2013

The Panicum mosaic virus-like 3\' cap-independent translation element: A translation enhancer that functions in mammalian systems

Mariko Sada Peterson
Iowa State University

Follow this and additional works at: http://lib.dr.iastate.edu/etd

Part of the Biochemistry Commons, Molecular Biology Commons, and the Virology Commons

Recommended Citation
http://lib.dr.iastate.edu/etd/13061
The Panicum mosaic virus-like 3’ cap-independent translation element: A translation enhancer that functions in mammalian systems

by

Mariko Sada Peterson

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Biochemistry

Program of Study Committee:
W. Allen Miller, Major Professor
Susan Carpenter
Mark Hargrove

Iowa State University
Ames, Iowa
2013

Copyright © Mariko Sada Peterson, 2013. All rights reserved.
# TABLE OF CONTENTS

ACRONYMS USED IN THIS THESIS:........................................................................................................ iv
ABSTRACT.................................................................................................................................................. v

CHAPTER 1. GENERAL INTRODUCTION: A REVIEW OF THE LITERATURE ..................... 1
   Historical Perspective.............................................................................................................................. 1
   Canonical translation initiation in plants and mammals ................................................................. 2
   Virus strategies of translation initiation: IRESes and CITEs.................................................. 8
   The Panicum mosaic virus-like translation element (PTE), an eIF4E-binding CITE........... 14
   eIF4E in plants and mammals ....................................................................................................... 16
   The Tombusviridae and the Umbraviruses................................................................................. 20
   Project Aims..................................................................................................................................... 25

CHAPTER 2: MATERIALS AND METHODS ................................................................................. 26
   eIF4E Sequence and Structure Alignment .................................................................................... 26
   RNA Preparation............................................................................................................................. 26
   T7-BHK cell culture and DNA transfection ............................................................................... 28
   HeLa cell culture and RNA transfection ................................................................................... 28
   In-vitro translation assays............................................................................................................. 29

CHAPTER 3: RESULTS AND DISCUSSION..................................................................................... 32
   eIF4E sequence and structural alignments.................................................................................... 32
   Aim 1: System Optimization.......................................................................................................... 36
   Aim 2: Performance of PTE-containing luciferase reporters in lysates and cells............. 45
   Aim 3: Performance of the PTE as an inhibitor ....................................................................... 52

CHAPTER 4: CONCLUSIONS ........................................................................................................... 55
   Relevance ......................................................................................................................................... 55
   Proposed model............................................................................................................................... 56
   Future directions............................................................................................................................ 56

APPENDIX 1: 3’ TRUNCATION OF PTE CONSTRUCTS ................................................................. 58
APPENDIX 2: PREPARATION OF HEla CELL EXTRACT ........................................................... 60
ACRONYMS USED IN THIS THESIS:

BHK: Baby hamster kidney (cell line)
BTE: Barley yellow dwarf virus-like cap-independent translation element
CarMV: Carnation mottle virus
CITE: Cap-independent translation enhancer
CMMV: Cocksfoot mild mosaic virus
EMCV: Endomyocarditis virus
HCRSV: Hibiscus ring spot virus
IRES: Internal ribosome entry site
JINRV: Japanese iris necrotic ringspot virus
MNSV: Melon necrotic spot virus
PEMV-2: Pea enation mosaic virus RNA-2
PFBV: Pelagornium flower break virus
PIC: Pre-initiation complex
PMV: Panicum mosaic virus
PTE: Panicum mosaic virus-like cap-independent translation element
RRL: Rabbit reticulocyte lysate
SCV: Saguaro cactus virus
TPAV: Thin paspalum asymptomatic virus
UTR: Untranslated region
VPg: Viral protein, genome-linked
WGE: Wheat germ extract
ABSTRACT

In canonical eukaryotic translation initiation, the recognition of mRNA by the translation machinery is facilitated through the binding of the 5’ m7G cap structure to eukaryotic initiation factor 4E (eIF4E). This recognition is necessary for efficient translation of the message to occur. In contrast, some viruses lack the 5’ cap structure but have alternative mechanisms to initiate translation and express the viral proteins necessary for replication. The positive-strand RNA plant viruses of the panico-, carmo-, umbra-, (1-3) and aureusviruses (4) contain an RNA structure in their 3’ UTRs that binds eIF4E (2) and through communication with the 5’ UTR, are able to efficiently initiate translation in their hosts. This interaction is proposed to occur through a guanosine residue that interacts with the cap-binding pocket of the protein in a similar way to the m7G cap (3). Control of eIF4E expression has important implications in plant susceptibility to viral infection and has been used to design more resistant crops (5-7). Its control is also important in mammalian systems: deregulated eIF4E has been implicated in many cancers including those of the breast (8), lung (9) and prostate (10), chronic human diseases such as diabetes (11), and developmental disorders such as autism (12). As eIF4E structure and function are highly conserved across kingdoms, and characterizing novel interactions with this protein have potential use in disease control, we sought to characterize this element in mammalian systems. We optimized translation in rabbit reticulocyte lysate and HeLa lysate for salt, RNA concentration and incubation time (Aim 1), compared the translation efficiency of reporter RNAs containing PTEs of 9 different viruses in lysates and whole cells (Aim 2), and investigated the ability of the PTE to act as an inhibitor (Aim 3).
CHAPTER 1. GENERAL INTRODUCTION: A REVIEW OF THE LITERATURE

Historical Perspective
In 1958, Sir Francis Crick coined the “Central Dogma of Molecular Biology” which revolutionized the thinking of the field (13). For the first time, it explicitly stated the way in which biological information flowed from nucleic acids to proteins, solidifying the assumptions that scientists had held while discussing the theories for protein expression from genes (14). It stated: “The central dogma of molecular biology deals with the detailed residue-by-residue transfer of sequential information. It states that such information cannot be transferred from protein to either protein or nucleic acid” (14).

At this time, the fields of translation and its control were in their early years. In the 1950s, the field progressed rapidly: it was accepted that proteins were composed of unique combinations of 20 naturally occurring amino acids that were fastened together in an ordered fashion (13). The first protein sequences were published by Sanger and Tuppy, and radioisotopes allowed tracking of the fate of certain amino acids (15-19). In 1951, the first cell-free translation extracts were produced from rat livers, and it was discovered that both GTP and ATP were consumed in the translation reaction (20). In the same time period, protein synthesis was determined to occur rapidly at the ribosome, however, it was not entirely clear if the process required a template (15,21-23). Adaptor RNA, identified as tRNA was discovered and characterized in the late 1950s (24).

In 1961, Jacob and Monod proposed that protein synthesis occurred in a template-directed fashion, suggesting RNA as the template, and the ribosome and other factors forming an elegant system that could both decode the template and synthesize the new peptide (25). Central to this idea was the concept of an unstable messenger molecule that could be easily degraded.
should the cell need to alter its gene expression in response to rapidly changing environmental conditions (25). The messenger was confirmed the same year in bacterial cells to be RNA (26,27).

Translation is a highly conserved process, present in all kingdoms of life and is catalyzed by the rRNA of the ribosome (28). Translation itself has been divided into three stages: initiation, which involves the recruitment and assembly of the ribosome on the mRNA at the start codon, elongation, which is the decoding of the messenger and formation of the peptide chain, and termination, at which point the translation machinery is disassembled and the peptide is released (29). This review will focus on translation initiation. First I will describe canonical translation initiation in higher eukaryotes, and then contrast it with some of the strategies that viruses use to circumvent the requirements for host translation initiation, focusing on a specific RNA structure, the Panicum mosaic virus-like 3’ cap-independent translation element (PTE). I will then discuss eukaryotic initiation factor 4E (eIF4E) because it binds the PTE and because it is highly conserved across eukaryotes. I will finish with a general discussion of the biology of the Tombusviridae and the Umbraviruses, the virus groups in which the PTE is found.

**Canonical translation initiation in plants and mammals**

The process of protein synthesis, known as translation, begins with initiation. Initiation is the rate-limiting step of translation (28,30), is the most highly regulated step (31), and involves more involvement by protein factors than the other two steps in translation (28). In higher eukaryotes, the structure of the mRNA plays an important role in translation initiation and thus regulation (32,33). Nearly all cellular mRNAs undergo modifications after transcription in the nucleus (30). These include the addition of a m7G[5’]ppp[5’]N cap (N being any nucleotide, and the first in the mRNA sequence) at the 5’ end of the mRNA (30,34,35) which allows for efficient
recruitment of the translation initiation machinery via eIF4E binding (34,36), and increases mRNA stability (37), and polyadenylation (38), splicing (39), and nuclear export efficiency (40). In addition to the 5’ cap, a run of adenosines is added to the 3’ end, and is known as the poly(A) tail (41). This structure also stabilizes the mRNA (41,42) and aids in increasing translation efficiency by interacting with the 5’ end of the mRNA through protein-RNA and protein-protein interactions (41,43,44).

In eukaryotes, translation does not occur co-transcriptionally as it does in prokaryotes due to the separation of the processes by the nuclear envelope (30). Instead, after processing, the mRNAs are exported out of the nucleus through pores in the nuclear membrane, into the cytoplasm where they are recognized by initiation factors and the ribosome is assembled onto the message (30).

The mRNA is recognized and bound by eIF4F, which in mammals is a heterotrimer composed of the cap-binding protein, eIF4E, the scaffolding protein, eIF4G, and the helicase, eIF4A (45-47). In plants, eIF4F is a heterodimer composed of eIF4E and eIF4G (48). Recognition of mRNA by eIF4F in both plants and mammals is facilitated through binding of the eIF4E subunit to the 5’ cap structure (47,49). Binding of eIF4G increases cap-binding activity (36,50). eIF4B, which assists in eIF4A helicase function, joins the complex (45,51,52). The mRNA is activated after these (and perhaps more) factors bind and when the poly(A) binding proteins (PABPs) associated with its poly(A) tail interact with eIF4G to circularize the message (43,44,53,54).

While some initiation factors bind to the 5’ end of the transcript, others first bind the small (40S) subunit of the ribosome, preparing it to interact with the message (30). The small ribosomal subunit differs in size from the prokaryotic version (which is 30S and composed of
different sizes of rRNAs and scaffolding proteins) (30,55-57). It is equipped with three tRNA binding sites: A, P, and E (29,56).

The ternary complex, composed of eIF2 bound to GTP and the initiator methionine-bound tRNA interacts with the 40S ribosome bound to initiation factors eIF3, eIF1A, and eIF1 (Fig. 1) to form the 43S preinitiation complex (PIC) (58). This interaction is presumably facilitated by eIF5, which binds eIF3 and eIF2 (30). The initiator tRNA is unique in that it binds to the P site of the 40S ribosome after ternary complex interaction (58,59). All other tRNAs interact with the A site. The interaction with the P site is important to initiating elongation because the conformational change in the ribosome upon formation of the peptide bond initiates the ratcheting motion of the ribosome across the mRNA (58,60) Association of eIF3 has many roles including to help block premature association with the 60S ribosome. While most of eIF3 binds on the solvent-accessible side of the 40S ribosome (61), one of its domains wraps around to obstruct part of the binding site of the 60S ribosome (62). Without eIF3 and other initiation factors, the PIC is disassembled by 60S ribosome binding, displacing the ternary complex to form 80S ribosomes with 40S subunits that are not associated with mRNA (63). However binding of these factors stabilizes the PIC against 60S joining (62,63) and dissociation allows ribosome assembly when appropriate (64,65). eIF3 also helps recruit and position mRNA in the 43S by interacting with eIF4GPIC (62), and to disassemble 80S ribosomes, along with other initiation factors (66). eIF1A is a homologue of the prokaryotic eIF1 and blocks the A site to prevent premature entrance of tRNAs (67,68), and eIF1 binds near to the P site (69). These factors are important in PIC formation and selection of the start codon (56,70-72). The ternary complex, 40S ribosome and initiation factors are known as the 43S ribosomal preinitiation complex (PIC) (30,56).
The 43S PIC is now ready to interact with the activated mRNA. Multiple factors (eIF4A, eIF4B, and possibly PABP) work to unwind the secondary structure naturally present in most 5’ untranslated regions (UTRs) of RNAs (33,73). In order for the ribosome to scan to the start codon, the mRNA secondary structures must be unwound (33,73-75).

eIF4G serves as a scaffolding protein, simultaneously binding the mRNA, the cap-binding protein, eIF4E, the ATP-dependent helicase eIF4A, and eIF3. It anchors the 43S PIC to itself by its interaction with eIF3 (76). Cap-bound eIF4 factors specifically direct the PIC to join the mRNA and direct it toward cap-dependent translation while discouraging an anomalous eIF3-dependent mechanism from occurring (77). The next task at hand is to scan to the start codon. In prokaryotes, start codons are distinguished from internal methionine codons via a context-dependent mechanism in which the rRNA of the ribosome interacts with the mRNA at the Shine-Dalgarno sequence (78). In eukaryotes, no such interaction occurs. Instead, the ribosome scans until it reaches the first start codon, usually an AUG. Initiation is enhanced in a context-dependent manner when the ribosome reaches a consensus sequence present in many RNAs known as the Kozak sequence: GCC(A/G)CCAUGG, with optimal bases in bold and underlined (79-83).

The ribosome carries the ternary complex with the initiator tRNA and the initiation factors with it as it scans (perhaps leaving eIF4E at the cap). eIF1 enhances selection of the correct codon by discriminating against incorrect codon-anticodon interactions between the message and the initiator tRNA (33). Codon selection and scanning is aided by eIF1A (70,71). When the first AUG is located, the anticodon of the initiator tRNA, CAU interacts with it (30). After the start codon is selected, the GTP associated with eIF2 is hydrolyzed, bringing about a conformational change that allows the initiation factors to leave, including eIF3 (31,64). eIF5B,
bound to GTP joins the vacated A site, preventing tRNAs from entering it before the 60S ribosome joins (84,85). The 60S ribosome joins the 40S ribosome’s newly solvent-accessible face, hydrolyzing the GTP molecule associated with eIF5B and releasing eIF5B (84). This marks the formation of the 80S initiation complex and the end of the initiation step in translation. Elongation of the peptide chain and eventual termination follow.

In the elongation phase, entering amino acyl-tRNAs (aa-tRNA) bind the message through an anti-codon-codon interaction in the A site(29,56). Peptide bond formation occurs between the aa-tRNA in the A site and the peptidyl-tRNA in the P site which is linked to the growing peptide chain. This transfers the growing peptide from the tRNA in the P site. After peptide bond formation, the translocation step occurs in which the ribosome ratchets in a 3’ direction along the transcript, the new peptidyl-tRNA is moved to the P site to make room for incoming aa-tRNA and the deacylated tRNA in the P site is transferred to the exit (E) site (60). Peptide bond formation is catalyzed by the ribosome which properly positions the amino acids to react (56). This positioning is made possible by an A/P hybrid state in which the amino acid attached to the aa-tRNA leans into the P site, which is the result of the small ribosome moving in the 3’ direction relative to the large ribosome (60). Elongation factors stabilize hybrid states and help catalyze the movement of the ribosome along the mRNA via GTP hydrolysis (56). The 5’-most position in relation to the mRNA is the E site, where the empty tRNAs are shuttled and then exit the ribosome (29,56).
Fig. 1: Cap-dependent translation
Cap-dependent translation begins with the formation of the ternary complex, the activated mRNA, and the pre-initiation complex (PIC). After the mRNA is recruited to the pre-initiation complex and the ribosome is properly assembled, translation may occur (adapted from Aitken, 2012).
The 80S ribosome dissociates from the mRNA and is recycled to the beginning of the process when the ternary complex, eIF3, eIF1, and eIF1A release the 40S subunit (66).

**Virus strategies of translation initiation: IRESes and CITEs**

Viruses are obligate intracellular parasites of cells, and all but the megaviruses (86) lack any components of the cellular translation system. No viruses encode ribosomal proteins or RNAs. Therefore, in order to express the proteins required for infection, the virus must be able to hijack the cell’s translation machinery. For viruses that lack a 5’ cap and a poly(A) tail, this is difficult since the translation initiation factors show preference for mRNAs with these features. Two of the strategies that have evolved to avoid these structural requirements for recognition and efficient translation initiation involve secondary structures in the viral genomic RNA: internal ribosome entry sites (IRESes) and cap-independent translation enhancers (CITEs).

**IRESes**

IRESes were first characterized in the 5’ UTR of the poliovirus (PV) (87). The IRES is highly structured, conserved in picornaviruses, and allows efficient recruitment of the ribosome to the 5’ UTR of the virus despite its lack of a 5’ cap (87-89). Since then, many IRESes have been discovered in viral and host mRNAs (90) and not all function by the same mechanism. The viral IRESes have been divided into four classes based on structure and mechanism (91-93).

Group 4 IRESes are illustrated by the picornavirus, poliovirus. These IRESes require many cellular factors to function, are located at the 5’ UTR, and are highly structured (91,94). Early studies showed that poliovirus lacks a 5’ cap and contains a VPg instead (95,96). The VPg of polioviruses is a 22-residue peptide that is covalently linked to the 5’ end of the genome by a tyrosine and is required for viral replication (95,96). Infection was also accompanied by drastic shut down of host translation. Upon infection, Poliovirus protease 2A cleaves eIF4G such that
one segment can bind eIF4E and the other can interact with other initiation factors (97,98). In this way, Poliovirus shuts down host translation by preventing eIF4E-mediated recognition of cellular mRNAs and recruitment of factors. The other segment of eIF4G is used by the virus, and the IRES recruits IRES-trans acting factors (ITAFs), as well as eIF4A, eIF4B and the 43S PIC (91,93). Translation initiates downstream of the IRES, with the ribosome scanning to the start codon (94). Due to their dependence on ITAFs, hypothesized to help to stabilize the IRES structure and protein interactions (94), this class of IRESes is unable to function efficiently in rabbit reticulocyte lysate (RRL), and needs to be supplemented by factors in HeLa lysate, specifically the La and PTB proteins to function (99,100). Optimal function also depends on interaction with the poly(A) tail (101).

Group 3 IRESes also include members of the Picornaviridae, such as EMCV. These IRESes are similar to their cousins in Group 4, except they do not scan to the start codon, and while they depend on ITAFs for function, they can function efficiently in RRL (91,93,94). Efficient functioning of this IRES relates to how well the virus is able to infect cells, and is dependent on a conserved region in the center of the IRES, the J-K stem loops, and a polypurine tract (102,103).

Group 2 IRESes are characterized by the Flaviviridae member, Hepatitis C Virus (HCV). HCV lacks a poly(A) tail, but contains a highly structured region at its 3’ UTR (104,105). The HCV IRES is important to both replication and translation (91). These IRESes bind the 40S ribosome directly, are less structured than the Group 3 and 4 IRESes and require fewer factors to function: only eIF3, eIF2, and the initiator tRNA (94). The start codon is part of the IRES (domain IV) and the ribosome is recruited directly to it, without the requirement for scanning
Upon interaction with the IRES, conformational changes position the start codon into the ribosome and allow elongation to occur (106-108).

Group 1 IRESes are the simplest of the IRESes in terms of structure, size, and the number of factors required for translation initiation (92,93). The Group 1 IRES of Cricket Paralysis Virus, a dicistrovirus is located not in the 5’ UTR, but in the intergenic region (IGR) between ORFs 1 (nonstructural proteins) and 2 (structural proteins) (109). Amazingly, this IRES requires only the 40S ribosomal subunit. The structure of the mRNA itself allows the virus to skirt the requirement for the initiator tRNA by molecular mimicry: the mRNA folds into a state that mimics the transition state after peptide synthesis and before the ribosome moves across the mRNA (92,110).

CITEs

While IRESes have been identified in some plant viruses (111-113), these viruses employ other mechanisms to thwart the cap-dependent monopoly on translation initiation (Fig. 2). One of these strategies involves the recruitment of translation components to structures in the CITEs. CITEs are present in the positive sense RNA genomes of the Luteoviridae, Tombusviridae, and the umbraviruses and nepoviruses (114-116). These CITEs share four main characteristics: first they are located in the 3’ UTR (the opposite end of the RNA at which scanning of the ribosome is initiated), second, they bind translation factors, third, they communicate in some way with the 5’ UTR, and last, facilitated by this interaction, the ribosome associates with the 5’ UTR and scans from the 5’ end (114-116).

CITEs are divided into six classes based on their secondary structures (Fig. 2). 1. The first CITE to be identified, the translation enhancer domain (TED) in satellite Tobacco necrosis virus of the necroviruses, interacts with eIF4F (117,118). Optimal translation is observed when
Fig. 2: 3’ CITEs spread across many genera
Secondary structures of known classes of 3’ CITEs are shown at top of table. Initiation factors to which they bind are shown (adapted from Nicholson and White, 2011 and Simon and Miller, 2013).

<table>
<thead>
<tr>
<th>Translation component</th>
<th>elf4F and elfiso4F</th>
<th>elf4E and elf4G</th>
<th>60S ribosome</th>
<th>elf4E</th>
<th>elf4E and elf4G</th>
<th>elf4G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tombusviridae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tombusvirus</td>
<td>+</td>
<td></td>
<td></td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Necrovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Aureusvirus</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carmovirus</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panicovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Avenavirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machlomovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dianthovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Luteoviridae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luteovirus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Umbraviruses</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
both the 5' and 3' sequences are present. When either the 3' or 5' UTR is mutated, translation can be restored upon addition of a 5' cap; however addition of either a 5' cap or poly(A) tail to a reporter containing the STNV TED and leader sequence did not enhance translation (119,120). Meulewaeter et al. proposed that this was due to the enhancer 1. Operating by a different mechanism than capped and poly adenylated RNAs, and 2. That this enhancer preferentially enhances uncapped RNA (120). While the STNV TED contains sequences that are complementary to a 5' end stem-loop, suggesting a possible kissing interaction, Meulewaeter et al. found that abolishing complementarity did not abolish translation (120). This could mean that a protein interaction could serve as a bridge mediator between the 5’ or 3’ UTRs, or that the interaction needs more STNV sequence than that which was used in Meulewaeter and colleagues’ experiment (116,120)

2. The I-shaped structure of *Tobrusviridae* also interacts with eIF4F (121). This CITE is interesting in that it interacts both with the 5’ UTR through a kissing stem-loop interaction and with the eIF4E subunit of eIF4F. Nicholson et al. showed that while the I-shaped structure of Melon necrotic spot virus (MNSV) can interact specifically with eIF4E, binding was strongest and translation enhancement optimal with eIF4F binding (121). In addition, a strong preference for eIF4F over eIFiso4F was shown which fits with Gallie and Browning’s finding that eIF4F preferentially assists in the initiation of translation of structured mRNAs (121,122). The finding that this element in MNSV binds eIF4F and interacts with eIF4E is compatible with the development of resistance in melons to viruses with similar CITEs (123).

3. The Barley yellow dwarf virus-like translation element (BTE) of the luteoviruses binds eIF4G (124). Past studies have demonstrated that long-distance basepairing between the BTE and a stem loop in the 5’ UTR are essential for efficient translation enhancement (125). Unlike
the HCV IRES, the ribosome is not recruited directly to the start codon, but must scan from the 5’ end (126). These elements contain a 17 nucleotide conserved sequence (127). The structure of this element is composed of a three to six helix helical junction, depending on the virus (2,116). Bases forming the junction are important to function, with significant reduction in function of the element when these are mutated (128). In 1995, this element was tested in rabbit reticulocyte lysate and found not to be functional, despite being very efficient in wheat germ extract (129).

4. The T-shaped structure of the Carmo- and Umbraviruses, binds the 60S ribosomal subunit and the 80S ribosome (130). Ribosome binding is essential to its activity (130). As with other CITEs, this element also communicates with the 5’ UTR, however it does not interact through kissing interactions as with the BTE (131). Rather, Anne Simon proposes that the bound ribosomal subunits are able to form a protein bridge that link the two UTRs together, delivering the ribosome to the 5’ UTR (131). Interestingly, this element’s three-dimensional structure resembles that of tRNA, and it may bind the ribosome in between the A and P sites (130). A T-shaped structure has also been identified in PEMV RNA-2 upstream of the PTE (132). It is thought to bind the ribosome and communicate with the 5’ UTR, thus serving as the mode of communication between the 5’ and 3’ UTRs in this virus (132).

5. The Y-shaped structure, found in several genera of the Tombusviridae interacts with eIF4F and can be replaced with the PTE or the I-shaped structure (4). This versatility illustrates the hypothesis that CITEs have been exchanged between viruses by recombination as their distribution does not follow classical taxonomy (4,116).

6. An additional CITE has been characterized in Blackcurrent reversion virus, a Nepovirus (113). As of yet, it has not been classified because its secondary structure is
unknown, however 5’-3’ UTR communication is required for function (113) and function may implicate an IRES that was identified in the 5’ UTR (133).

The Panicum mosaic virus-like translation element (PTE), an eIF4E-binding CITE

The final CITE category is the Panicum mosaic virus-like translation element (PTE). The PTE was first identified in Panicum mosaic virus (PMV), a panicovirus (1). Since then it has been identified in the two other known panicoviruses (Fig. 3): Thin paspallum asymptomatic virus (TPAV) and Cockfoot mild mosaic virus (CMMV), five carmoviruses: Hibiscus chlorotic ring spot virus (HCRSV), Japanese iris ringspot virus (JINRV), Saguaro cactus virus (SCV), Carnation mosaic virus (CarMV), Pelargonium flower break virus (PFBV), and an umbrevirus: Pea enation mosaic virus (PEMV) RNA 2 (3,134). The element has also been recently identified in an additional Tombusviridae genus: the aureusviruses Pothos latent virus and Cucumber leaf spot virus (4,115).

The PTE is a T-shaped structure consisting of a stalk-like stem structure with a G-rich bulge. There are two stem loops, SL1 which was found to interact with upstream sequences near the 5’ end of the viral genome in carmoviruses (132), and SL2 which are divided by a C-rich bulge. The C and G rich bulge domains are hypothesized to interact to form a pseudoknot, and a hypermodifiable G residue (named for its reactivity in SHAPE structure probing) is theorized to interact with the cap-binding pocket of eIF4E (3,134). The PTE was shown to interact very strongly with eIF4E with a Kd of ~ 58nM, which is comparable to eIF4E’s interaction with the cap analog m7GTP at Kd ~ 11 nM (3,134,135). The PTEs, unlike the 17 nucleotide conserved region of the BTEs share no sequence identity (except for the C and G rich domains), but are
Fig. 3: Representatives of the PTEs from three different genera.
The PTEs investigated in this work are those of the umbra-, panico-, and carmoviruses. (Wang, 2011).
highly conserved three-dimensionally, with the PTE presumably adopting a very stable coaxial stacking structure (3).

Communication with the 5’ UTR is thought to occur one of two ways, depending on the genus of the virus. In the *Umbravirus*, PEMV, there is a Turnip crinkle virus-like T-shaped structure, the kI-TSS, that not only binds a stem loop in the 5’ region of the coding region for p33, but also binds both ribosomal subunits for dual-action delivery of the ribosome and communication between the 5’ and 3’ UTRs (132). Interaction of this element with this stem loop also brings the PTE bound to eIF4E proximal to the 5’ region of the genome (132). Interestingly, in the Simon lab, it was this interaction that proved to be more important for translation enhancement than the existence of the PTE itself (132) in contrast to the results obtained in the Miller lab (134). The PTE in other genera is thought to communicate directly with the 5’ UTR, not through a ribosome-binding intermediate. In SCV, SL1 has been shown to basepair with a stem loop in the coding region of the genome, close to the 5’ UTR (not in the 5’ UTR itself) (136). It is thought that panicoviruses communicate in a similar way (136).

**eIF4E in plants and mammals**

eIF4E, the translation initiation factor that binds the PTE, is the cap-binding protein that recognizes the 5’ end of the mRNA by binding to the 7-methyl guanosine (m7G) cap structure located there (49,137). In both plants and mammals, the cap structure is recognized and sandwiched between two parallel tryptophan residues (Trp-62 and -108 in wheat, and Trp-56 and -102 in human) in the cap-binding pocket (137-139). The two tryptophan residues and the m7G form π-π interactions that result in a highly stable interaction (138,139).
Fig. 4 eIF4E alignment.
Muscle sequence alignment (Jalview) of the cap-binding pocket region of eIF4E sequences acquired from NCBI from the indicated species (A) was used to make a neighbor joined phylogenetic tree (B) of eIF4E (Mega). The cap-binding tryptophans are highlighted in red and blue (A). Structure representatives of Animalia (Human, C), Plantae (Wheat, D), and Fungi (Baker’s Yeast, E) were overlaid (Swiss-PDB Viewer) and rendered (Pov-Ray) to show structure conservation (F).
The cap binding pocket is composed of alpha helix and loop regions on either side, and beta sheet regions at the bottom, when the protein is oriented such that the pocket is facing up (30,138,139) (Fig. 4C-F). The protein itself is composed of 3 alpha helices and 8 parallel beta sheets (138,139). It is an important protein in translation initiation because it facilitates recognition of the mRNA by other factors (especially the eIF4G and eIF4A subunits of eIF4F) as well as the 43S ribosome (30). eIF4G itself interacts very tightly with eIF4E, forming a cuff around a “wrist-like” structure toward the “fist-shaped” N-terminus of eIF4E (36). This interaction increases eIF4E’s affinity for the cap structure and is extremely stable (36). PABP interactions also increase cap interactions (140).

Efficient translation of most cellular mRNAs cannot occur without its binding to the cap structure (49). While sequence alignments of eIF4Es from diverse organisms show significant homology (138), the three-dimensional structure shows amazingly strong conservation across the three eukaryotic kingdoms: Plantae, Mammalia, and Fungi (138,139,141), with the termini of the protein showing some variation (138,139,141) (Fig. 4F). Illustrating this structural homology, eIF4E from mouse can be substituted for that in yeast without significant detrimental effects to the cells’ growth and reproduction (142).

The rate of cell division is intimately linked with the rate of translation (143). As eIF4E is part of the repertoire of proteins that play a role in initiation of translation, it is highly regulated in mammals (143,144). Mnk1 phosphorylates eIF4E on Serine 209 and is part of the MAP kinase (MAPK) pathway (143,144). This phosphorylation leads to increased translation levels. In metazoans, eIF4E is also negatively regulated through the 4E-BP (eIF4E binding protein) family (143,144). These proteins bind eIF4E at a site that overlaps with that of eIF4G. When they are unphosphorylated, they bind eIF4E, competing with eIF4G and slowing rates
Loss of eIF4E regulation resulting in over-activation has been implicated in many diseases of mammals including multiple types of cancers such as breast (8), prostate (10), lung (9), and cervical cancers associated with HPV (145,146), chronic human diseases such as diabetes (11) and developmental disorders such as autism (12). Viruses also manipulate eIF4E. In herpesvirus infection, for example, eIF4E phosphorylation is increased through virus-induced alteration of cellular signaling pathways (144,147). In picornavirus infection, hypophosphorylated 4E-BP accumulates to sequester eIF4E (144,148,149). In adenoviruses, the opposite occurs, and 4E-BP is inactivated to increase protein synthesis in early infection (150). Interestingly, plants appear to lack any homologs of 4E-BP (151).

Plants contain another layer of initiation factors that are absent from animals: that is, they contain isoforms of eIF4F and its subunits eIF4G and eIF4E (48). These are referred to as eIFiso4F, eIFiso4G, and eIFiso4E. eIFiso4E is similar in size to eIF4E, however, eIFiso4G is about half the size of eIF4G (48,122). These isoforms are highly abundant in the cell (with eIF4F at about 11% the levels of eIFiso4F) and capable of interacting with capped mRNA. However eIFiso4F is significantly less efficient (about 20-30 fold) than eIF4F at engaging capped mRNA and facilitating translation initiation (48,122). It is thought eIF4F and eIFiso4F interact with different classes of mRNA.

eIF4E is a substrate for phosphorylation under certain environmental conditions (for example under oxygen deprivation in maize), however no homologues to the mammalian 4E-BP have been found (48).

The presence of two forms of eIF4E has been important in the evolution of virus resistance and crop development(152-154). The RNA genome of viruses in the Potyviridae
family contains a protein (VPg) covalently attached to the 5’ end in place of a cap (48,155). This protein interacts with either eIF4E or eIFiso4E, depending on the virus, and mutant versions or deletions of eIF4E or iso4E in several plant species including Arabidopsis (48,156), pepper (6), tomato (7,48) and lettuce (5,48) have been observed to confer resistance to potviruses (48,152-154).

**The Tombusviridae and the Umbraviruses**

While the PTE has been identified in the aureus-, carmo-, and panicoviruses of the Tombusviridae as well as the umbraviruses, the work in this thesis focuses on the carmo-, panico-, and umbraviruses (Fig. 5). This section will discuss characteristics of the viruses harboring PTEs that are relevant to this work.

Carmo- and panicoviruses are classified in the Tombusviridae family (157). They possess icosahedral capsids with T = 3 symmetry and they lack envelopes. The genome of viruses in the Tombusviridae consists of one molecule of positive-sense, linear RNA. It is uncapped and lacks a poly(A) tail (157). The genome contributes a significant amount to the capsid weight (at 17%) and ranges in size from 3.7 to 4.8 kb (157).

Carmovirus genomes contain four open reading frames (ORFs) (Fig. 5A). The first ORF contains a leaky stop codon that allows some ribosomes to translate the downstream ORF resulting in a population of proteins with two different C-termini (158,159). ORFs 2, 3, and 4 overlap slightly (157). ORF 1 codes for two proteins: a 28 kD protein, and an 88 kD polymerase (when readthrough occurs) (157,159). ORF2 encodes an RNA binding protein which is basic and shares similarity with other Tombusviridae genera (157,160). ORF3 encodes an integral membrane protein (157,160). Both of these proteins are important in cell-to-cell movement within the infected host (157,160). The coat protein (CP) ORF is located toward the 3’ end of the
genome, and requires sgRNA synthesis for expression (157). It is required for long distance
movement of the virus, and in some viruses is a silencing suppressor, e.g., HCRSV (161) and
Carmovirus species have a limited natural host range (162), but the genus as a whole infects
many different species of plants – both monocots and dicots (157). Transmission occurs easily
via mechanical means, especially CarMV which has been spread through contaminated carnation
cuttings (157,163). Other members of the genus may be spread through infected seeds or soil
(157).

Carmoviruses include: PFBV (158,164), JINRV (165), HCRSV (166), CarMV(167), and
SCV (168). These examples contain PTEs (3). The impacts of these viruses vary, but many of
them affect ornamental flower production by causing unsightly pathologies. PFBV, for instance
is one of the most common viral agents infecting plants of the species Pelargonium, such as
geraniums (158). Common symptoms are flower streaking, leaf chlorosis and stunting (158).
JINRV is known for its “spindle-shaped” necrotic lesions on the leaves of Japanese irises (165).
HCRSV was first identified in the United States from hibiscus plants imported from El Salvador
(166). Common symptoms include stunting, distortion of flowers, chlorotic ring spots (as the
name implies), and leaf streaking (166). Previous studies have shown that translation
enhancement activity occurs from the 3’ UTR (112,169). This activity was proposed to originate
from an IRES (169), with the sequence GGGCAG essential for activity (169). This sequence,
located at nt 3659-3664 corresponds with the G domain of the HCRSV PTE (3,169). CarMV is
the type virus of this genus (157,167), and causes mild symptoms in its native host, carnation
(163). It was shown to be able to express proteins in rabbit reticulocyte lysate in 1985 (170).
SCV in contrast to the other Carmoviruses studied in this work does not cause a symptomatic
infection (168). It was first isolated in Arizona during a survey of cactus species and infects only
the giant saguaro cactus naturally (168).

The Panicoviruses contain 5 ORFs, with a single subgenomic (sg) RNA encoding ORFs
2-5 (157) (Fig. 5B). ORF 2 encodes the polymerase, which is similar to other Tombusviridae
except for a longer N terminal domain which makes the protein larger (112 kD) in most of the
members of the genus (157). ORF2 encodes the movement protein, which is 8 kD in size (157).
sgRNA also encodes a 6.6 kDa protein important in movement (157). The coat protein (ORF 3)
and p15 from ORF4 are also implicated in movement (157). Hosts of this virus are the Panicea
tribe of Poaceae grasses (157,171). Transmission occurs mechanically, mainly through
contaminated soil (157).

The Panicoviruses that contain PTEs are TPAV, a newly characterized virus of tall grass
prairies (3), PMV, and CMMV (172). Little is known of TPAV (3), however CMMV and PMV
are relatively well-characterized. PMV was first characterized in the 1950s in the United States
in switchgrass (171). It is a pathogen of St. Augustinegrass, and is the causative agent of St.
Augustine decline which drastically affected production of the grass in the 1960s in Texas (171).
The virus also causes widespread disease in centipede grass, along with a satellite virus that is
dependent on PMV for propagation, but codes for its own capsid protein (171). Symptoms
involve stunting, chlorosis, and reduced seed production (171). CMMV infects grasses such as
Timothy and cocksfoot grass, and is closely related to other viruses such as Phleum mottle (172).
The virus causes mottle symptoms in grass (172).

Umbraviruses have not yet been assigned to a family of viruses, but are closely related to
the carmoviruses of the Tombusviridae (173). They do not form their own capsids, and depend
on members of the Luteoviridae for packaging and dissemination (e.g., PEMV-2, an Umbravirus
depends on PEMV-1, a member of the *Luteoviridae* family (174). Infection by an Umbravirus is stable when it is alone, and possible by mechanical inoculation of infectious RNA (175). Studies have suggested that its RNA is enveloped due to its sensitivity to organic solvents (173). These membrane-bound structures are proposed to protect the RNA, or to serve as compartments for replication (or both) (173). Umbravirus genomes are composed of a single positive stranded RNA 4 – 4.2 kb in length, and are uncapped and unpolyadenylated (173). The genome contains no structural proteins, but the virus contains 4 ORFs. ORF1 and ORF2 partially overlap, with a frameshift site which is predicted to allow translation of the RdRp from ORF2 (173). ORFs 3 and 4 almost completely overlap, but are in different reading frames and each encode a 26-29 kDa protein (173). The ORF3 product is predicted to protect the RNA and ORF 4 shows similar to movement proteins (176). Umbravirus-infected tissues are commonly associated with viral dsRNA that becomes infective after it is heat denatured (173). These viruses have very narrow host ranges, are transmitted in nature by a helper virus via aphids, and are present worldwide (175).

PEMV-2 is the only Umbravirus with a PTE (3,134,177). Its helper virus is PEMV-1, an Enamovirus in the *Luteoviridae*, and causes disease in legumes with significant crop loss (177). PEMV-1 provides PEMV-2 with structural proteins, as PEMV-2 does not encode its own capsid, and vector competence; However, PEMV-2 provides PEMV-1 with long-distance movement properties in the plant (177).
Fig. 5: Viral genome organization. Genomes of the three genera studied in this thesis are shown (Richon, 2012; Ryabov, 2012).
Project Aims

The main goal of this work was to determine if the PTE is able to enhance translation in mammalian systems. We sought to accomplish this goal through three aims: the first was to optimize the in vitro mammalian systems for translation of PTE-containing mRNA reporters. The second was to test the reporters in cis and trans and compare the relative translation levels of these RNAs to those in in vitro plant systems. The last was to test PTEs in whole cells.
CHAPTER 2: MATERIALS AND METHODS

eIF4E Sequence and Structure Alignment

Twenty eIF4E sequences from organisms from all four eukaryotic kingdoms were collected from NCBI (Accession numbers contained in Table 1). These were then aligned using the MUSCLE sequence alignment algorithm (178) in Jalview v. 2.8 (179). Default settings were used. Alignments were then analyzed in MEGA v. 5.05 (180) to create a neighbor joined phylogenetic tree (181). Bootstrap value option was chosen and set at 1000. 3D structures were obtained from Protein Data Bank (Accessions numbers: Human: 1IPC, Yeast: 1AP8, Wheat: 2IDV), overlaid using the “iterative Magic fit” tool in Swiss PDB Viewer v. 4.03 (http://www.expasy.org/spdbv/), and the RMS values were calculated pair wise for each protein (182). 3D renderings for sharper cartoon images were accomplished with Persistence of Vision Raytracer (Pov-Ray) v. 3.62 (http://www/povray.org) (183).

RNA Preparation

The Firefly luciferase reporter RNAs (3,134) were in vitro transcribed using the Ambion T7 Megascript™ Kit after linearization of the template DNA with EcoICRI in the case of PMVlucPMV, and SmaI for the others. Reactions were assembled according to manufacturer’s protocol, with 1 μg of linearized DNA, reaction buffer, NTPs, and T7 polymerase in the reaction mixture. They were then incubated at 37 °C for four hours. The reaction was stopped by adding DNase to degrade the DNA template, and incubated for 15 minutes and 37 °C. Afterward, the RNA was precipitated by adding lithium chloride and storing the RNA overnight at -20°C. The RNA was washed with ethanol, dried and then resuspended in water. In the case of RNAs that were used for RNA transfection, the RNA was phenol chloroform extracted and then ethanol precipitated.
As a control, the construct, CAluc, was used (184). It contains vector sequence flanking the Firefly luciferase gene and a 62 adenosine poly(A) tail. There are no known enhancing elements in this construct. The DNA was linearized using EcoI CRI. This construct was transcribed using the Ambion T3 Megascript™ Kit, with identical protocol except using T3 polymerase and appropriate T3 buffer. After precipitation and reconstitution in water, the RNA was capped post-transcriptionally using Cell Script’s T7 mScript™ Standard mRNA Production System to serve as a capped, poly(A) control. This kit has the capacity to assemble a cap 1-RNA out of a cap-0 RNA, which is important because cap 1-RNA is the natural cap structure of higher eukaryotes. This structure is more representative of the behavior of RNAs native to the systems that we are using. Briefly, the RNA is added to a mixture of capping enzyme, 2’-O-methyltransferase, SAM, GTP, and buffer and incubated for 1 hour at 37 °C. Manufacturer’s instructions were followed. After the reaction is complete, the RNA was either lithium chloride precipitated, or phenol-chloroform extracted and ethanol precipitated depending on its downstream use.

RNAs used for cell experiments were capped using a G(5’)pppA(5’) cap analog structure (A-cap) from New England Biolife. RNAs capped with this analog were co-transcriptionally capped using the T7 Ambion Megascript™ Kit a modified protocol (101). Briefly, 4.2 μL of 40 nM G(5’)pppA(5’) structure was added to a standard 20 μL Megascript reaction, adjusting the GTP concentration to 0.4 μL per reaction to increase capping efficiency and accommodate the volume change. All RNAs used in cell experiments were phenol chloroform extracted and ethanol precipitated.
T7-BHK cell culture and DNA transfection

T7-BHK cells were cultured as previously described (185,186). Briefly, cell growth media was prepared by supplementing Glasgow Minimal Essential Media (GMEM) with amino acids, 10% fetal bovine serum (FBS), and penicillin-streptomycin antibiotic. To maintain the T7 expression plasmid, the antibiotic G418 (geneticin), was added every other passage at a concentration of 1 mg/mL (185,186). Cells were maintained on 75 cm$^2$ flat bottom culture flasks, at 37° C, and at an atmospheric CO$_2$ concentration of 7%.

When the cells reached about 80% confluency, they were passed. To transfect the cells, the Invitrogen Lipofectamine™ Transfection Reagent was used. Manufacturer’s protocol was followed. Briefly: Cells were seeded onto 6-well plates 24 hours before experiments at a density of 2 x 10$^5$ cells/well and 80% confluency was attained. 4 ug of DNA was added 250 uL Opti-MEM/plate and 10 uL Lipofectamine was added to 250 uL Opti-MEM/plate. The transfection mixture was allowed to incubate at room temperature for 5 minutes, then added to the media of the cells. Significant cell death was not observed. After 24 hours, the cells were placed on ice, removed from the plate by scraping, and lysed using Lysis Buffer (Promega) and frozen at -80 C until analysis following manufacturer’s instructions. To read, lysate was diluted 1:4 in lysis buffer, and 20 uL were added to 100 uL of Luciferase Assay Reagent (Promega) and read with a GLO-MAX Luminometer.

HeLa cell culture and RNA transfection

HeLa-S3 cells were cultured as described with some alterations by Rakotondrafara et al. (187). Briefly, Dubelcco’s Modified Eagle Medium with Gluta-MAX was supplemented with 10% fetal bovine serum and 1% Penicillin/Streptomycin (CellGro: 10,000 IU/mL penicillin;
10,000 ug/mL streptomycin). Cells were incubated at 37°C, and 7% CO₂ and passed when they reached about 90% confluency (187).

To transfect with RNA, the TransMessenger Transfection kit (Ambion) was used, and manufacturer’s protocols were followed. Briefly: 5 x 10⁵ cells/well were seeded 24 hours before transfection in a 6-well plate and allowed to grow to 80 – 90% confluency. 2 ug A-capped RNA at a minimum concentration of 0.1 ug/mL was added to pre-mixed 4 uL Enhancer R in Buffer EC-R to a final volume of 100 uL and mixed thoroughly. This mixture was incubated on the bench top at room temperature for 5 minutes. 8 uL TransMessenger Transfection Reagent was added to this mixture and mixed thoroughly. This mixture was incubated for 10 minutes at the bench top for transfection complexes to form. While this incubation took place, the growth medium of the 6 well plates was discarded, avoiding drying the cells. After incubation of the mixture, 900 uL growth medium without supplements was added to the tubes, gently mixed and added to the cells. Cells were incubated for 3 hours, removed by scraping, mixed with Lysis Buffer (Promega), and read following manufacturer’s protocols (20 uL of lysate was added to 100 uL LAR and read with a GLO-MAX Luminometer).

**In-vitro translation assays**

In vitro translation assays were performed in HeLa cell extract (188) and commercially-available wheat germ extract (WGE) and rabbit reticulocyte (RRL) lysate.

*Wheat germ extract translation:*

Wheat germ extract was obtained from Promega (Madison, WI). Reaction set-up followed manufacturer’s instructions (Promega), and as described previously, with optimization of potassium ranging from 50 mM to 200 mM. Final concentration of RNA was 8 nM. Briefly, 5 uL WGE, 0.8 uL kit-provided amino acids –leu, 0.8 uL amino acids –met, optimized volume of
kit-provided potassium acetate, 0.2 uL RNasin, 2 uL diluted RNA, volume brought up to 10 uL total by adding nuclease-free water. Reactions were set up on ice. After reactions were set up, they were allowed to sit on the bench top (25° C) for one hour, then stopped by placing on ice for at least 15 minutes. 2 uL of reaction was mixed with 40 uL of LAR buffer (Promega) and the luciferase output was measured using a GLO-MAX Luminometer. For trans-inhibition reactions, experiments were performed as described before (3,134). Briefly: 100x of the inhibitor PTE was added to .2 pmol reporter RNA (CA luc or PMV+lucPMV) in 20 uL reactions scaled up preserving the ratios of components described above.

Rabbit reticulocyte lysate:

Rabbit reticulocyte lysate is commercially available from Promega. Manufacturer’s protocols were followed. Briefly, final reaction components: 12.5 uL rabbit reticulocyte lysate, 0.5 uL kit-provided amino acids –Leu (1 mM), 0.5 uL kit-provided amino acids –Met (1 mM), 1 uL RNasin, 2 uL diluted RNA (final concentration 8 nM), final volume brought up to 50 uL by adding RNase-free water. Reactions were set up on ice, incubated at 30° C and stopped by placing on ice for at least 15 minutes. Translation reactions were read by mixing 2 uL of reaction with 40 uL of LAR buffer (Promega) and reading with a GLO-MAX Luminometer.

HeLa cell extract translation: translation extract was obtained from Aurelie Rakotondrafara (Plant Pathology Department, University of Wisconsin). Final reaction: 4 uL HeLa cell lysate, 1 Supermix [final concentration: amino acid mixture (2 mM), ATP (100 mM), GTP (10 mM), HEPES (1 M), Creatine Phosphate (1 M), Spermadine (5 mM)], 1 uL Creatine Kinase, optimized volume of Potassium acetate (1 M), 0.2 uL Magnesium acetate (100 mM)], 1 uL diluted RNA (final concentration of 8 nM), final volume of 10 uL reached by adding nuclease-free water (188). Reactions were set up on ice. Translation reaction was incubated for 1 hour at 37° C,
stopped on ice for 15 minutes, and read with a GLO-MAX Luminometer [2 uL reaction added to 40 uL LAR (Promega)]. Trans-inhibition reactions were set up as described above.
CHAPTER 3: RESULTS AND DISCUSSION

**eIF4E sequence and structural alignments**

In 2009, it was discovered that the PTE binds eIF4E in plants (134). The high structural conservation of eIF4E led us to believe that the PTE may bind mammalian eIF4E, and therefore function in mammalian systems. To show this rationale, a MUSCLE (multiple sequence comparison by log-expectation) protein sequence alignment (178) of eIF4E from 20 different species spanning the eukaryotic kingdoms Protista, Animalia, Fungi, and Plantae was conducted in Jalview (179). To more easily visualize the relatedness of the protein, a neighbor-joined phylogenetic tree (181) was constructed from this alignment in MEGA (180). Finally, since sequence conservation does not necessarily translate to 3D structure conservation, a structural alignment was performed and the RMS values calculated in Swiss PDB Viewer and rendered in Pov-Ray (182, 183).

*Protein Sequence Alignment*

The MUSCLE protein sequence alignment algorithm was developed by Robert Edgar in 2004 as a way of mitigating the problem of aligning sequences accurately and quickly (178). This algorithm uses three steps. First, a guide tree is constructed of unaligned sequences using kmers. A kmer is a pre-defined sequence of residues (178). Instead of comparing pair-wise sequences residue by residue, the kmer method compares “words”, or sequences of residues (178). Sequences with more similarity contain more “words” that are alike. This reduces the time required to align sequences since it reduces the number of units being compared (178). A rough phylogenetic tree is constructed from the kmer-aligned sequences (178). Because this preliminary step is done quickly, there is space for errors to occur, so a second step is performed to refine the results (178). The Kimura distance is calculated using the aligned sequences and a
new tree is formed (178). Finally, edges of the trees are individually analyzed to optimize the sum of pairs score (178).

The MUSCLE alignment was chosen out of three other protein sequence alignments (Mafft, Probocon, and T-Coffee) because it gave the highest bootstrap values and a tree that made the most biological sense when a neighbor-joined phylogenetic tree was constructed (Fig. 4B).

Twenty eIF4E protein sequences were selected from NCBI (Table 1). These were selected due to their representation of not only all four kingdoms of Eukarya, but also due to their representation of the kingdoms of Animalia (especially the class Mammalia) and Plantae (both monocots and dicots were chosen) which are relevant to the central aim of this work. While the overall sequence identity of these aligned sequences is small, the cap-binding tryptophans which differ by species in amino acid position, are very highly conserved (highlighted in red and blue, Fig. 4A). Because the PTE is thought to bind with the cap-binding pocket by interacting with these two tryptophans, we hypothesized that the PTE would be able to interact with eIF4E outside of the plant world.

**Neighbor-joined Phylogenetic Tree**

To visualize how the eIF4E sequences are related to one another, a neighbor-joined phylogenetic tree (181) was constructed using MEGA (180). This tree uses an algorithm that performs pair-wise comparisons of alignments to determine the best tree based on a bootstrap value. The tree organizes the members of Kingdom Animalia together, with mammals all grouping closely together. Kingdom Plantae is also organized together with a separation between monocots and dicots. While Fungi are thought to be more closely related to animals due
<table>
<thead>
<tr>
<th>Scientific Name</th>
<th>Common Name</th>
<th>Kingdom</th>
<th>Accession Number (NCBI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Rattus norvegicus</em></td>
<td>Norway Rat</td>
<td>Animalia</td>
<td>NP_446426</td>
</tr>
<tr>
<td><em>Mus musculus</em></td>
<td>House Mouse</td>
<td>Animalia</td>
<td>BAC38660</td>
</tr>
<tr>
<td><em>Oryctolagus cuniculus</em></td>
<td>Rabbit</td>
<td>Animalia</td>
<td>NP_001095180</td>
</tr>
<tr>
<td><em>Bos taurus</em></td>
<td>Cow</td>
<td>Animalia</td>
<td>NP_776735</td>
</tr>
<tr>
<td><em>Pan troglodytes</em></td>
<td>Chimpanzee</td>
<td>Animalia</td>
<td>JAA39337</td>
</tr>
<tr>
<td><em>Homo sapiens</em></td>
<td>Human</td>
<td>Animalia</td>
<td>NP_001959</td>
</tr>
<tr>
<td><em>Xenopus laevis</em></td>
<td>African Clawed Frog</td>
<td>Animalia</td>
<td>P48597</td>
</tr>
<tr>
<td><em>Danio rerio</em></td>
<td>Zebrafish</td>
<td>Animalia</td>
<td>NP_571529</td>
</tr>
<tr>
<td><em>Drosophila melanogaster</em></td>
<td>Fruit fly</td>
<td>Animalia</td>
<td>NP_524829</td>
</tr>
<tr>
<td><em>Triticum aestivum</em></td>
<td>Bread Wheat</td>
<td>Plantae</td>
<td>P29557</td>
</tr>
<tr>
<td><em>Hordeum vulgare subsp. vulgare</em></td>
<td>Domesticated Barley</td>
<td>Plantae</td>
<td>AAV80393</td>
</tr>
<tr>
<td><em>Zea mays</em></td>
<td>Maize</td>
<td>Plantae</td>
<td>NP_001105612</td>
</tr>
<tr>
<td><em>Solanum lycopersicum</em></td>
<td>Tomato</td>
<td>Plantae</td>
<td>NP_001234459</td>
</tr>
<tr>
<td><em>Arabidopsis thaliana</em></td>
<td>Thale cress</td>
<td>Plantae</td>
<td>NP_193538</td>
</tr>
<tr>
<td>* Pisum sativum *</td>
<td>Pea</td>
<td>Plantae</td>
<td>AAR04332</td>
</tr>
<tr>
<td><em>Cucumis melo</em></td>
<td>Muskmelon</td>
<td>Plantae</td>
<td>ABD57970</td>
</tr>
<tr>
<td><em>Dictyostelium discoideum AX4</em></td>
<td>N/A</td>
<td>Protista</td>
<td>XP_647593</td>
</tr>
<tr>
<td><em>Plasmodium falciparum</em></td>
<td>N/A (causative agent of malaria)</td>
<td>Protista</td>
<td>XP_001351220 3D7</td>
</tr>
<tr>
<td><em>Candida albicans</em></td>
<td>N/A (causative agent of thrush)</td>
<td>Fungi</td>
<td>XP_719387</td>
</tr>
<tr>
<td><em>Saccharomyces cerevisiae</em></td>
<td>Baker’s Yeast</td>
<td>Fungi</td>
<td>NP_014502</td>
</tr>
</tbody>
</table>

Table 1 eIF4E Protein Sequence Accession Numbers.

eIF4Es of 20 species spanning all eukaryotic kingdoms were acquired by NCBI. Above are the species used in the protein sequence alignment and neighbor-joined phylogenic tree with their scientific name, common name, kingdom, and NCBI accession number.
to the presence of chitin in their cell walls, the eIF4Es of plants and animals were more closely related to each other than to those of fungi (Fig. 4B).

Structure Alignment of eIF4E

Because sequence conservation is not always a direct measure of structural and therefore functional conservation of proteins, a 3D protein alignment was performed using Swiss PDB Viewer (182). The crystal structures of human (accession number: 1IPC) and wheat (accession number: 2IDV) eIF4E and the NMR structure of yeast eIF4E (accession number: 1AP8) were chosen because they were the available structure representatives of three different kingdoms (other plant and animal representatives were omitted for simplicity and relevance).

In plants, there is a highly conserved disulfide bond that is reduced in reducing conditions allowing the cap to bind the cap-binding pocket (138). Under oxidizing conditions, this disulfide bond can form, resulting in a structure that is not able to bind cap analog (138). The wild-type wheat eIF4E (accession 21DR), shows this condition, however the C113S mutant (accession number 21DV) which prevents this disulfide bond from forming allows cap analog into the pocket under all conditions. This mutant was chosen for the overlay because it fits the overall pattern observed in other structures and because the disulfide bond was shown to be an artifact of crystallization (138). The mode NMR structure of yeast was chosen after amino acids 1–36 were omitted from the analysis (these form the N-terminal tail region which is highly variable and truncated in the crystal structures) to be overlaid with the crystal structure.

The proteins were overlaid by using the “iterative Magic fit” option which is able to directly compute the structural alignment of the proteins, rather than using a rough alignment option. The result (Fig. 4C-F) shows high structural conservation, especially in the cap-binding
pocket and in the orientation of the cap-binding tryptophans: W56 and W102 in human (139), W62 and W102 in wheat, (138), and W43 and W166 (137,141).

In agreement with the sequence alignment and the phylogenetic tree (Fig. A and B), the RMS values were smaller for the pairwise alignment between plant and human eIF4E at 0.86 Å. RMS values between human and yeast and wheat and yeast were similar at around 2.6 Å. RMS values close to 0 indicate high sequence homology. Because of the especially high structure homology of eIF4E, we sought to test if the PTE which is proposed to bind its cap-binding pocket, functioned to enhance translation in mammal systems.

**Aim 1: System Optimization**

The ultimate goal of this thesis was to characterize the ability of the PTE to enhance translation in mammalian systems. Preliminary studies by Dr. Zhaohui Wang in rabbit reticulocyte lysate and BHK-T7 cells showed that the PTE was capable of enhancing translation in mammalian systems (Wang, unpublished data). To investigate this further, it was necessary to conduct some experiments in plant lysates to compare the results obtained in mammalian systems to the PTE’s performance in the systems in which it evolved. Salt titrations, time courses, and concentration reactions were performed to optimize the PTE-containing construct, TPAVlucTPAV against the mutant TPAVlucTPAV-m2, and BlucB. The TPAV construct was chosen because it was the most efficient enhancer in Wang’s preliminary mammalian system translation studies. The m2 mutant is predicted to disrupt a key pseudoknot in the PTE and has been demonstrated to have a drastic reduction in activity in the PEMV PTE (3). Thus it should serve as a negative control. After these optimization steps, all-PTE translation studies were done to compare the translation efficiencies of the PTEs against each other in wheat germ extract. To test how well these PTEs bind the initiation factor (known to be eIF4E in plants (134)), trans-
Fig. 6 Reporter gene constructs. Two PTE luciferase reporter constructs were used. PTE luciferase reporter constructs designed and cloned by Dr. Zhaohui Wang include viral flanking regions and a firefly luciferase reporter (A). The CAluc control as described by Iizuka et al (Iizuka, 1994) (B).
inhibition studies including m2 mutants were conducted in wheat germ using three reporters: a capped Firefly luciferase reporter with flanking vector sequence, CAluc (184), and P2lucP2, as was done in Wang et al., 2009.

The constructs used for all of our studies include a firefly luciferase reporter from Promega flanked by the viral 5’ and 3’ UTRs of each virus. These and their respective mutants were designed and cloned by Dr. Zhaohui Wang (Fig. 6A). Diagrams of representatives of the PTEs showing the proposed pseudoknot and the m2 mutation are shown in Fig. 7.

**Potassium Salt Optimization:**

Optimal translation levels for capped RNAs usually occur at less than 80 mM of K+ ions (188-190). However, probably due to a requirement for proper folding of CITE elements (128), optimal concentration of K+ ions tends to be higher. For instance, the BTE of Barley yellow dwarf virus (BYDV) has a potassium salt optimum at 130 – 140 mM (128,129). To test the most efficient salt concentration of the PTE, three constructs were used: the PTE of TPAV, its m2 mutant, and the BTE of BYDV (using BlucB, a construct that contains the Firefly luciferase gene with flanking BYDV UTRs (129)). The BTE’s salt optimum is known in WGE, and was used as the positive control for the salt experiment. In WGE, these five concentrations were tested: 53 mM (the concentration that is default in commercially available wheat germ extract), 93 mM, 133 mM, 163 mM, and 203 mM (Fig. 8A). As expected, BlucB showed a parabolic curve with a maximum around 133 mM (128). TPAVlucTPAV showed a similar curve, also with a maximum of around 133 mM. Interestingly, mutant TPAVlucm2 exhibited a hyperbolic curve with a maximum around 50 – 100 mM after which translation activity rapidly decreases as potassium levels were increased. As this element shows a similar dependence on K+ ions as BlucB, we assumed that its Mg2+ ion concentration was similar to BlucB, at around 2 mM Mg2+ (Peterson,
Fig. 7: PTE Interactions.
Representative PTEs from three viral genera, Panicoviruses (A, B), Carmoviruses (C) and Umbraviruses (D) are labeled with bases of interest. Included are the C-rich (blue) and G-rich (orange) domains which may interact with each other to form pseudoknot (dotted bracket). The loop region of SL1 is proposed to interact with upstream regions in the Carmo- and Panicoviruses (green). To study the affect mutation has on potential base-pairing between the C-rich and G-rich domains, two bases are mutated in the C domain to As, known as the m2 mutant (solid bracket). The G proposed to interact with eIF4E in the same manner as the m7G cap is marked with an orange circle (Wang, 2011).
unpublished data and (128)). We used this data to perform future PTE experiments in wheat germ extract.

Because our main goal was to determine if the PTE enhances translation in mammalian systems, we chose to use the commercially available, rabbit reticulocyte lysate (RRL) to screen the PTE constructs. We also sought to expand on Dr. Zhaohui Wang’s trans-inhibition studies in this lysate. This was not the first time that a PTE-containing virus was tested in a translation reaction in RRL: in 1985, the Carmovirus, CarMV genome was shown to express proteins when added to the lysate (170), however salt, RNA concentration, and time course assays were performed to optimize the system for our constructs. Because the PTE had only been very briefly studied in rabbit reticulocyte lysates (Zhaohui Wang, unpublished data), we sought to optimize the system for potassium, similar to what we did in the wheat germ extract (Fig. 8A and B). Both BlucB and TPAVlucm2 showed curves with maximums around 79 mM. However, as salt concentration increased, their translation levels decreased rapidly. TPAVlucTPAV did not show the same trend as it did in wheat germ extract, but instead, translation remained largely steady between 79 mM and 139 mM and then decreased rapidly. Perhaps the wild type version of this element has a higher tolerance to higher salt conditions than do the other constructs, as opposed to having a salt optimum.

HeLa cell extract is derived from a human cervical carcinoma cell line and is a good system to test dependence of translation on cap-poly(A) synergy (101,188). Because of its sensitivity to cap-dependent translation, we reasoned that this extract would be a good lysate to test the degree to which the PTEs enhance translation. HeLa cell extract is also not commercially available or optimized for protein expression, and should give a relatively low background level of translation with optimization. Two controls, a capped and poly-adenylated luciferase reporter,
Fig. 8 Potassium optimization in three in vitro translation systems.
WGE (A), RRL (B), and HeLa lysate (C) were optimized for potassium. 8 nM TPAVlucTPAV, BlucB and TPAVlucm2 mRNAs were translated at the indicated potassium concentrations at 2.0 mM Mg for WGE, 0.5 mM Mg for RRL, and 2.5 Mg mM HeLa lysate.
CA luc (184) and a luciferase reporter containing the EMCV IRES were used to compare the PTEs against a) capped-poly(A) synergistic translation and b) a known, powerful cap-independent element (the EMCV IRES). BlucB was used as what I thought would be a negative control was, because it does not function in rabbit reticulocyte systems (Fig. 8B and(129)).

Initial studies sought to identify the optimum parameters in which the element functions. We first tested K+ concentrations. The results were interesting in that none of the constructs exhibited the same salt-dependent curves as in the other systems. TPAVlucm2 had an optimum at 90 mM, but was about 75% of the level of its wild type version at that concentration. TPAVlucTPAV had an optimum between 90 and 130 mM. This supports other results with these constructs with the TPAVm2 mutant functioning between 50 – 75% of the TPAV wild type element in HeLa cell extracts (Fig. 8C, Fig. 12D), and gel shift assays performed by collaborator Franck Martin of Institut de Biologie Moléculaire et Cellulaire in Strausbourg, France showing binding of both m2 and wild type element to human eIF4E (unpublished). BlucB surprisingly translated in this system, with its lowest translation levels exhibited at 50 mM and its highest at 170 mM (Fig. 8C).

The Thoma protocol (188) states that 2.5 mM Mg2+ is an optimal concentration to use for the HeLa system, but that this is RNA-dependent. To determine which Mg concentration was most appropriate for our PTE constructs, I performed a Mg2+ titration using 0.5, 1.0, 1.5, 2.0, and 2.5 mM Mg2+(Fig. 9). While five points were collected, they only caught half of the curve. In any case, 2 – 2.5 mM Mg2+ appears to be at the peak of the efficiency curves of the RNAs tested (BlucB, TPAVlucTPAV and TPAVlucm2). We continued to use 2.5 mM as our Mg concentration.
Fig. 9 Magnesium titration of HeLa lysate.

8 nM TPAVlucTPAV, BlucB as a control, and TPAVlucm2 mRNAs were translated at the indicated magnesium concentrations at 130 mM K conc. for 60 min.
Fig. 10 RNA Concentration.
RNA concentration titrations were performed with BlucB, TPAVlucTPAV and TPAVm2 in HeLa lysate with indicated concentrations, at optimal salt conditions.
**RNA Concentration Optimization:**

We also wished to optimize RNA concentration. A danger to using too much RNA in a reaction is that it becomes saturated and it isn’t possible to differentiate between translation levels because excess RNA can make up for deficiencies in translation efficiency. This is especially important when differences between constructs may be subtle. We tested final RNA concentrations at 40, 20, 10, 8 (what is traditionally used in the lab), and 4 nM. I found that none of them appear to be saturating (Fig. 10), as the curves do not level off.

**Time Course Experiment:**

To determine an appropriate time to incubate the samples in wheat germ extract, rabbit reticulocyte lysate, and HeLa cell lysate, a time course was run for four hours with time points being taken at every half an hour increment. The level of luciferase translated leveled off between an hour and an hour and a half in all constructs tested. Therefore, we determined that an hour was an appropriate time point to use to run our translation experiments in all systems (Fig. 11).

**Aim 2: Performance of PTE-containing luciferase reporters in lysates and cells**

*Translation Enhancement by the PTE in wheat germ extracts:*

We wished to know how well the PTE enhanced translation of a firefly luciferase reporter in comparison to the BTE from BYDV, a known translation enhancer in plants (129). Dr. Jelena Kraft and I tested PTEs from three panicoviruses (TPAV, CMMV, and PMV), five carmoviruses (CarMV, JINRV, HCRSV, PFBV, and SCV), and one umbravirus (PEMV) along with the m2 mutants of TPAV and PEMV in wheat germ extract and compared their relative efficiencies. In terms of enhancement, TPAV and PFBV translated most efficiently, at about 70 to 80% the
Fig. 11 Time Course.
A time course experiment was run at optimal salt conditions with 8 nM TPAVlucTPAV, BlucB, and TPAVm2 mRNAs, and time points were taken at the indicated intervals.
percentage of the BYDV BTE. The JINRV PTE gave the lowest yield of luciferase, at less than 10% of the BYDV BTE. The m2 mutation in both TPAV and PEMV abolished translation. HCRSV, CarMV, and PMV performed at around 50% as efficiently as the BYDV construct, and CMMV and SCV performed at about 20 – 30% (Fig. 12B). It cannot be ruled out that the different efficiencies of PTEs was related to their specific adaptation to monocot or dicot systems. Wheat germ extract is derived from monocots and TPAV, one of the viruses with the most efficient PTEs is a monocot virus. Conversely, JINRV is a virus of irises, a dicot, and it functions the least. Preliminary data by Jelena Kraft in BY2 extracts, a dicot cell extract derived from tobacco supports this hypothesis (Kraft, unpublished).

This experiment was repeated in rabbit reticulocyte lysates. Translation of PTE firefly luciferase RNAs at the default cap-dependent conditions (79 mM K+) was done. All experiments were normalized to the translation level of CarMVluc because CarMV had been previously shown to express protein in this extract (170). BlucB was used as a negative control because it had been shown to not function as an enhancer in RRL (129). Interestingly, at this salt concentration these two mRNAs and the m2 mutant of PEMV all translated similarly (Fig. 12C). TPAV functioned the most efficiently at about 6-fold higher than CarMV, followed by JINRV. One problem with this system is that it is highly optimized for protein expression, and all mRNAs translate efficiently in it, even without a 5’ cap. This makes it difficult to tell if the translation efficiency of the translation enhancers vary because they are all being translated relatively well. For this reason, we decided to use a different system: HeLa cell extract.

Translation experiments were performed at conditions recommended by the Thoma protocol (188). First, the PTEs were tested using a firefly luciferase construct in cis. The reporter gene (firefly luciferase ORF) was flanked by viral UTRs including the PTE. Results
Fig. 12 Cis Translation of PTE Reporters A schematic of the experimental design of cis translation is shown (A). Cis translation from three different systems is shown: WGE at 2.0 mM Mg, 133 mM K, and 8 nM RNA (B), RRL at 0.5 mM Mg, 70 mM K, and 8 nM RNA (C), and HeLa lysate (D) at 2.5 mM Mg, 50 mM K, and 8 nM RNA. All experiments were incubated for one hour.
show that at the low salt concentrations called for (2.5 mM Mg, 50 mM K), the PTE functioned at or above the level of the EMCV IRES. While the EMCV IRES is the strongest IRES known in mammalian systems, the comparison is difficult to make because the experiment as done at salt concentrations that were below the optimum for the IRES. Another interesting result was that the m2 mutant of TPAV functioned at about half the efficiency of wild type TPAV (Fig. 12D). While many of the constructs seemed to translate at lower levels, giving a clearer picture from the RRL results at this concentration, the mutants did still work, which in the case of TPAVm2 agrees with Franck Martin’s gel-shift data which showed binding of both wild type and TPAVm2 PTEs to human eIF4E.

Translation in mammalian cells.

Translation extracts, as in vitro systems have many limitations along with their advantages. While they are simple and easily manipulated, these characteristics can also lead to results that are not representative of what happens in whole cells. Because of this, we decided to test the PTE luciferase reporters in two cell lines: HeLa cells and T7-BHK (BSR-T7/5) cells. The latter had been used in a single experiment by my predecessor on this project, Dr. Zhaohui Wang. We used both DNA and RNA transfection methods.

Baby hamster kidney (BHK) cells are widely used in virus propagation, and have a deficiency in RIG-I, a key protein in the innate immune system that recognizes double-stranded RNAs – a signature of viral replication (191), that may explain why they are efficient host cells for viruses (186). T7-BHK cells (BSR-T7/5) are BHK-21 cells that have been stably transformed with a plasmid stably expressing the T7 polymerase (186). This cell line was first developed to express a cDNA clone of Bovine respiratory syncytial virus for study of that virus (185). As we do not have a CO₂ incubator, I took advantage of a collaboration formed by Dr. Zhaohui Wang
Fig. 13: Cell Translation of the PTE.
Translation in HeLa cells resulting from an RNA transfection incubated in 6-well plates at 37°C for three hours.
with Dr. Cathy Miller (Department of Veterinary Microbiology and Preventive Medicine), who studies reovirus replication in multiple mammalian cell lines. We used the fact that the cells are transformed with the T7 polymerase, which produces uncapped transcripts, to our advantage and transformed the DNA expressing our PTE constructs into the cells. After 24 hours, the results were similar to those obtained by Dr. Zhaohui Wang (unpublished) in that the TPAV PTE luciferase reporter yielded about half as much luciferase as an EMCV luciferase reporter.

After discussing these results with Dr. Aurélie Rakotondrafara, it was brought to our attention that cytoplasmic capping of the T7 transcripts is possible (192). It was found that cleavage products of hepatitis B virus produced by activation of RNase H had been capped after the authors first noticed that the transcripts were very stable (192,193). The substrate of the cytoplasmic capping enzyme is a 5’ monophosphate at the 5’ end of the RNA message (192). As the mRNA transcribed off the plasmid that was transfected is uncapped, and a degradation product of RNA is 5’ monophosphorylated, we could not rule out the possibility that some of the transcripts had been capped in the cytoplasm. To test if the BHK-T7 results had been significantly affected by the possibility of expression of a cap, we chose to do RNA transfections of HeLa cells. While the existence of cytoplasmic capping may occur in these cells as well, the incubation time for RNA transfections is much less (3-6 hours versus 18-24 hours), and we capped our RNAs with the A-cap (G[5’]pppA) which does not bind eIF4E or stimulate translation, to increase stability while avoiding interaction with eIF4E. Thus, A-capped transcripts must translate cap-independently (101).

Preliminary results comparing the positive control, CAluc with reporters containing PTEs show that the TPAV PTE translates between 10 and 15% as efficiently as the CAluc (capped polyadenylated) control with other PTEs translating at or below 5% (Fig. 13). Differences in cell
density in the wells could account for this difference because the experiment was performed without a Renilla luciferase reporter control to normalize translation levels.

**Aim 3: Performance of the PTE as an inhibitor**

If the PTE is functioning to enhance translation by binding an initiation factor, then adding PTE RNA in trans to the translation of a reporter mRNA will decrease translation levels of the reporter. This is because the PTE will bind the initiation factor, making it unavailable for use in translation. To determine the efficiency at which different PTEs from various viruses of the Panico- and carmoviruses, and one umbravirus inhibit translation of a reporter RNA, P2lucP2 (3) and compare them to their respective mutants, Dr. Jelena Kraft and I performed trans-inhibition experiments. PTE RNA was added at 100x the concentration of reporter RNA and the level of inhibition was compared to the level of translation of the reporter without inhibitor added. In all cases in which a mutant was available for comparison, the wild type PTE inhibited translation better than its respective mutant. For PMV, it inhibited translation to about 15% of the uninhibited reporter, while its mutant inhibited to 70%, for SCV, the wild type RNA inhibited to about 30%, while its mutant inhibited to about 45%. TPAV inhibited at about 35% and its mutant at around 80%, and PEMV inhibited translation at about 35% and its mutant at about 90% (Fig. 14B). Nonspecific RNA interactions could explain variable translation inhibition between mutants.

This experiment tested the PTE’s ability to inhibit translation that occurs through eIF4E recruitment by the PTE mechanism. We also sought to test the efficiency with which the PTE could inhibit canonical translation that occurs via eIF4E recruitment by capped RNA (Fig. 14C). We used capped, polyadenylated CA1uc RNA that lacks any viral sequence(184) as reporter RNA, and repeated the trans-inhibition experiment. Interestingly, wild type PTEs were also able
A schematic of the design of a trans-inhibition experiment is shown (A). Trans-inhibition experiments from two systems including HeLa (B) and WGE (C) and (D). Two reporters were used for the WGE experiments, a capped polyadentylated reporter (Iizkuka, 1994) (C) and P2lucP2 (D) (Wang, 2009; Wang, 2011).

Fig. 14: Trans-inhibition of a reporter by the PTE
to inhibit translation by this capped RNA. The strongest inhibitor was the PMV PTE which inhibited to a level of about 10% of the uninhibited CAluc (Fig 14). All of the wild type PTEs inhibited the translation of CAluc at 25% or less except for JINRV and PFBV. As expected, the mutant PTEs did not inhibit.

While the PTE could perform efficiently in cis in HeLa cell extracts, it did not inhibit efficiently in trans in this system (Fig. 14D). A trans inhibition experiment like the one described for the wheat germ experiments was set up using P2lucP2 as a reporter. None of the PTE constructs functioned to inhibit translation below 50%. In contrast, Zhaohui Wang showed previously that many PTEs inhibited efficiently in rabbit reticulocyte lysate (unpublished data). This cannot be explained by relative levels of eIF4E in HeLa extract and reticulocyte lysate. eIF4E is limiting in HeLa cell extract and present in excess in rabbit reticulocyte lysates (194,195). Rather, these results may be explained by differences in conditions between the two experiments such as relative concentration between inhibitor and reporter and salt concentrations.
CHAPTER 4: CONCLUSIONS

Relevance

The study of viruses provides an elegant and simpler model for studying new phenomena in molecular biology than cell models, and as such, the characterization of virus structural elements have led to novel discoveries in cellular systems. For instance, the viral mRNAs of vaccinia and reovirus were among the first to be discovered to have a 5’ cap before it was accepted as the rule rather than the exception that eukaryotic mRNAs carry this 5’ modification (196-198). Alternative methods of translation such as IRES-mediated initiation were also characterized in viruses first (94). From a basic science perspective, the characterization of the PTE follows a similar course. Interestingly, an eIF4G-binding CITE was recently identified to function in mammalian cell extracts to boost translation (199).

The characterization of CITEs also has applied science relevance. Understanding of virus-host interactions has led to the development of resistant crops. This is especially true in the development of melon varieties with mutated eIF4E that are no longer susceptible to certain viral disease (123,153). Understanding which sequences or structures are interacting with host proteins is important to design crops that no longer contain adequate targets for viral interaction. Characterization of a CITE by Nicholson et al. provided interesting insight into the possible reasons for resistance that previously designed crops enjoyed (123,153).

As discussed by Terenin et al., understanding alternative methods in translation initiation and control provides scientists a “new paradigm” with which to work (199). The bridge between basic science and human health can also be bridged by understanding the importance that eIF4E has in human disease. As discussed before, eIF4E regulation has far reaching consequences. The PTE, as an element that interacts with eIF4E through a novel mechanism (3) expands on
Terenin’s idea of a paradigm shift and may open the door to discovering new ways to control eIF4E.

**Proposed model**

While further investigations need to take place, the PTE shows promise as at the least, a weak translation enhancer in mammalian systems. It is difficult to compare EMCVluc to the levels of translation in HeLa lysate because the optimal concentration for this IRES is higher than what was used (200). While the m2 mutants very clearly abolished translation in plants and reticulocyte lysates at higher salt concentrations (133mM K+), surprising results in HeLa lysates revealed that the m2 mutant does function at 50% the efficiency of the wild type, and this was corroborated by gel shift data by our collaborator, Franck Martin. Perhaps the PTE is not binding the eIF4E through a pseudoknot-mediated contact (3) in a human system (though it is clear that the interaction is important in plants) or more sequence is needed to visualize more drastic translation changes (such as the additional sequence Anne Simon proposes is important in carmovirus and umbravirus translation (132,136)). Additionally, it is known that HeLa lysates contain ITAFs that are important in IRES function such as the La protein (99). These are not present in rabbit reticulocyte lysate and may serve to stabilize the PTE even with a disrupted pseudoknot and still recruit factors (94,99). Other than human eIF4E (Franck Martin), it has not yet been tested to which proteins the PTE binds in mammalian systems, and this could be tested by performing a pull-down experiment in lysates.

**Future directions**

Some future directions for this project should include mutation of the hypermodifiable G in the G-rich domain of the PTE to determine if it is interacting with eIF4E in the hypothesized
way through π-π interactions with the eIF4E tryptophans (3). As the interaction between the cap binding tryptophans is very specific

Another way to test if the PTE can compete with the cap for eIF4E binding is to titrate cap analog into the translation reaction and observe the effect this has on PTE-mediated translation in relation to the EMCV IRES which does not require eIF4E to function, and a capped and polyadenylated control RNA (199). If eIF4E interacts with the PTE in a similar way as it does mRNA, then translation levels of the reporter mRNA should mirror those of the capped, polyadenylated control – they should both decrease as the cap analog concentration decreases. However, if eIF4E is not required for PTE function in mammalian systems, then its translation should follow the trend set by the translation of reporters controlled by the EMCV IRES: they should not be affected dramatically by increasing cap analog concentrations (199). To test the specificity of the interaction, GTP would be used as a control.

Preliminary gel shift assays that have been performed by our collaborator Franck Martin in the lab of Gilbert Eriani have shown that the PTE interacts with both wheat and human eIF4E. While m2 mutants were compared, hypermodifiable G mutants have not been, and may be more critical than the pseudoknot mutants to testing the interaction hypothesis. It would also be interesting to ascertain a Kd for the PTE-human eIF4E interaction. While the PTE may bind mammalian eIF4E, it did not evolve in this system, and it would provide a more complete picture to see how far apart the Kd for the mammal eIF4E-PTE and the plant eIF4E-PTE differ. Ultimately, we would like to explore host RNAs of both plants and mammals to ascertain the presence of PTE or PTE-like translation enhancers. With the importance that control of regulation of initiation factors has on the health of the cell, it would be interesting to see if this is, indeed a paradigm shift as the discovery of IRESes was.
APPENDIX 1: 3’ TRUNCATION OF PTE CONSTRUCTS

The 5’ UTR of the PTE-containing viruses is extremely short: less than 20 bases in most cases. Most cellular UTRs are structured and longer, and require the helicase activity of eIF4A for ribosome scanning to occur. In the absence of long, structured UTRs, the ribosome is able to scan to the start codon and initiate translation without the help of eIF4A and other initiation factors involved in cap-dependent translation initiation (33). To determine how much of the enhancement effect of the PTE was due to this 5’-scanning effect, I truncated the PTE constructs, digesting them at the XbaI site, right after the luciferase gene and added them to rabbit reticulocyte lysate translation reactions with 0.5 mM Mg2+, and 0.79 mM K+ concentrations. Surprisingly, the constructs translated very well without the 3’ UTR, with PMVlucXbaI translating at about 80% of PMVlucPMV, and TPAVlucXbaI and CMMVlucXbaI showing similar results to PMV in relation to their full-length constructs (Fig. 15). This shows that under these lower salt conditions, 5’ scanning constitutes a significant portion of the translation of this construct in rabbit reticulocyte lysates.

Interestingly, when the experiment was repeated in wheat germ, under slightly salt conditions (~90 mM K+), the truncated TPAVlucXbaI construct translated at only 25% of the efficiency of TPAVlucTPAV.
Fig. 15: Truncations.
Truncated PTE luciferase constructs cut with XbaI were tested in WGE (A) and RRL (B) using default salt conditions.
HeLa cell extract preparation was attempted using two different protocols from Matthias Hentze’s lab (187,200). Our lab lacks access to a bioreactor to grow large quantities of liquid culture-adapted cells. Therefore, I attempted to use a hybrid between Thoma’s protocol which uses liters of cell culture and Rakotondrafara’s protocol which is designed to prepare small amounts of extract from cells that have been treated with interfering RNAs. To start, I used eight 150 cm\(^2\) flasks. However, this yielded only a small pellet (less than 400 uL) and was not worth the labor and materials. I then contacted the Cell and Hybridoma facility and was able to grow two 1700 cm\(^2\) roller bottles of cells. As per advice by the staff there, I deviated from the protocol and did not use trypsin, but rather a chelator-based non-enzymatic cell remover to remove the cells. To ensure that no chelator was present, I rinsed the cells twice with PBS and twice with Hank’s Balanced Salt Solution. I prepared the lysis buffer as described in Rakotondrafara, 2011. I found that using a small needle to lyse the cells was dangerous, so I opted to use a dounce homogenizer. I had difficulty lysing the cells (fig. 16). When tested, the extract was not functional (data not shown). Future improvements could include not using the chelator-based cell remover, and also douncing for less time (no more than 20 strokes).
**Fig. 16 HeLa Cell Lysate.**
Preparation of lysate was attempted. HeLa cells at 80% confluency (A) are shown, along with cells after trypsinization and dying with trypan blue (B). Figures C-F show different stages during lysis with a dounce homogenizer. Lysis is obtained when there are few round cells left (F).
APPENDIX 3: EFFECT OF POTASSIUM IONS ON BTE TRANS INHIBITION

Because CITE function depends on their ability to faithfully adopt structure, we sought to explore the role that salt concentration has on structure and function. We investigated mutations made in the four way junction region of the BTE and Jelena Kraft showed that as salt concentration increased, the wild type and mutant BTEs adopted more compact structures using native polyacrylamide gel electrophoresis (128). To test if these structural results correlate with function as salt increases, we tested the ability of BTEs from three different viruses: Barley yellow dwarf virus (BYDV), Tobacco necrosis virus-D (TNV-D), Red clover necrotic mosaic virus (RCNMV) to trans inhibit a capped reporter (184) at two different salt conditions (53 mM and 133 mM K). As a control, human 18S rRNA was used to determine the level of nonspecific trans inhibition.

Under “low” salt conditions, ie, 53 mM K, the BYDV BTE inhibited at about 40 - 50% of the non-inhibited control, with the other BTEs inhibiting at much less levels. However, under the “high” salt condition, 133 mM K, the BYDV BTE nearly abolishes translation, with TNV-D and RCNMV at around 50%.

These results support Kraft’s structural data in a biologically relevant context and were included as Fig. 4B in Kraft et al.’s 2013 Nucleic Acids Research paper and is shown in Fig. 17.
**Fig. 17 Potassium ion effect on BTE trans inhibition of BTE translation.**

Trans inhibition performed at two salt concentrations with 8 nM reporter construct and inhibitor at 200X. BTE inhibitors shown (adapted from Kraft et al., 2013).
REFERENCES CITED


70


ACKNOWLEDGEMENTS

First and foremost, I’d like to thank my advisor Dr. Allen Miller for giving me the freedom and having the confidence in me to be able to work on such a challenging, but interesting project. I thank him for his guidance, wisdom, support, and the growth he has facilitated in me as a scientist. I’d also like to thank my committee members, Dr. Susan Carpenter and Dr. Mark Hargrove for their constructive criticism and for the time that they spent guiding me to becoming a better focused scientist.

I would also like to thank my lab mates, Dr. Jelena Kraft, Dr. Jimena Carrillo-Tripp, Dr, Krzysztof Treder, Nikki Krueger, and Alice Hui. Thank you all for the enlightening discussions, the emotional support, and for helping me to grow as a better person.

Lastly, I’d like to thank my teachers, my family, and my partner AJ. Thank you for having the faith in me that I sometimes didn’t.