA CCATGG AAAAAAACGACAAACAAGTGA
	R: TA ACTAGT CAACCCCAGCAGCTGG