THE DISSERTATION COMMITTEE FOR NICOLA ANN COLE CERTIFIES THAT THIS IS THE APPROVED VERSION OF THE FOLLOWING DISSERTATION:

Identifying the Role of the Cap-binding Complexes in the Regulation of Translation in *Arabidopsis thaliana*

Committee:

Karen S. Browning, Supervisor

David Hoffman

Edward Marcotte

Rick Russell

Claus Wilke

Identifying the Role of the Cap-binding Complexes in the Regulation of Translation in *Arabidopsis thaliana*

BY

NICOLA ANN COLE

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

DECEMBER 2017

DEDICATION

To my beloved husband, Robert,

and my children, Nathan, Nicholas, Bonnie and Robert

ACKNOWLEDGEMENTS

I would like to first and foremost acknowledge and thank my advisor, Dr. Karen Browning, for her kind and gracious advice and guidance over the course of my research and writing of this dissertation. I have learned much from her and her dedicated enthusiasm and easy-going manner has been a blessing to work with. I would also like to thank my committee members, Dr. David Hoffman, Dr. Edward Marcotte, Dr. Rick Russell and Dr. Claus Wilke. I have appreciated the energy they have put into our committee meetings, their input and their encouragement. I also thank Dr. Julia Bailey-Serres for so graciously allowing me to come and work in her lab at UC Riverside, and for her valuable input. Thank you also goes to Dr. Maureen Hummel and graduate student Lauren Dedow for their sacrifice of time and effort in our collaboration. Very special thanks go to my sister Dr. Kathy Newton-Ishmiel, my daughter and PhD candidate Bonnie Cole and Dr. Jessica Momb. Words cannot express how valuable their input and encouragement over the years has been to me. Dr. Gail Grabner has been another source for encouragement and support I am so grateful for. I would also like to thank my lab members, Dr. Andrew Lellis, Dr. Ryan Patrick and Laura Mayberry for all of their input and help over the course of my research. They have been invaluable. Thank you too for the help of undergraduates Peyton Busby and Hailey Van Hoorn, and from the BioBricks FRI, Dr. Soo-Hyun Yang for her help and support. And to Claire McWhite for her help with the kinase project. I would be remiss in not thanking Dr. Dean Appling. I would not have come this far if it were not for his longsuffering, patience and encouragement while working for him as an undergraduate. His guidance and mentoring has served me well during this experience. Thank you too to

Dr. Anne Tibbets for her sweet attitude, her effort and enthusiasm in mentoring me as an undergraduate.

Most of all, I am thankful for such a loving and supportive husband who has sacrificed much while I have pursued this path. His encouragement and support has been invaluable to me and I am so grateful. And for my children, Nathan, Nicholas, Bonnie and Robert. They have also sacrificed much over the years in my pursuit of this degree. They are my joy and my life and I am so thankful for them. I hope they are inspired to continue to challenge themselves academically and go after their pursuits and growth in knowledge with passion and vision. Above all, I am thankful to God for His grace and strength to help me accomplish this.

Identifying the Role of the Cap-binding Complexes in the Regulation of Translation in *Arabidopsis thaliana*

Nicola Ann Cole, Ph.D.

The University of Texas at Austin, 2017

Supervisor: Karen S. Browning

There is a fundamental gap in our understanding of the regulation of translation in plants. None of the canonical eukaryotic methods found in mammals, such as the use of 4E-Binding protein phosphorylation and inhibition of eIF2B function by phosphorylation of eIF2α, have been found in plants to date. In the very early steps of eukaryotic initiation of translation, the cap-binding complex, eIF4F, binds to the m^7G of the mRNA and positions it for binding by the 43S preinitiation complex. This complex consists of the large scaffolding protein, eIF4G, and the smaller cap-binding protein, eIF4E, along with the DEAD box helicase eIF4A and the RNA binding protein eIF4B. In addition to eIF4F, plants have a well-conserved second cap-binding complex, eIFiso4F, not found in other eukaryotes and comprised of eIFiso4G and eIFiso4E, isoforms of the scaffolding protein and the cap-binding protein, respectively. *Arabidopsis thaliana* plants lacking *eIFiso4F* have stunted growth, reduced chlorophyll, reduced fertility and are more sensitive to light stress. In *in vitro* translation assays, there is a preference for eIF4F by capped messages. In this work, several possible methods for the regulation of translation in *Arabidopsis thaliana* by eIF4F and eIFiso4F are investigated. A previously discovered eIF4G wheat kinase known to phosphorylate eIF4G, eIFiso4G, eIF4B and a 60S ribosomal protein is

further characterized to determine its possible role in the regulation of translation. It is found to be a calcium dependent kinase that phosphorylates eIFiso4G in the C-terminal domain most likely at a consensus sequence for calcium dependent protein kinases. Additionally, there is a conserved pair of cysteines located in the cap-binding pocket of both eIF4E and eIFiso4E, found only in plants. In an *in vitro* labeling assay, these cysteines are tested for their thiol reactivity in eIF4E versus eIFiso4E. eIFiso4E is found to be more thiol reactive than eIF4E, possibly indicating differences in redox activity between the two proteins. Finally, two methods for identifying specific populations of mRNA associated with plant protein synthesis initiation factors eIF4F and eIFiso4F are described, RNA immunoprecipitation and ribosome profiling.

TABLE OF CONTENTS

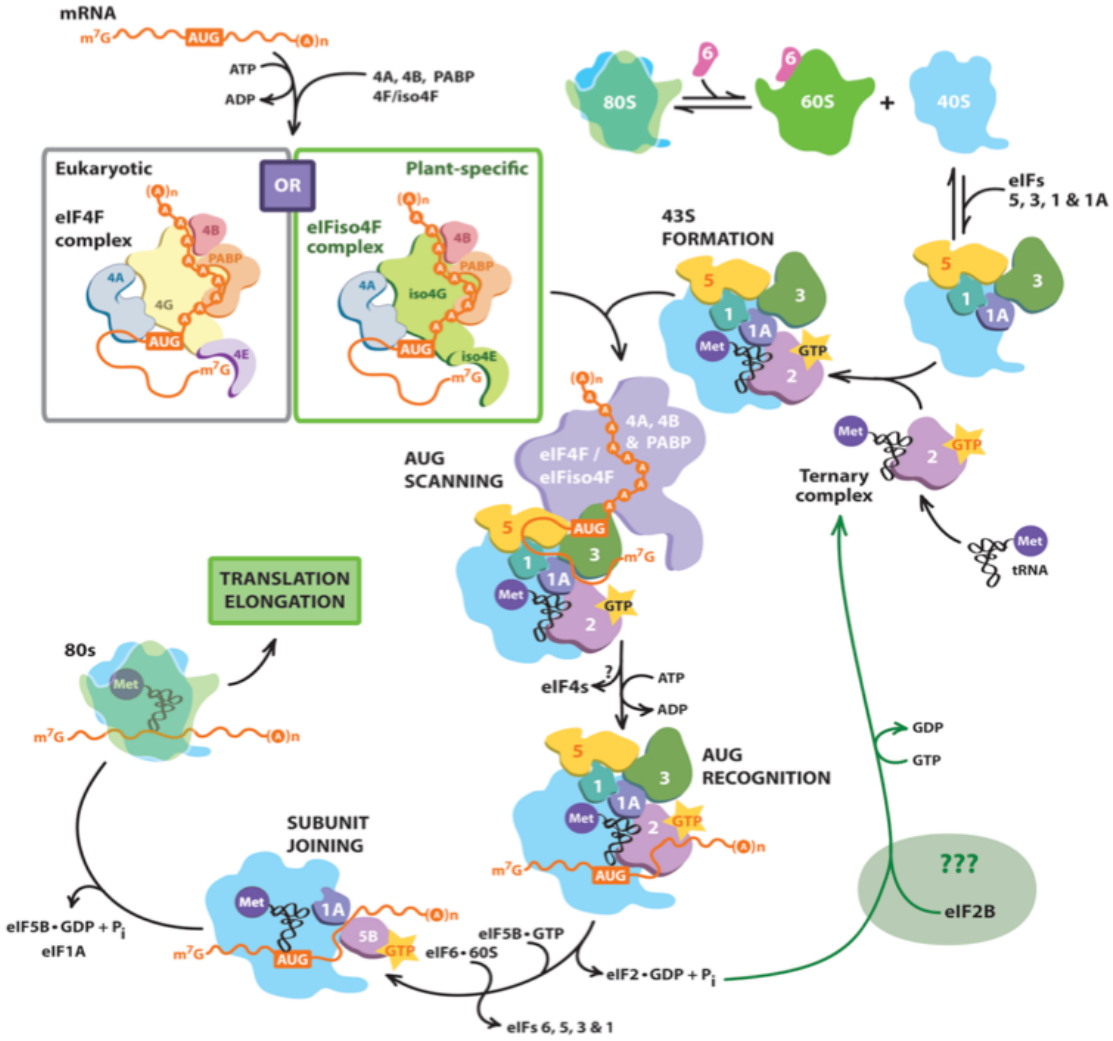
LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 Overview of Translation

The path to protein synthesis within the eukaryotic cell starts with the conversion of information contained within the DNA in the nucleus into the messenger RNA (mRNA) that is sent to the protein making machinery in the cytoplasm. mRNA is the vehicle from which the information is "read" by the highly-conserved ribosome and translated into protein (Browning & Bailey-Serres, 2015; Sonenberg, 1994). This "reading" takes place in three steps: initiation, the preparatory steps leading up to the start of protein synthesis, elongation, the joining together of the amino acids to form the polypeptide, and termination, the end of protein synthesis and the dismantling of the protein synthesis machinery.

The initiation of translation is the bringing together of the necessary factors required to assemble and prepare the final initiation complex in order for the elongation step to begin. Prokaryotic initiation of translation requires only three single-polypeptide initiation factors in order to form the final 70S initiation complex from which protein synthesis begins (Sonenberg & Hinnebusch, 2009). In eukaryotes, these steps are carried out by at least 12 eukaryotic initiation factors (eIFs) made up of more than a total of 30 polypeptides in a very coordinated and complicated fashion (Figure 1.1) (Browning & Bailey-Serres, 2015; Jackson et al., 2010; Sonenberg & Hinnebusch, 2009). The plant translational machinery is similar to that found in mammals and *S. cerevisiae* (Browning & Bailey-Serres, 2015). In the very early steps of eukaryotic initiation, the mRNA is bound at its cap, a 7-methyl guanosine (m^7G) that is bound to the 5' end of the mRNA in a 5' to 5' triphosphate linkage, by the cap-binding complex (eIF4F).



From (Browning & Bailey-Serres, 2015)

Figure 1. 1 **Schematic of the initiation of plant translation.** This schematic depicts the steps leading up to the start of protein synthesis. It begins with the disassembly of the 80S ribosome into its 2 subunits. This allows re-formation of the 43S preinitiation complex (PIC). The PIC is then recruited by eIF4F, the cap-binding complex, which has the mRNA bound to it. The plant specific isoform of the eIF4F complex, eIFiso4F, is featured in the green box. After recruitment, the small subunit scans for the start codon on the mRNA. Once the AUG has been located, the large ribosomal subunit joins the complex, with a loss of several factors. With this final formation of the 80S initiation complex, translation is now ready to begin.

(Aitken & Lorsch, 2012; Browning & Bailey-Serres, 2015; Shatkin, 1976). This mRNA/cap-binding complex then recruits other eIFs such as eIF4A and eIF4B and binds to the 43S preinitiation complex made up of the ternary complex (eIF2•initiator $tRNA^{Met}$•GTP), eIF3, eIF5, eIF1 and eIF1A (Browning & Bailey-Serres, 2015; Hinnebusch & Lorsch, 2012; Jackson et al., 2010). The correct positioning of the start codon on the mRNA in the 40S ribosomal subunit, followed by the release of P_i, brings about the formation of the 48S complex and the 60S ribosome is now able to bind. The newly formed 80S initiation complex (IC) is now ready for protein synthesis to take place.

While the three prokaryotic initiation factors (IFs), IF1, IF2 and IF3, are structurally and/or functionally similar to eIF1A, eIF5B and eIF1, respectively, prokaryotes and eukaryotes use very different mechanisms for locating the start codon (Browning & Bailey-Serres, 2015; Sonenberg & Hinnebusch, 2009). In prokaryotes, a segment of the 16S rRNA of the prokaryotic small ribosomal subunit base pairs with the conserved Shine-Dalgarno sequence on the bacterial mRNA, which is located about 10 nucleotides upstream of the start codon. This alignment leads to the correct positioning of the start codon in the ribosome, ready for the initiator tRNA to bind. In eukaryotes, after recruitment by the eIF4F/mRNA complex, the 43S preinitiation complex starts a scanning (5' to 3') process, in an ATP-dependent fashion, for the start codon in the 5' untranslated region (5' UTR) of the mRNA. Although the AUG context can vary depending on the organism, the Kozak sequence bias of -3A/G and +4G is valid for plants along with a bias for +5C and -2A/C, giving higher order plants a general consensus sequence of GCNAUGGC (Gupta et al., 2016).

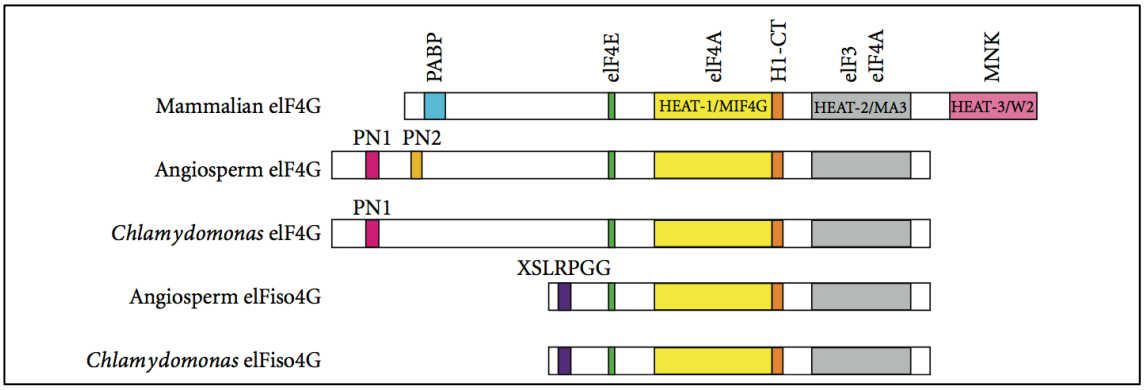
1.2 The Cap-Binding Complex, eIF4F

eIF4F, which binds the mRNA and recruits the 43S PIC, is composed of eIF4G, eIF4E and eIF4A in mammals and yeast, although only eIF4G and eIF4E are considered part of the complex in plants, as eIF4A does not co-purify with the complex (Mayberry et al., 2011). eIF4G is a large scaffolding protein onto which the other eIFs such as eIF4E, eIF4A, eIF4B and poly-A binding protein (PABP) bind to. Its molecular mass is 168 kDa and contains various motifs for protein interactions. eIF4E, a much smaller protein with a molecular mass that is approximately 24 kDA, binds to the "cap" of the mRNA, located at the 5' end of the mRNA. eIF4E and eIF4G form a strong complex with subnanomolar affinity (Mayberry et al., 2011). eIF4A is a DEAD box RNA helicase that is involved in unraveling the mRNA in preparation for its scanning by the ribosome, for which eIF4B, an RNA binding protein, enhances the helicase activity of eIF4A (Hinnebusch & Lorsch, 2012; Rozovsky et al., 2008).

1.2.1 eIF4G

Judging from available plant genome data, eIF4G is well conserved in plants (Patrick & Browning, 2012). Plant eIF4G contains most of the same C-terminal core structure as mammalian eIF4G, having the eIF4E, H1-CT, HEAT-1/MIF4G and HEAT-2/MA3 domains but lacking the third HEAT-3/W2 domain, which in mammals binds MNK1 (MAP kinase activated protein kinase), a kinase known to bind eIF4G and phosphorylate eIF4E (Figure 1.2) (Korneeva et al., 2016; Lellis et al., 2010; Patrick & Browning, 2012). Plant eIF4G also has the long N-terminus found in mammalian eIF4G that also possesses a PABP binding site similar to the mammalian eIFG, but is considerably longer in sequence, 203 bp in plants versus 29 bp in mammals (Figure 1.2, showing only the mammalian PABP binding domain) (Cheng & Gallie, 2013). The wheat PABP binding site also overlaps with a eIF4B binding site, with eIF4B and PABP binding competitively

with each other in this N-terminus domain (Cheng & Gallie, 2013). In wheat eIF4G, PABP and eIF4B also have overlapping binding sites with a small portion of the HEAT-1 C-terminal domain, binding competitively with each other but not competitively with eIF4A (Cheng & Gallie, 2013).



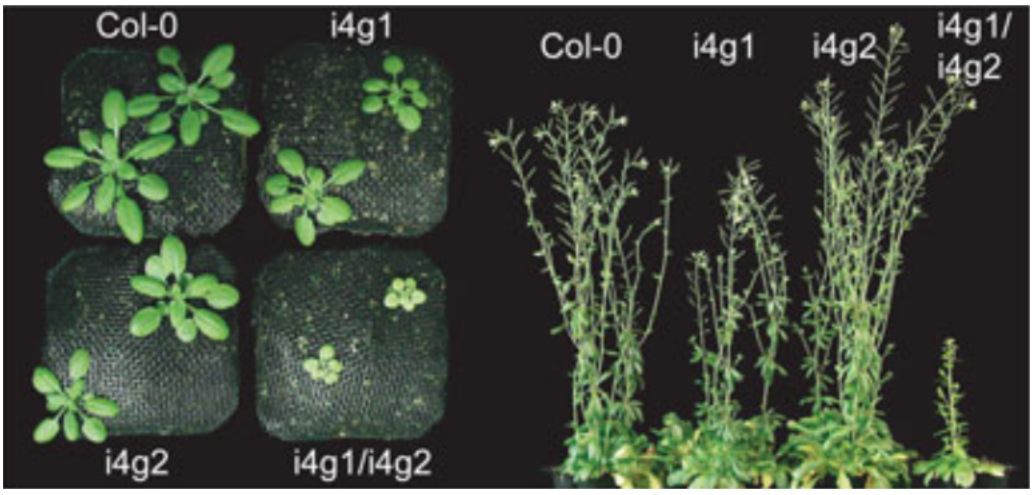
From (Patrick & Browning, 2012)

Figure 1.2 **Schematic comparing the domain organizations of eIF4G and eIFiso4G in Mammals, higher order plants (Angiosperm) and lower order plants (Chlamydomonas).** In common between the two eIF4G isoforms and between all organisms are the eIF4E, eIF4A and eIF3 binding domains, the H1-CT motif and two of the three HEAT domains seen in Mammalian eIF4G, MIF4G and MA3. The third HEAT domain in mammals, W2, is found in an extension of the C-terminal sequence not present in plants. Not shown is a PABP binding domain in plant eIF4G as is seen in mammalian eIF4G, but is considerably larger than the mammalian PABP binding site and serves also as a competitive binding site for eIF4B. eIF4G and eIFiso4G also bind PABP and eIF4B in the HEAT 1 domain. Plant eIF4G contains an N-terminal extension of amino acids not seen in mammals, with higher order plants possessing two very conserved motifs, PN1 and PN2, intervened by poorly conserved sequences. Lower order plants contain only either the PN1, as shown here, or PN2 motif. Found only in the plant specific eIF4G isoform, eIF4iso4G, and conserved across all plants, including green algae, is the sequence XSLRPGG found in the N-terminal domain.

Also found in the N-terminus, making the plant eIF4G N-terminus longer than the mammalian eIF4G N-terminus, are two conserved plant specific motifs, plant eIF4G N-terminal motif 1 and 2 (4G-PN1 and 4G-PN2) (Patrick & Browning, 2012). In Arabidopsis, 4G-PN1 and 4G-PN2 consist of 17 amino acids and 15 amino acids, respectively, with 4G-PN1 appearing 180 amino acids into the sequence and 4G-PN2 at 375 amino acids (Patrick & Browning, 2012). Outside of these two motifs in the N-terminal region, the sequences are quite variable between plant species. The function of these two sequences and of the N-terminal region is unknown with deletion of a large portion of the N-terminal region having minor effect on *in vitro* translation (Patrick & Browning, 2012). Evidence of these conserved sequences is observed back to the beginning of land plant evolution. This amount of conservation in plant eIF4G implies possible importance for the sequences such as a regulatory role or some other function such as protein-protein interactions (Patrick & Browning, 2012).

Plants differ from all other eukaryotes in that they contain a second cap-binding protein complex, eIFiso4F, consisting of eIFiso4G and eIFiso4E (Figure 1.1) (Browning et al., 1992). Most angiosperm, or flowering plants, have more than one eIFiso4G gene (Patrick & Browning, 2012). Arabidopsis has two genes, eIFiso4G1 and eIFiso4G2. eIFiso4G differs from eIF4G in missing a significant portion of the N-terminal sequence of eIF4G yet retain the binding sites and motifs that eIF4G has, including overlapping PABP and eIF4B binding sites within the HEAT-1 domain (Figure 1.2) (Gallie, 2016; Lellis et al., 2010; Patrick & Browning, 2012). Arabidopsis eIFiso4G1 and eIF4G2 have molecular weights of 86 kDa and 83 kDa, respectively, showing approximately 27% identity and 41% similarity with eIF4G; however, eIFiso4G2 is expressed at lower levels than eIFiso4G1 (Gallie, 2016; Lellis et al., 2010). Interestingly, in TDNA knockout lines, single mutant plants expressing either *eIFiso4G1* or *eIFiso4G2* show very little phenotypic

difference as compared with wild type Arabidopsis plants, yet eIFiso4G1/eIFiso4G2 double mutants have a significant phenotype (Figure 1.3). The double mutant displays slow and stunted growth, reduced chlorophyll and lower fertility (Figure1.3) (Lellis et al., 2010). This severe phenotypic effect of the double mutant as opposed to the single knock-outs also suggests at least some functional redundancy between the two isoforms; however, in a study to isolate genes involved in fatty acid synthesis in Arabidopsis it was only the mutant lacking *eIFiso4G1* that was seen to play a prominent role in seed oil biosynthesis (Li et al., 2017). This would suggest there may be both overlapping and independent functions for the two genes in Arabidopsis.



From (Lellis et al., 2010)

Figure 1.3 **Image exhibiting the phenotypes of wild type Arabidopsis (Columbia; Col-0), and the *eIFiso4G1* (i4g1) and *eIFiso4G2* (i4g2) single and double T-DNA knockouts.** Left: overhead comparison of 21 day old plants. Right: profile view of 56 day old plants.

Also, unique to the plant eIF4G isoform, is a conserved N-terminal sequence not found in eIF4G (Figure 1.2). This sequence is XSLRPGG with X being a hydrophobic amino acid I, L, or V and occurs before the first 20 amino acids of the N-terminus. While

this sequence is conserved in all green algae and plant eIFiso4Gs, complete removal of these sequences cause no noticeable effects in the binding of eIF4E, eIF4A and can carry out protein synthesis at wild type levels (Metz & Browning, 1996). The function of this conserved sequence is unknown (Patrick & Browning, 2012).

While most flowering plants have more than one eIFiso4G gene, usually one of the copies is found to carefully preserve the consensus eIF4E binding sequence while other copies are found to diverge from the sequence (Patrick & Browning, 2012). Both Arabidopsis eIFiso4G isoforms carry out comparable *in vitro* translation activity and the reason for such differences is unknown (Mayberry et al., 2009).

An H1-CT (C-terminal from the HEAT-1/MIF4G domain) site seen in animals and fungi has also been identified in both eIF4G and eIFiso4G (Figure 1.2) (Patrick & Browning, 2012). While the conserved sequences in both isoforms are similar to those found in animal and fungi, there are differences between the conserved sequences in eIF4G and eIFiso4G that set them apart from each other. The pronounced phenotype of the double eIFiso4G mutant compared to an apparent lack of phenotype for plants lacking eIF4G suggests that in plants, eIFiso4G performing some function that eIF4G alone cannot provide (Figure 1.4). Overexpression of eIF4G in plants lacking both eIFiso4G1 and eIFiso4G2 does not appear to rescue the severe phenotype (Lellis and Browning, unpublished data) suggesting that there is some specific translation function attributable to the eIFiso4F complex that eIF4F cannot carryout.

eIF4E x eIF4G
(Only *eIFiso4F1/2*; no *eIF4F*)
***eIF4F* mutant**

eIFiso4E x eIFiso4G1 x eIFiso4G2
(only *eIF4F*)
***eIFiso4F* mutant**

Figure 1.4 **Arabidopsis plants exhibiting the phenotypes of plants mutant in *eIF4F*, *eIFiso4F* and various combinations of these complexes.** Starting at the top, *eIF4E x eIF4G,* and *eIFiso4E x eIFiso4G1 x eIFiso4G2* double and triple T-DNA knockouts. Overhead view of 50 day old plants grown under 18 hour light/6 hour dark cycle.

1.2.2 eIF4E

eIF4E is the cap-binding protein in the eIF4F cap-binding complex that binds to the m^7G cap of the mRNA (Browning et al., 1987). There are three genes for Arabidopsis eIF4E: *eIF4E, eIF4E1B and eIF4E1C* (Patrick & Browning, 2012; Patrick et al., 2014; Rhee et al., 2003). eIF4E is a class I family member of the eIF4E-family of cap-binding proteins, classified as such due to a conserved pair of tryptophans that are positioned equivalently to the Trp-43 and Trp-56 of human eIF4E (Joshi et al., 2005; Patrick et al., 2014). eIF4E1B and eIF4E1C are the result of a Brassicaceae-specific gene duplication event and are located near each other on chromosome 1 (Joshi et al., 2005; Patrick et al., 2014). While eIF4E1B and eIF4E1C can both bind to the m^7G of mRNA, they both have low affinities for eIF4G and have lower activity in *in vitro* translation assays than eIF4E. Additionally, plants lacking both *eIF4E* and *eIFiso4E*, another class I eIF4E-family member in Arabidopsis, is lethal, providing evidence that eIF4E1B and eIF4E1C alone cannot support *in vivo* protein synthesis (Callot & Gallois, 2014; Patrick et al., 2014). This

is most likely due to the fact that eIF4E1C has negligible levels of expression in Arabidopsis and eIF4E1B is only enriched in reproductive tissues, having very low expression levels in all other tissues, suggesting that eIF4E1B may have some unknown specialized roles in those tissues (Patrick et al., 2014).

eIFiso4E is the counterpart of the Arabidopsis eIF4E cap-binding protein and binds to eIFiso4G in the eIFiso4F cap-binding complex (Figure 1.1) (Browning & Bailey-Serres, 2015; Browning et al., 1992). eIFiso4E is found only in higher plants and not other eukaryotes and appears to have evolved around the same time as flowering (Patrick & Browning, 2012). It shares approximately 41% amino acid identity with eIF4E, both having a molecular mass of about 24kD (Browning & Bailey-Serres, 2015; Lellis et al., 2010). Although both eIF4E and eIFiso4E share the same structure and function, and display no obvious phenotype in Arabidopsis T-DNA insertion plants, differences between the two isoforms in plants have been revealed (Monzingo et al., 2007; Patrick et al., 2014). In wheat (*Triticum aestivum*), when eIF4E and eIFiso4E are complexed with their respective binding partners, eIF4G and eIFiso4G, eIFiso4F preferred dimethylated cap analogues 1.3 – 1.9-fold more than a single methyl group at the N-7 position, whereas eIF4F had a 1.2 – 1.7-fold lower affinity for the trimethylated cap relative to the single methylated cap (Figure 1.1) (Carberry et al., 1991). Also, while eIF4F preferentially supported translation from a structured 5' leader and from uncapped monocistronic and dicistronic mRNA's over eIFiso4F, eIFiso4F preferentially translated an unstructured mRNA (Gallie & Browning, 2001). Additionally, they have been found to be differentially expressed in Arabidopsis (Patrick et al., 2014; Rodriguez et al., 1998). *In situ* hybridization carried out on 2-week-old Arabidopsis seedlings revealed that eIF4E is present in all tissues except in root tissues, whereas eIFiso4E is mostly found in flowers and young developing tissues (Rodriguez et al., 1998). Interestingly, while *eIF4E* is conserved throughout plants,

eIFiso4E only started appearing at the time of flowering plants (Patrick & Browning, 2012). Finally, resistance to some viruses has been associated with the lack of either *eIF4E* or *eIFiso4E,* as well as *eIF4G* or *eIFiso4G* (Browning & Bailey-Serres, 2015; Jiang & Laliberte, 2011).

1.3 Plant Viral Initiation of Translation

Many plant viruses exploit the two isoforms of eIF4F-family of translation initiation factors and are genes involved in natural recessive resistance to many viruses (Truniger & Aranda, 2009). This is seen most notably in *Potyviridae* (potyviruses), a single stranded positive-sense RNA virus, the most abundant and effective type of eukaryotic virus, making up the largest group of plant viruses with RNA genomes (Miras et al., 2017; Simon & Miller, 2013; Wang & Krishnaswamy, 2012). Potyviruses require the use of a covalently bound VPg (viral protein genome-linked) at the 5' end of the viral RNA for infectivity (Truniger & Aranda, 2009). The VPg either directly or indirectly interacts with the cap-binding protein(s) to presumably allow for protein synthesis to take place, in place of a m^7G cap (Jiang & Laliberte, 2011). While some potyviruses can use either eIF4E or eIFiso4E, a large number of them are dependent on only one of the eIF4E isoforms (Truniger & Aranda, 2009). For example, a mutant form of the gene encoding eIFiso4E in Arabidopsis has been linked to resistance to *Tobacco etch potyvirus* (TEV), *Turnip mosaic potyvirus* (TuMV) and *Lettuce mosaic virus* (LMV), while loss of susceptibility to *Clover yellow vein virus* (CLYVV) has been mapped to a mutant form of eIF4E in Arabidopsis (Lellis et al., 2002; Truniger & Aranda, 2009). Furthermore, TEV requires eIF4E in pepper and tomato while LMV requires eIF4E in lettuce. Interestingly, all resistance-breaking potyvirus genes contain mutations in the same general area of the protein, suggesting that this could be the point of contact between the VPg and the cap-binding protein (Truniger

& Aranda, 2009). This has been confirmed by yeast two-hybrid assays, enzyme-linked immunosorbent assays (ELISAs) and GST (Glutathione S-transferase) pull-downs used in the study of various potyviruses using recombinant proteins and mutant isolates.

Potyvirus resistance has also been linked to mutations in eIF4G and eIFiso4G, though not as often (Truniger & Aranda, 2009). For example, eIF4G has been linked to gene resistance in CLYVV and eIFiso4G has been linked to TuMV and LMV. The resistance seen with these viruses are usually associated with their need for their respective binding partners, eIF4E and eIFiso4E, indicating that the entire cap-binding complex is necessary for the initiation of protein synthesis in those viruses but not necessarily, as is seen in the case of *Rice yellow mottle virus*, in which virus resistance is controlled by eIFiso4G alone.

Outside of potyviruses, most other positive strand RNA plant viruses use cap-independent translation elements (3'-CITE) (Miras et al., 2017; Simon & Miller, 2013). These viruses make use of structures in the 3' untranslated region that are made up of various conformations such as pseudoknots or Y-, I-, or T-shaped structures that are able to functionally substitute for the m^7G cap. The 3' structure is used to circularize the RNA by binding to eIF4G or eIFiso4G and/or an element in the 5' UTR to use in recruitment of the ribosome at the 5' end of the viral RNA for scanning to begin for the initiation of translation. For many of these viruses in which both isoforms are tested, eIF4F translated the virus RNA more efficiently than eIFiso4F (Nicholson et al., 2010; Nicholson et al., 2013; Treder et al., 2008; Wang et al., 2009b).

Interesting to note, neither *eIF4E1B* and *eIF4E1C* have ever been found to be viral resistant genes to date, further suggesting they do not have a pivotal role in canonical translation initiation (Browning & Bailey-Serres, 2015).

1.4 Eukaryotic Regulation of Translation

The regulation of translation is conserved across most all eukaryotes and mostly takes place at the level of initiation (Browning & Bailey-Serres, 2015; Sesma et al., 2017; Sonenberg & Hinnebusch, 2009). There are a number of strategies that can be used in regulation such as phosphorylation of initiation factors, inhibition of initiation factors by binding other proteins, secondary structure of mRNA and/or mRNA binding proteins (Muench et al., 2012; Sonenberg & Hinnebusch, 2009). One prominent regulatory mechanism in mammals is through the phosphorylation of eIF2, the core member of the ternary complex (Figure 1.1) (Sonenberg & Hinnebusch, 2009). Formation of the ternary complex requires that eIF2 is bound to GTP which becomes hydrolyzed to GDP after the start codon recognition step during the initiation phase of translation. The guanine nucleotide exchange factor (GEF) required to carry out this exchange is eIF2B. Phosphorylation of Ser51 of the alpha subunit of eIF2 (eIF2α) prevents the release of eIF2 from eIF2B, therefore decreasing the amount of eIF2 available for ternary complex formation and therefore decreasing protein synthesis (Sonenberg & Hinnebusch, 2009). In mammals, this phosphorylation occurs by general control non-derepressible 2 (GCN2) kinase, heme-regulated inhibitor (HRI), PKR-like ER kinase (PERK) and protein kinase double-stranded RNA-dependent (PKR), with each of the kinases responding to a particular type of stress (Browning & Bailey-Serres, 2015; Donnelly et al., 2013; Hinnebusch, 2005). Yeast only use GCN2 in response to stress (Hinnebusch, 2005).

Other kinases known to be important for the regulation of translation in mammals are mTOR (mammalian target of rapamycin) and Mitogen-activated Protein Kinase (MAPK) Interacting Kinase 1 (MNK1) (Sonenberg & Hinnebusch, 2009). mTOR is a serine/threonine kinase that regulates cell proliferation and growth (Holz et al., 2005; Sonenberg & Hinnebusch, 2009). mTOR phosphorylates a subunit of eIF3 to allow its

joining to the PIC. mTOR also phosphorylates ribosomal protein S6 kinase I (S6K1), activating it, which in turn leads to the phosphorylation of eIF4B, allowing it to join with eIF4A in the initiation of translation. eIF4E-binding proteins are also phosphorylated by mTOR, discussed below, to upregulate protein synthesis. In mammals, MNK1 is known to associate with eIF4G to allow the phosphorylation of eIF4E, which in turn enhances the binding of eIF4E to mRNA (Pyronnet et al., 1999).

While plants do not have MNK, plants use mTOR to activate subunit h of eIF3 (Browning & Bailey-Serres, 2015). They also use brassinosteroid insensitive receptor kinase (BRI1) to activate eIF3i for the initiation of translation (Ehsan et al., 2005). Furthermore, plants also have a GCN2 kinase that phosphorylates the conserved serine in eIF2α as seen in mammals and can complement a yeast *gcn2* mutant (Browning & Bailey-Serres, 2015; Sesma et al., 2017; Zhang et al., 2003). Induction of GNC2 in Arabidopsis has been seen in response to amino acid starvation, purine starvation, several types of stress such as cadmium, wounding and cold shock. However, the use of eIF2B in the exchange of GDP for GTP in eIF2α appears to be unnecessary as this activity has yet to be found in plants, along with the finding that in wheat germ, affinity for GDP by eIF2 is only 10 times more than it is for GTP whereas in rabbit reticulocyte, the affinity for GDP over GTP is approximately 100-fold (Shaikhin et al., 1992). This observation suggests that a GEF (guanine exchange factor) may not be needed by plants and thus explains why eIF2B may not function similarly in plants as in highly studied mammalian system.

In *in vitro* assays, CK2 (formerly casein kinase II) has been seen to phosphorylate Arabidopsis eIF2α, eIF2ß, eIF3C, eIF4B and eIF5 (Dennis & Browning, 2009). In a study that analyzed the phosphoproteome to see the effects of light/dark on the translational apparatus in Arabidopsis, eIF3, eIF4A, eIF4B, eIF4G and eIF5 were found to be involved in the regulation of translation (Boex-Fontvieille et al., 2013). While both yeast and

mammals have CK2, neither appear to use CK2 in the phosphorylation of eIFs for the regulation of translation suggesting potential differences in how plants may regulate translation (Hinnebusch & Lorsch, 2012).

The second major mechanism used by mammals to regulate protein synthesis is through the use of 4E-binding proteins (4E-BPs) (Roux & Topisirovic, 2012; Sonenberg & Hinnebusch, 2009). 4E-BPs compete with eIF4G to bind with eIF4E to prevent the formation of the cap-binding complex. Phosphorylation of 4E-BPs by target of rapamycin (TOR), a nutrient sensor and growth signaling kinase, decreases their ability to effectively bind to eIF4E, allowing the formation of eIF4F and therefore upregulating the initiation of cap-dependent translation. While there is also TOR activity in plants, there is thus far no evidence of 4E-BPs in plants (Menand et al., 2002; Sesma et al., 2017).

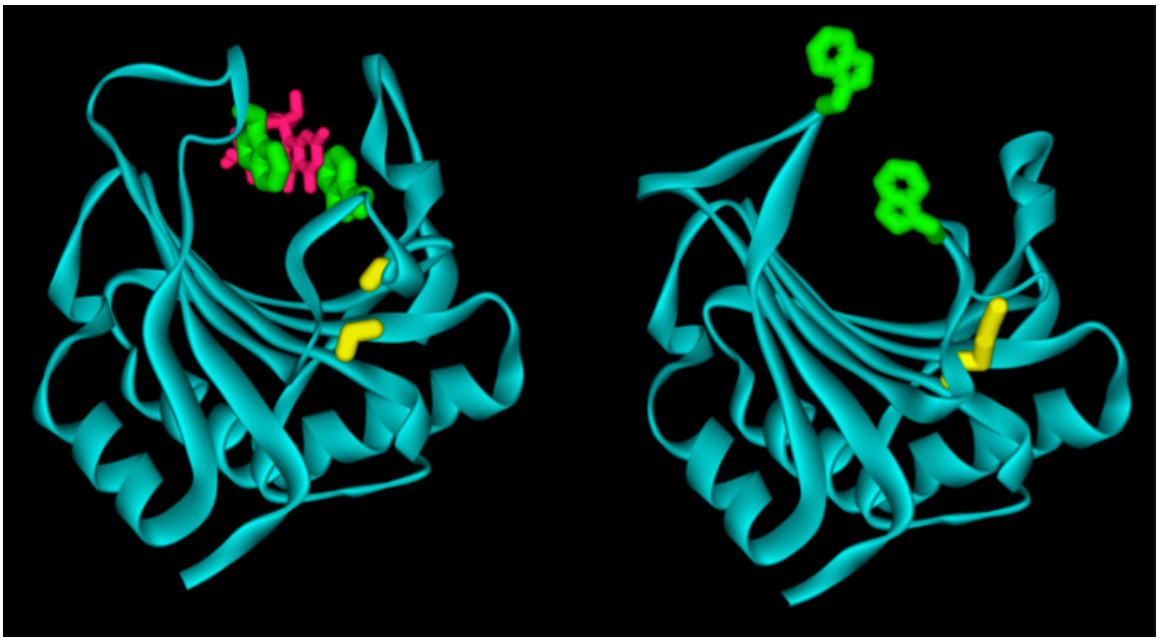
While eIF4E/eIFiso4E and eIF4G/eIFiso4G apparently have some redundancy due to the fact that a knockout of any single gene does not cause any noticeable phenotypic effects, the apparent lack of the typical mammalian or yeast methods of the regulation of translation and the fact that most of the regulation of protein synthesis occurs at the point of the initiation of translation, suggests that there are still some undiscovered mechanisms through which plants regulate translation (Patrick et al., 2014).

1.5 A Novel Disulfide Bond

Sequence alignments carried out on the eIF4E-family of cap-binding proteins show that plants have a conserved pair of cysteines that other eukaryotes do not have (Monzingo et al., 2007; Patrick & Browning, 2012). Using x-ray crystallography, it was observed that an intramolecular disulfide bridge was able to be formed with these two conserved cysteines (Cys) in wheat eIF4E (Figure 1.5) (Monzingo et al., 2007). Furthermore, NMR

spectroscopy was used to determine that the disulfides are reduced in solution, but can form a disulfide bond in intentionally oxidizing conditions (Monzingo et al., 2007).

Plants use redox in the regulation of many activities such as in the regulation of germination, seed development and photosynthesis (Buchanan & Balmer, 2005). This conserved pair of cysteines, along with the fact that there are 11 very conserved amino acids on the surface of eIF4E in plants that differ from mammals and yeast, raises the possibility that plant eIF4F group I members could possibly regulate the initiation of translation through redox changes in the protein structure (Monzingo et al., 2007).



From (Monzingo et al., 2007)

Figure 1.5 **Ribbon diagrams of wheat eIF4E.** Wild type (right) and C113S mutant (left). The Trp residues that form the binding site are in green in both diagrams with the mutant bound to a m^7GDP, which is colorized in magenta. The Cys/Ser residues, shown in yellow, are shown in the mutant in their reduced form, left, and in the disulfide bond, right.

Chemical probes may react differentially with proteins when they are transitioning to a different conformation (O'Brien et al., 2013a). A chemical probe designed to react with primary amines for use in UVPD mass spectrometry was used to investigate the solvent exposed lysines of wheat eIF4E and wheat eIFiso4E for structural changes upon the binding of m^7GTP and in different oxidation states (O'Brien et al., 2013a; O'Brien et al., 2013b; Vasicek et al., 2012). Both eIF4E and eIFiso4E have 11 lysines in which 7 are perfectly conserved and 4 that are within 1 or 2 residues. The study used eIF4E, eIF4E bound to full-length eIF4G, eIF4E bound to a truncated version of eIF4G and eIFiso4E to map surface accessibility under reducing and oxidizing conditions (O'Brien et al., 2013a). The findings indicated that both eIF4E and eIFiso4E exhibited changes in their structure in

response to changes in the oxidation state of the disulfide bonds. In the presence of excess (1 mM) dithiothreitiol (DTT) versus without excess DTT (0.1 mM), and 10 μM m^7GTP, lysines showing significant probe reactivity were located in areas that could be affected by disulfide bond formation. In the case of eIF4E, only one lysine (K203), located proximal to the binding pocket, demonstrated a significant change, approximately 7-fold, while eIFiso4E contained 3 lysines, 1 located within the binding pocket (K110), one proximal (K156) and one more distal (K186) to the binding pocket, having significant reactivity with the probe, approximately 2 – 6-fold. Additionally, the same lysine significantly affected in eIF4E in excess DTT was also the only lysine found to change significantly with the binding of the truncated form of eIF4G. Even more lysine reactivity was observed with the binding of the full length eIF4G. This indicates that the binding of eIF4G may induce a shift in the cap-binding site that may increase the affinity of eIF4E for m^7GTP which has been seen to modestly occur in eIF4E in *in vitro* reducing conditions (Monzingo et al., 2007). Overall though, the lysine amines of eIFiso4E were more reactive to the chemical probe than the lysine amines of eIF4E (O'Brien et al., 2013a). In a study done to investigate the binding preferences of eIF4G and eIFiso4G for either eIF4E or eIFiso4E, each displayed a strong preference for their binding partner, eIF4E and eIFiso4E, respectively (Mayberry et al., 2011). Using surface plasmon resonance (SPR), the mixed complex of eIF4G-eIFiso4E was observed to have a K_D value 80-fold higher than for the eIF4G-eIF4E complex. Likewise, the mixed complex of eIFiso4G-eIF4E had a K_D value 148-fold higher than eIFiso4G-eIFiso4E. When both complexes have the correct binding partners, the observed K_D values for eIF4F and eIFiso4F are subnanomolar (0.18 nM and 0.08 nM, respectively). While mixed complexes are able to carry out translation in *in vitro* translation assays, only the mixed complex of *eIFiso4G2* x *eIF4E* are able to survive in Arabidopsis T-DNA knockout plants lacking eIFiso4G1, eIF4G and eIFiso4E (Mayberry

et al., 2011; Patrick & Browning, 2012). While this lends itself to the idea that eIFiso4G most likely partnered itself with eIF4E before the evolution of eIFiso4E, it could possibly also indicate that the correct binding partner is important for the *in vivo* functionality of the respective cap-binding protein, due to the differences observed between eIF4E and eIFiso4E (Patrick & Browning, 2012). That is, that the respective scaffolding protein imparts some kind of advantage to its correct cap-binding partner that the other one does not provide. Furthermore, plants with the mixed complex of eIF4G and eIFiso4E only are not viable and this suggests specificity of the functions for the complexes (Patrick et al., 2014).

Could the structural changes that occur in the two cap-binding isoforms that are so pronounced in relation to changes in the oxidation state of the disulfide bonds be a unique way in which plants regulate translation?

1.6 Plants are Different

The differences found in the molecular machinery and its regulation in plant protein synthesis as opposed to those used in other eukaryotes such as yeast and mammals, most likely reflects the physiological differences found in this kingdom, such as photosynthesis and a plant's static existence in an ever-changing environment (Browning & Bailey-Serres, 2015). There is an onslaught of oxidation occurring due to the exposure of light to the plant (Barber & Andersson, 1992). In levels of high light the plant must be able to respond quickly and efficiently to high levels of oxidation. Other stresses such as high salt, cold and heat, also have an effect on the photosynthetic apparatus of the plant. The specialized biochemistry found only in plants brings about needs that are unique to plants and therefore specialized ways to deal with these differences and regulate processes such as protein synthesis (Browning & Bailey-Serres, 2015).

Seeing that plants do not appear to utilize either of the two major sources of translation regulation (4E-BPs or eIF2B inhibition by eIF2α phosphorylation) that are utilized by other eukaryotes, but do have a unique isoform for the cap-binding complex, it is possible to speculate that plants regulate translation differently than other eukaryotes. Could those very conserved cysteines in the cap-binding proteins be important players in the regulation of translation? Or could there other uncharacterized phosphorylation events that aid in the regulation of translation (Humphreys et al., 1988)? And finally, could the two different cap-binding isoforms preferentially bind specific mRNAs as another method of regulation? The following studies seek to answer these questions.

CHAPTER 2: IS THERE AN UNCHARACTERIZED KINASE INVOLVED IN THE REGULATION OF PLANT TRANSLATION?

2.1 INTRODUCTION

The phosphorylation-dephosphorylation of proteins is a widely-used mode for the regulation of cellular processes in the eukaryotic cell. This is true for some of the initiation factors involved in the very early steps of protein synthesis, such as 4E-BPs or eIF2α as was previously discussed for mammalian systems. Previously, an unidentified wheat germ kinase was found to phosphorylate eIF4G, eIFiso4G, eIF4B and a 60S ribosomal protein, but not eIF4A, eIF3, eIF2, elongation factors or 40S ribosomal proteins (Humphreys et al., 1988). Using the native wheat eIF4F, it was determined that this wheat germ kinase had no effect on the binding of mRNA to the 40S ribosomal subunit. Although there were no observable effects on binding of mRNA to the 40S ribosome *in vitro*, there could be other uncharacterized regulatory effects.

The identification of the type of kinase was not possible in 1988 and efforts in 2006 did not yield any actionable data to identify the type of kinase. Recently, we had this kinase analyzed again by more sensitive mass spectrometry. The results indicated that the wheat kinase fraction contained peptides (over 300 proteins were identified) with sequence similarity to at least two different Arabidopsis calcium dependent protein kinases (CDPKs), CDPK2 and CDPK4, and to an Arabidopsis SNF1 kinase regulatory subunit, a subunit of a highly conserved heterotrimeric kinase related to the SNF1/AMPK/SnRK1 family of protein kinases involved in cellular energy homeostasis (Bouly et al., 1999; Emanuelle et al., 2015). Other kinases were also identified, but either had fewer peptides or encoded a kinase that did not fall within the size range of the wheat germ kinase, ~55 kDa. Analysis of the native wheat kinase in the presence of EGTA suggested that the kinase was indeed calcium-dependent (see Fig. 2.2). Most Arabidopsis CDPKs are ~54 –

66 kDa, and combined with the apparent calcium dependence makes the wheat kinase very likely to be a member of the CDPK family of kinases in plants (Cheng et al., 2002).

Calcium signaling is also a widely-used second-messenger that can be used to in eukaryotic cells to activate signaling cascades (Cheng et al., 2002). In plants, it can be used to generate responses to abiotic stress, pathogen response, hormones, light and mechanical disturbances. Plants contain four families of kinases that are calcium responsive: Calcium- and calmodulin-dependent protein kinases (CCaMKs), sucrose non-fermenting-related kinases 3 (SnRK3s), Calcium-dependent protein kinases (CDPKs) and CDPK-related protein kinases (CRKs) (Simeunovic et al., 2016). CDPKs are the largest family of these kinases and are only found in plants and in some protozoans. CDPKs are a unique kinase in that they are calmodulin independent in their ability to phosphorylate targets. They have 3 domains, an N-terminal variable domain, a kinase domain and a CDPK activation domain (CAD) and are thought to have been a fusion protein of the N-terminal portion of an ancestral Calcium/calmodulin dependent Ser/Thr protein kinase (CaMK) and the C-terminal portion of an ancestral calmodulin gene, containing the EF-hand calcium binding domain that is seen in calmodulin (de Souza et al., 1998; Schulz et al., 2013; Zhang & Choi, 2001).

There are 34 Arabidopsis CDPKS, and they can be grouped into 4 main groups based on their sequence homology (Cheng et al., 2002). In the plant cell, the specificity of calcium responsive phosphorylation by CDPKs is achieved by a combination of several factors such as the particular tissue and the developmental stage they are expressed in, their localization, that is, how broadly or narrowly a particular CDPK is known to be distributed, and the different calcium binding requirements needed for each of the CDPKs (Boudsocq et al., 2012; Simeunovic et al., 2016). The N-terminal variable domain determines their subcellular localization (Simeunovic et al., 2016). They are known to be mostly membrane

localized exclusively or soluble and membrane associated, with a few being only soluble. CDPKs have been linked to or are believed to be linked to several functions such as the regulation of pollen tube growth, hormone signaling, the transition to flowering, senescence, and biotic and abiotic stress signaling (Levy & Dean, 1998; Simeunovic et al., 2016).

CDPKs have been shown to play a major role in the innate immunity of plants (Simeunovic et al., 2016). In the early steps of pathogen-associated molecular pattern (PAMP)-triggered immunity (PTI), distinct patterns of calcium are released that lead to the generation of ROS (reactive oxygen species) by the NADPH oxidizing respiratory burst homologue D (RBohD), which requires both the high levels of calcium in the cell and its phosphorylation for activation. In potatoes (*Solanum tuberosum*), StCDPK4 and 5 are known to phosphorylate StRBohD at a pair of conserved serines which are also present in the Arabidopsis RBohD. Four Arabidopsis CDPKs have also been identified as participants in a regulatory role in the early PTI transcriptional response (Boudsocq et al., 2010; Simeunovic et al., 2016). In effector-triggered immunity (ETI), several CDPKs have been shown to be activated and to work in conjunction with other kinases, RBohD, RBohF and transcription factors (TFs) (Simeunovic et al., 2016). Taken together, the data suggests that there is crosstalk occurring at the points of transcription factors and RBohs to bring about the proper immunity response. This crosstalk with other kinases is also seen in the regulation of additional metabolic enzymes, TFs and 14-3-3 proteins in the regulation of primary metabolism during both normal growth and starvation conditions.

It is expected that CDPKs are multifunctional and have numerous phosphorylation targets (Curran et al., 2011). Could this also include the regulation of translation? More specifically, could this unidentified, calcium dependent, wheat germ kinase that phosphorylates eIF4G, eIFiso4G and eIF4B function in the regulation of translation? The

specificity and role of the phosphorylation of the apparent single site of eIFiso4G by CDPKs is the topic of this chapter.

2.2 MATERIALS AND METHODS

2.2.1 Kinase Assay

The kinase assay was adapted from Humphreys et al. (1988). Briefly, in 10 μL final volume, each reaction consisted of 45 mM Hepes, pH 7.6, 9 mM $Mg(OAc)_2$, 3.8 mM DTT, 0.38 mM ATP ([γ-^{32}P] ATP (60 Ci/mM)), ~0.25 mg/mL substrate and 600 ng/mL of kinase. Alternatively, unlabeled ATP can be used without the addition of [γ-^{32}P] ATP. To test for calcium dependence, 2 mM (final concentration) EGTA was added. The reaction is incubated at 27˚C for 1 hour. 10 μL of 2X SDS PAGE loading buffer containing 180 mM DTT is added to each reaction and 10 μL of the reaction is separated on a 4 – 20% SDS PAGE regular TGX Precast Gel or TGX Stain-Free Precast Gel (Bio-Rad). In the case of using [γ-^{32}P] ATP, the gel is first imaged for protein loading on a Bio-Rad Chemidoc Imager to activate the stain-free ligand and then placed in a phosphor screen overnight and radioactivity imaged the following day on a Typhoon 9500 Imager. For assays in which unlabeled ATP was used instead of [γ-^{32}P] ATP, the gel was stained using Pro-Q Diamond Phosphoprotein Gel Stain (Invitrogen), according to the manufacturer's instructions.

2.2.2 Identification, Cloning and Purification of the Wheat eIF4G Kinase

Mass spectrometry was carried out on the native wheat kinase. Briefly, ~50 μg of protein was run about 1 cm into a TGX 4-20% Precast Gel (Biorad) and stained with colloidal commassie blue. The band of protein was excised and analyzed at the ICMB mass spectrometry core facility. 475 proteins were identified in the "purified" native kinase preparation. To facilitate identification of the proteins with kinase signatures, Claire

McWhite graciously assisted in sorting out all the proteins that could be kinases. There were ~20 protein kinases that sorted out and the two families with highest number of peptides were the Arabidopsis calcium-dependent kinase family and snf1-related kinase family. Both of these kinase families average size fell within the known size (~55 kDa) for the wheat kinase. An easy method for determining calcium dependence was to check for sensitivity to EGTA. The kinase did show sensitivity to EGTA and so the calcium-dependent kinase was selected for cloning. Using the wheat database and the peptide sequences and again with the assistance of Claire McWhite using the wheat genomic database a wheat calcium dependent kinase (Triticum aestivum ABC59621.1) was selected for cloning (see Appendix I for protein sequence).

The cloning and purification of the recombinant eIF4G wheat kinase was carried out similarly as in Boudsocq et al. (2012). A codon optimized full-length cDNA of a wheat sequence that most closely matched a wheat calcium dependent kinase was used in the construction of the expression clone for the eIF4G kinase. The cDNA was cloned into pGEX-4T-1 expression vector (GE Healthcare) using BamH1 and Not1 restriction enzymes (Genscript). This vector contained a GST fusion at the N-terminus for use in the purification of the expressed protein, as described in Boudsocq et al. (2012).

4G kinase pGEX-4T-1 vector was transformed into chemically competent *E. Coli* BL21 cells. One colony was selected from an LB agar plate containing the antibiotic ampicillin, 100 μg/ml, to inoculate 5 ml of LB medium containing ampicillin and incubated at 37°C with shaking for approximately 4 hours. One ml was used to inoculate 50 ml of LB medium containing 100 μg/ml ampicillin and was incubated at 37°C with shaking overnight. The 50 ml culture was then used to inoculate 1 liter of LB medium containing 100 μg/ml ampicillin and was incubated at 37°C with shaking until the cells had grown to an OD_{600} of ~0.5. Isopropyl β-D-1-thiogalactopyranoside (IPTG) at a final concentration

of 1 mM was added and the cells were grown for 3 more hours in the same conditions. The cells were collected by centrifugation at 8,000 rpm for 10 min at 4°C in a JA10 rotor. The supernatant was discarded and the cell pellets were collected in a 50 ml conical tube and frozen at -80 °C until the purification was carried out. One mL pre-induction and post-induction aliquots were taken, collected at 18,000 x g for 3 minutes, resuspended in sample buffer and stored at -80 °C.

The cell pellets were thawed on ice and resuspended in 10 ml of 20 mM Hepes-KOH, pH 7.6, 100 mM KCl, 1 small protease inhibitor caplet (Thermo Scientific) and phenylmethane sulfonyl fluoride (PMSF) (1 mM). The cells were lysed using a Vibra Cell, Model VC40 sonicator, on ice using a 2 mm probe for 3 x 30 seconds at 70%, 2 x 30 seconds at 90% and 1 x 12 sec at 90%. 10% Triton X-100 was added to a final concentration of 1% before the centrifugation. The lysed cell pellets were centrifuged at 48,000 x g using an ultracentrifuge for 45 min. using a TI 70.1 rotor (Beckman). The supernatant was then used as described in Boudsocq et al. (2012). Briefly, a 0.5 mL glutathione-agarose (Glutathione Sepharose 4B, GE Healthcare Life Sciences) column was overlayed with a thin Sephdex G25 layeron top of the glutathione-agarose column with a fret between the two layers. The Sephadex G25 prevented clogging of the GST resin. The column was washed with 20 mM Hepes-KOH, pH 7.6, containing 100 mM KCl. The supernatant was added to the column and the column was then washed with 20 mM Hepes-KOH, pH 7.6, 100 mM KCl. The column was then incubated in 10 mM reduced glutathione, 50 mM Tris-HCl, pH 7.6 and the column flow stopped for 15 minutes to maximize interaction of the glutathione and the slow resumed to collect the protein. Fractions of 250 μL were collected and 10 μL aliquots were analyzed by SDS PAGE.

2.2.4 Purification of Wheat eIFiso4G Truncations

The truncations of wheat eIFiso4G cDNA have been previously described (See Figure 2.1) (Metz & Browning, 1996). pET3d vectors carrying the truncation sequences were transformed into chemically competent *E. Coli* BL21 cells. For each protein, one colony was selected from an LB/ampicillin agar plate to inoculate 5 mL of LB containing 100 μg/mL of ampicillin. The culture was then incubated at 37° C with shaking for four hours. 1 mL of the culture was added to 50 mL of LB-amp (100 μg/mL) and grown overnight at 37° C with shaking. The 50 mL culture was then added to 250 mL of LB medium containing ampicillin (100 μg/mL) and incubated at 37° C with shaking until the cells were had grown to an OD_{600} of 0.500 – 0.900. 1 mM IPTG (final concentration) was added to the culture and incubated for three more hours at 37° C with shaking. The cultures were then cooled quickly on ice to stop growth and centrifuged at 8,000 rpm at 4° C for 10 minutes in a JA10 rotor (Beckman). The pellet was frozen at -80° C.

The pellet was slowly thawed on ice and resuspended in 7 mL of N^1 buffer (20 mM Hepes-KOH, pH 7.6, 0.1 mM EDTA, 1 mM DTT, 10% Glycerol) containing 0.6 M KCl, one dissolved protease inhibitor tablet (Thermo Scientific) and 1 mM PMSF. The cells were sonicated using a Vibra Cell, Model VC40 Sonicator, on ice using a 2 mm probe for 3 x 30 sec. at 70%, 2 x 30 sec. at 90% and 1 x 15 sec. at 90%, with 1 minute pauses between each burst. The sonicated cells were centrifuged at 48,000 x g for 45 min. at 4° C using a TI 70.1 rotor (Beckman). The supernatant was diluted to 0.1 M KCl using the N^1 buffer with no added salt.

All procedures were carried out at 4° C. The supernatant was added to a 1.5 mL phosphocellulose (Whatman) column, with a thin layer of G25-Sephadex (Sigma-Aldrich) added to the top of the column, that had been equilibrated with N^1 buffer containing 0.1 M KCl followed by washing with the same buffer. N^1 buffer containing 0.3 M KCl was then

added to the column and the flow stopped for 15 min. prior to elution in 250 μL fractions. Aliquots of 10 μL were analyzed by SDS PAGE.

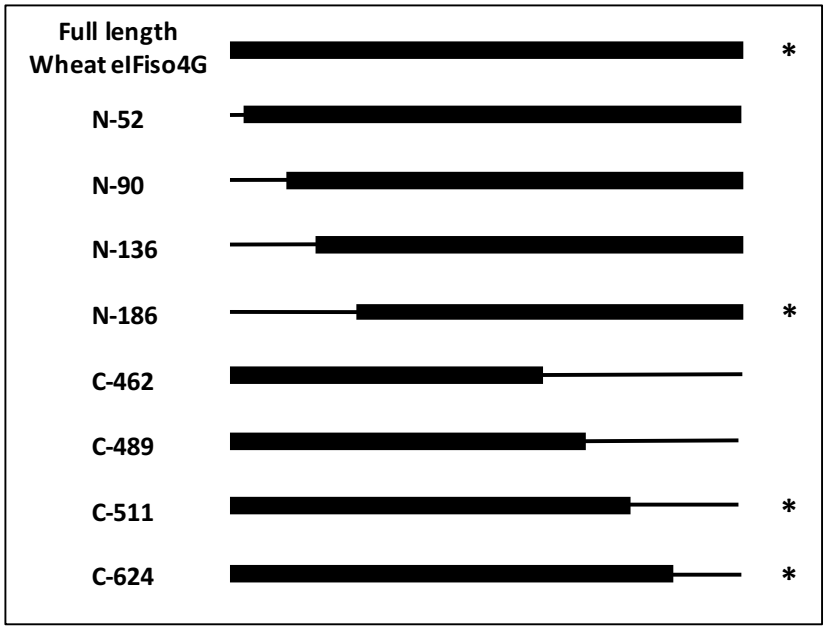


Figure 2.1 **Schematic of the wheat eIFiso4G truncations.** Full length eIFiso4G is represented by the black rectangular bar. The portions of amino acids missing in the subsequent truncations are indicated by thin black lines. Truncations used in following experiments are indicated by asterisks, *. Full length wheat eIFiso4G is 788 amino acids (aa); N-52, 736 aa; N-90, 698 aa; N-136, 652 aa; N-186, 602 aa; C-462, 462 aa; C-489, 489 aa; C-511, 511 aa; C-624, 624 aa. Figure adapted from (Metz & Browning, 1996).

2.2.3 Expression and Purification of Arabidopsis CDPKs

Clones for Arabidopsis CDPKs 2, 3, 5, 9, 11, 13, 25, 30, and 32 were generously provided by Dr. Marie Boudsocq and are described in Boudsocq et al. (2012). Briefly, the sequences for these 9 different CDPKs were each cloned into pGEX-2T vectors, which contain a GST fusion at the N-terminus. Each of the proteins were expressed and purified as was described in Boudsocq et al. (2012) and above in section 2.2.2, *Cloning and Purification of the Wheat Germ eIF4G Recombinant Kinase*.

2.3 RESULTS

2.3.1 Kinase Assay for Calcium Dependency

Wheat eIF4G, eIFiso4G and eIF4B were used as substrates for native wheat kinase. 2 mM EGTA was added to chelate calcium and determine if the kinase was calcium dependent. A strong dependency on calcium was seen for the phosphorylation events suggesting this kinase is a calcium dependent kinase as indicated by the mass spectrometry peptide analysis (Figure 2.2). No signal was seen in the lane for the reaction that did not include a substrate, indicating that no detectible auto-phosphorylation had taken place by the wheat kinase.

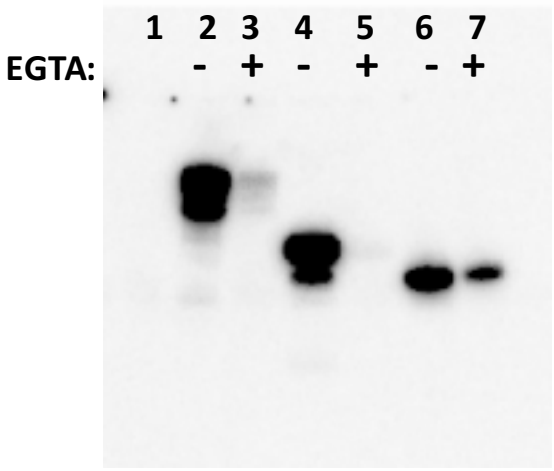


Figure 2.2 **Phosphor image of the native wheat kinase assay for calcium dependency.** EGTA at a final concentration of 2 mM was added to specified reactions. Radiolabeled [γ-^{32}P] ATP was used. (1) no substrate, (2) Wheat eIF4G, (3) Wheat eIF4G + EGTA, (4) Wheat eIFiso4G, (5) Wheat eIFiso4G + EGTA, (6) Wheat eIF4B, (7) Wheat eIF4B + EGTA. ~0.25 mg/mL substrate used in the lanes including substrate.

2.3.3 Cloning and Purification of the Wheat eIF4G Recombinant Kinase

A 55 kD wheat germ kinase that had previously been isolated and found to phosphorylate eIF4G, eIFiso4G and eIF4B, has now been determined to be calcium dependent (Figure 2.2) based on peptide similarity to Arabidopsis CDPK 2 and 4. The native wheat kinase has a molecular weight in range with the majority of Arabidopsis CDPKs, 54 – 66 kDa (Cheng et al., 2002). Taken together, this indicates that this kinase is most likely related to the CDPK family of plant kinases. Using a wheat sequence (ABC59621.1) that most closely matched the Arabidopsis CDPK2, the sequence was codon-optimized (Genscript) and used for cloning into a pGEX-4T-1 plasmid (GE Healthcare) which contained a Glutathione S-transferase (GST) tag at the N-terminus portion of the gene (Figure 2.3). The sequence was cloned into the vector using the BamH1

and Not1 restriction enzymes (Genscript). The use of the GST-tag has been successfully used for other CDPKs (Boudsocq et al., 2012).

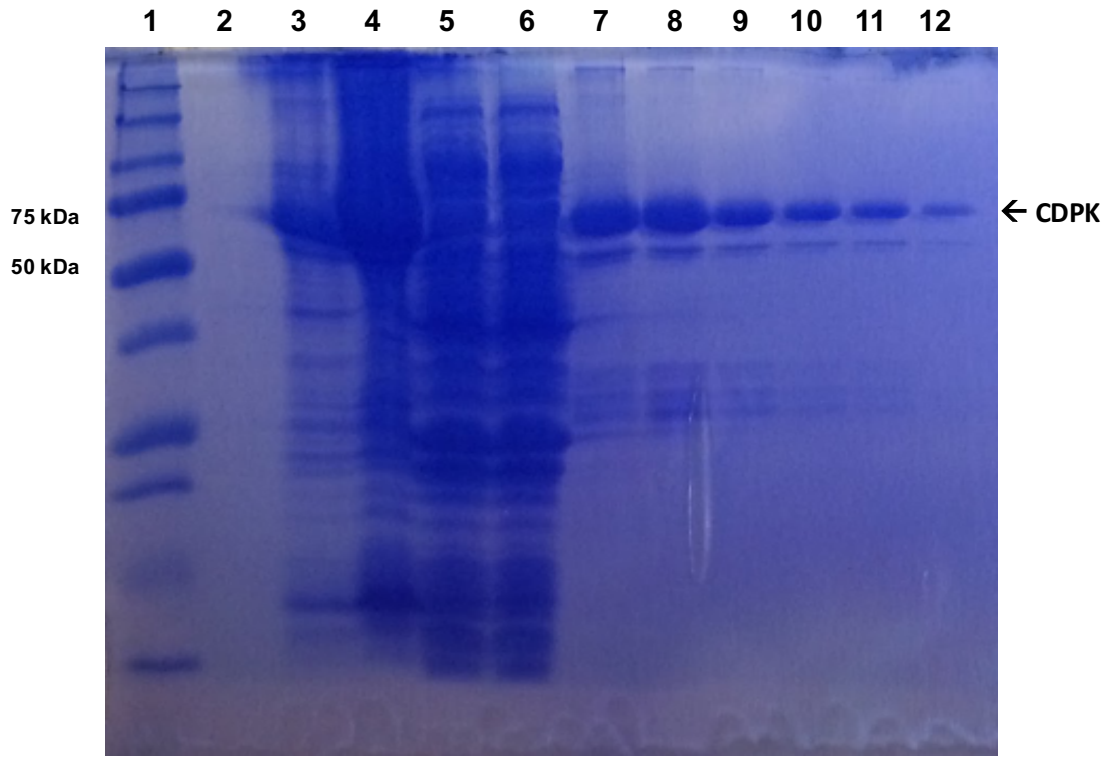


Figure 2.3 **Purification of the recombinant wheat germ kinase.** Coomassie Blue stained SDS Polyacrylamide (PAGE) gel of the GST-tagged protein. (1) Markers (Precision Plus Protein Standards, Bio-Rad), (2) Pre-induction, (3) Post-induction, (4) Pellet, (5) Input, (6) Flow-through, (7 – 12) Fractions 1 – 7, eluted with reduced glutathione.

2.3.4 Verification of Recombinant Wheat eIF4G Kinase Activity

The activity and calcium sensitivity of the recombinant kinase was determined using the native wheat kinase with the substrates eIF4G, eIFiso4G and eIF4B in the presence and absence of EGTA. The recombinant kinase exhibited activity with all three substrates similar to that of the native wheat kinase (see Figure 2.4). These results indicate that the correct type of kinase was selected and the recombinant kinase recapitulated the native kinase activity.

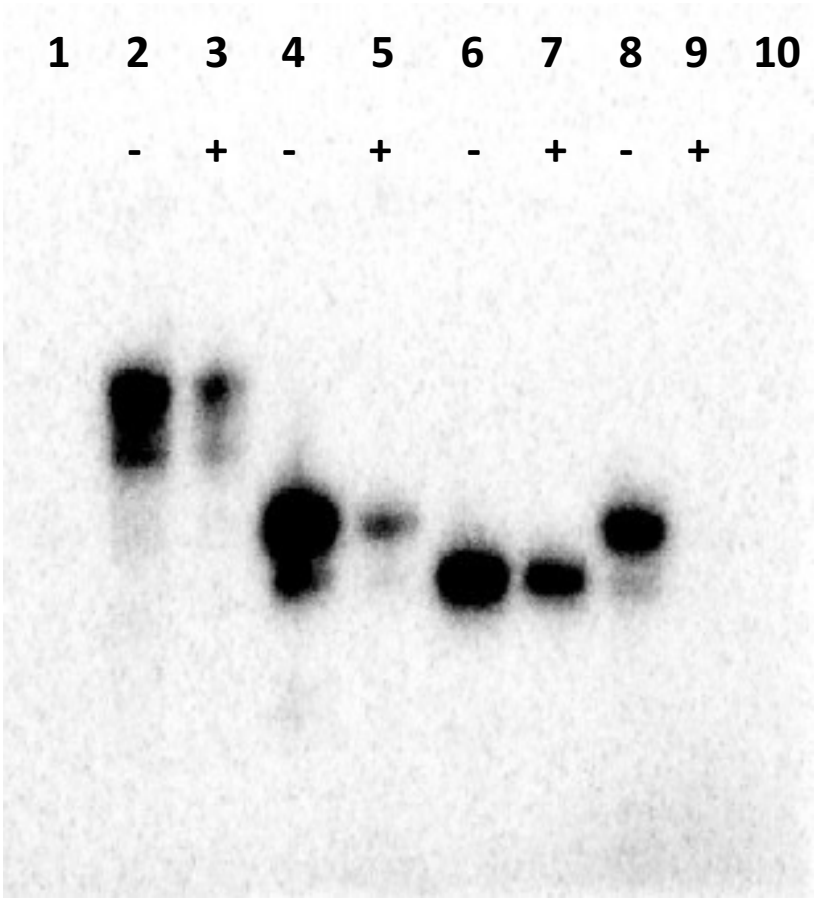


Figure 2.4 **Phosphor image verifying the activity of the recombinant wheat kinase.** The lanes are as follows, (1) Recombinant kinase without substrate, (2) Recombinant kinase with wheat eIF4G, (3) Recombinant kinase with wheat eIF4G + EGTA, (4) Recombinant kinase with wheat eIFiso4G, (5) Recombinant kinase with wheat eIFiso4G + EGTA, (6) Recombinant kinase with wheat eIF4B, (7) Recombinant kinase with wheat eIF4B + EGTA, (8) Native kinase with wheat eIFiso4G, (9) Native kinase with wheat eIFiso4G + EGTA, (10) Native wheat kinase without substrate. ~0.25 mg/mL substrate used in the lanes including substrate. Kinase final concentration, ~600 ng/mL per lane.

2.3.5 Mass Spectrometry on the Wheat Kinase to Determine Phosphorylation Sites

Full-length wheat eIFiso4G was phosphorylated by the recombinant kinase with unlabeled ATP and analyzed by mass spectrometry. Several sites were identified but the sites in the C-terminus and contained a conserved CDPK phosphorylation site (Figure 2.5) (Curran et al., 2011). The consensus site is (basic-X-X-S/T) and identified phosphoserines at 522 and 524 and 543 all fall within the consensus and all are located in the C-terminal portion of the protein.

Sequence | Modification List

Position	Target	Modification	Classification	Highest PTM Score	Highest Peptide Confidence	Sequence Motif
34	S	Phospho	Post-translational	49.5	High	AFAVAAsGSGDFL
36	S	Phospho	Post-translational	49.5	High	AVAASGsGDFLRP
48	S	Phospho	Post-translational	0.2	High	PHGGGAsGVSRIG
51	S	Phospho	Post-translational	0.8	High	GGASGVsRIGDLH
222	T	Phospho	Post-translational	71.1	High	GILNKLtPEKFDL
236	S	Phospho	Post-translational	86.7	High	KGQLLDsGITTAD
239	T	Phospho	Post-translational	71.1	High	LLDSGItTADILK
240	T	Phospho	Post-translational	71.1	High	LDSGITtADILKD
522	S	Phospho	Post-translational	1.8	High	NWEVQRsRSMPRG
524	S	Phospho	Post-translational	98.2	High	EVQRSRsMPRGDP
543	S	Phospho	Post-translational	100	High	LINKVPsINKPSP
548	S	Phospho	Post-translational	0	High	PSINKPsPINPRL
559	T	Phospho	Post-translational	100	High	RLLPQGtGALIGK

Figure 2.5 **Mass spectrometry results of the eIF4G wheat kinase on full-length wheat eIFiso4G.**

2.3.6 Localization of Phosphorylation Site Using Wheat eIFiso4G Truncations

To confirm the location of the consensus phosphorylation sites in the C-terminus, a series of truncations previously made for wheat eIFiso4G were utilized (Figure 2.1) (Metz & Browning, 1996). Expression and purification of the truncations were done as previously described (Metz & Browning, 1996). Specifically, N-186, C-511, C-624 were prepared for use along with the full-length protein which would confirm sites in the C-terminus between amino acids 511 and 624. The purified truncated wheat eIFiso4G proteins are shown in Figure 2.6.

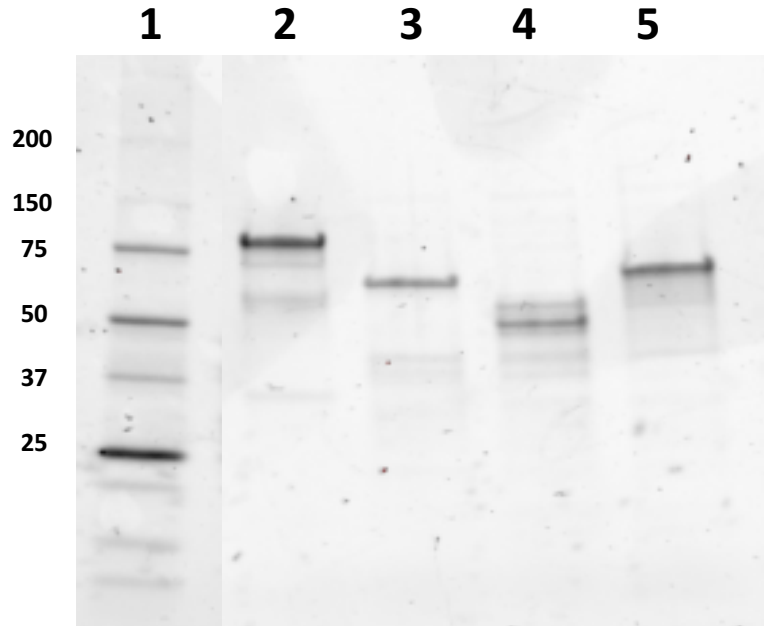


Figure 2.6 **SDS PAGE of the purified wheat eIFiso4G truncations to confirm phosphorylation of the CDPK consensus sequence sites.** (1) Markers (Precision Plus Protein Standards, Bio-Rad); (2) Full length eIFiso4G; (3) N-186; (4) C-511; (5) C-624. Stain-Free Polyacrylamide Gel (4% - 20%) (Bio-Rad). Each lane contains ~0.25 mg/mL protein.

Because the conserved CDPK consensus phosphorylation sites, (basic-X-X-S/T), that had been found to be phosphorylated in the mass spectrometry results were found in the region between amino acids 511 and 624, the wheat eIFiso4G truncations C-511 and C-624 were used as substrates along with the full-length wheat eIFiso4G and N-186 truncation (to eliminate sites identified in the mass spectrometry)(Figure 2.5) (Curran et al., 2011). The reactions that included the proteins possessing the consensus sequence sites, the full-length protein, N-186 and C-624, were all seen to be phosphorylated (Figure 2.7, A.). C-511, which would not contain those sites but still maintains a significant portion of the 788 aa of the full-length protein that does contain a number of other potential phosphorylation sites, did not have any phosphorylation, eliminating the other sites

identified as false positives. The Stain-Free Polyacrylamide Gel (4% - 20%) (Bio-Rad), (Figure 2.7, B.) shows that the substrates were present and loaded evenly.

These data confirm the location of the CDPK site to be in the C-terminus between AA 511 – 624, most likely at AA 522, 524 and/or 543 that fall within the consensus site for CDPKs. Further analysis by mutagenesis will be necessary to determine which site(s) are phosphorylated.

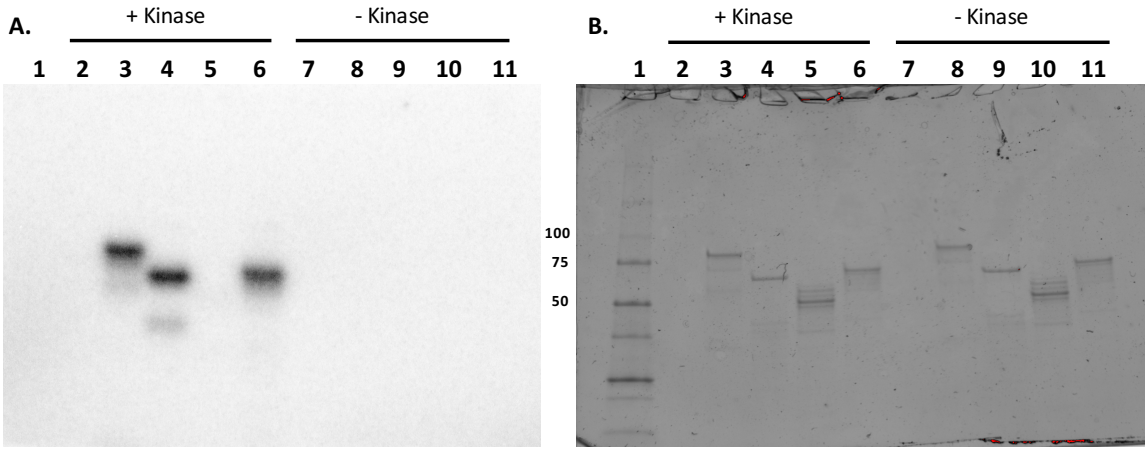


Figure 2.7 **Results of the kinase assay to confirm phosphorylation of the CDPK consensus sequence sites region.** Shown are the phosphor image (A) and the Stain-Free Polyacrylamide Gel (4% - 20%) (Bio-Rad) image (B) of the results of the Kinase assay (Section 2.2.1). (1) Markers, (2) Wheat kinase + no substrate, (3) Wheat kinase + full-length wheat eIFiso4G, (4) Wheat kinase + N-186, (5) Wheat kinase + C-511, (6) Wheat kinase + C-624, (7) no kinase + no substrate, (8) no kinase + full-length wheat eIFiso4G, (9) no kinase + N-186, (10) no kinase + C-511, (11) no kinase + C-624. ~0.25 mg/mL substrate and ~600 ng/mL kinase, final concentration, used per lane indicated.

2.3.7 Expression, Purification and Analysis of Arabidopsis CDPKs

In order to investigate the activity of Arabidopsis CDPKs oneIF4G, eIFiso4G and eIF4B, we obtained 9 Clones for Arabidopsis CDPKs 2, 3, 5, 9, 11, 13, 25, 30, and 32, described in Boudsocq et al. (2012) and were generously provided by Dr. Marie Boudsocq. These 9 CDPKs were representative of 3 of the 4 subgroups of CDPKs found in Arabidopsis, based on sequence homology, (Cheng et al., 2002). The cDNA sequences for these 9 different CDPKs had each been cloned into pGEX-2T vectors, which contain a GST fusion at the N-terminus. Each of the proteins were expressed and purified as was described in Boudsocq et al. (2012) and above in section 2.2.2, *Cloning and Purification of Arabidopsis eIF4G Recombinant Kinase*. The purified proteins were analyzed on a

Stain-Free Polyacrylamide Gel (4% - 20%) (Bio-Rad) and imaged for purity on a Bio-Rad Chemidoc Imager (Figure 2.8).

To analyze the functionality and specificity of the CDPKs the kinases were assayed using native wheat eIFiso4G and histone octamer (positive control), as substrates (Figure 2.9). A third reaction for each CDPK contained no substrate as a negative control. It was hypothesized that some classes might show substrate specificity to eIFiso4G. However, all 9 Arabidopsis CDPKs were seen to phosphorylate wheat eIFiso4G as well as the histone octamer positive control. The kinases also have varying degrees of autophosphorylation (no substrate control lane). No obvious differences were seen between the different classes of CDPKs in this assay which suggests there is no substrate specificity *in vitro* for eIFiso4G or histones. However, differences may occur *in vivo* between various tissue types due to controlled expression of the CDPKs. These CDPKs will be used in further studies on Arabidopsis eIFiso4G1 and eIFiso4G2 to determine if there are any differences in specificity and to identify phosphorylation sites.

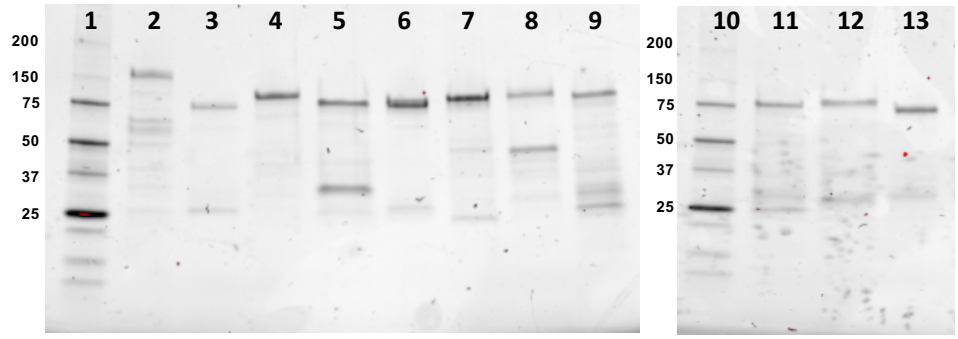


Figure 2.8 **Purification of Recombinant CDPKs.** Image of Stain-Free Polyacrylamide Gel (4% - 20%) (Bio-Rad). Each lane contains ~1 μg of protein. Gel lanes: 1, Markers; 2, CDPK 2; 3, CDPK 3; 4, CDPK 5; 5, CDPK 9; 6, CDPK 11; 7, CDPK 13; 8, CDPK 25; 9, CDPK 30; 10, Markers; 11, CDPK 30; 12, CDPK 32; 13, Recombinant eIF4G wheat germ kinase.

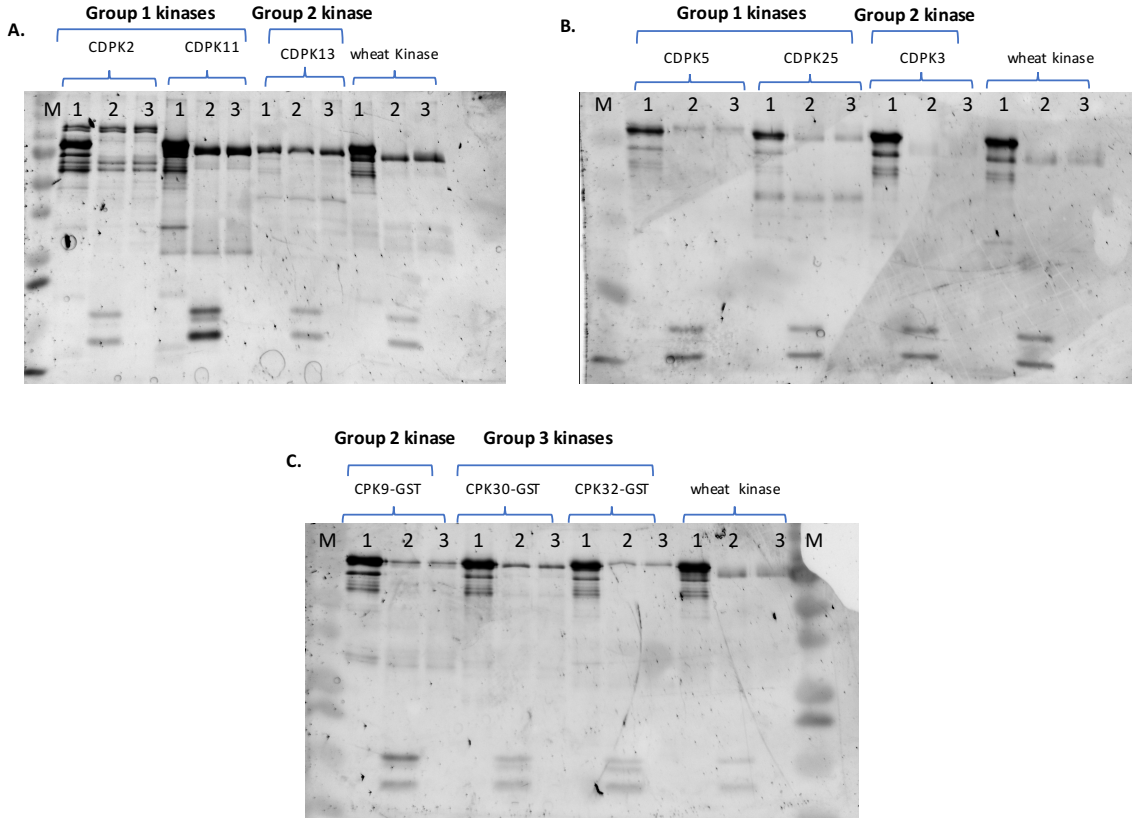


Figure 2.9 **Phosphorylation of wheat eIFiso4G by Arabidopsis CDPKs.** For all three panels, *A, B* and *C*, wheat eIFiso4G (*Lanes 1)* and histone octamer (Lanes 2), and no substrate (Lanes 3). Recombinant wheat eIF4G (4G) kinase was used in all three panelsas a positive control. CDPK2 and CDPK11 from Group 1 kinases and CDPK13 from Group 2 kinases are in Panel A. CDPK5 and CDPK25 from Group 1 and CDPK3 from Group 2 Panel B. CDPK9 from Group 2 and CDPK30 and CDPK32 from Group 3 are in Panel C. A 12% PAGE gel was stained with Pro-Q Diamond Phosphoprotein Gel Stain (Invitrogen).

2.4 DISCUSSION

The full picture of the biological function of plant CDPKs is far from being understood as relatively very little work has been done to fill in the gaps in our understanding (Cheng et al., 2002; Simeunovic et al., 2016). There is also a dearth in our understanding of the regulation of plant protein synthesis (Sesma et al., 2017). Plants use calcium as a second messenger to mediate the plant's response during biotic and abiotic stress and in the growth and development of the plant (Cheng et al., 2002; Curran et al., 2011). Phosphorylation is known to be used in a number of regulatory roles in the life of a plant and is implicated in many more (Curran et al., 2011). Plants have a plant-specific translation initiation factor along with a very unique photosynthetic biochemistry in a stationary existence that most likely utilizes unique methods to regulate the synthesis of proteins needed to address the needs of the plant in a timely way (Browning & Bailey-Serres, 2015). Based on the knowledge that plants have a special group of kinases, CDPKs, and that it is widely known that mammals and yeast use phosphorylation in the regulation of translation, based on our preliminary findings here in this work, it is a plausible hypothesis that plants use CDPKs in the regulation of translation (Simeunovic et al., 2016; Sonenberg & Hinnebusch, 2009).

Due to the increased sensitivity of mass spectrometry in recent years, we have been able to identify an unclassified wheat kinase that had previously been shown to specifically phosphorylate proteins involved in the early steps in the initiation of protein synthesis, wheat eIF4G, eIFiso4G,eIF4B (Figure 1.1) (Humphreys et al., 1988). This kinase had peptides identified by mass spectrometry with high identity with two different Arabidopsis CDPKs and an Snf1-related kinase regulatory subunit. The native wheat kinase was shown to have a calcium dependency (Figure 2.2). The assistance of Claire McWhite (Dr. Edward Marcotte Lab, The University of Texas at Austin), was invaluable to the identification of

all the protein kinases in proteomic data and the selection of a wheat kinase sequence to clone.

The recombinant wheat kinase was able to successfully phosphorylate its target proteins eIF4G, eIFiso4G and eIF4b (Figures 2.3 And 2.4). Mass spectrometry of the phosphorylated substrate, wheat eIFsio4G, revealed it to be phosphorylated at two conserved CDPK phosphorylation sites, (basic-X-X-S/T) in the C-terminus of the protein (Figure 2.5). A kinase assay performed on various truncations of wheat eIFiso4G along with the full-length protein confirms this finding (Figure 2.7).

In order to further investigate if there is any substrate specificity among the various classes of CDPKs we obtained the clones for 9 Arabidopsis CDPKs from Dr. Marie Boudsocq (Boudsocq et al., 2012). All 9 Arabidopsis CDPKs were seen to phosphorylate wheat eIFiso4G along with the histone positive control. Unfortunately, with the wheat eIFiso4G substrate no difference was observed with the Arabidopsis CDPK classes; however, further analysis with Arabidopsis eIFiso4G1 or eIFiso4G2 might show specificity *in vitro*.

CDPKs are thought to carry out multiple functions on various substrates (Curran et al., 2011). CDPKs have already been shown to bind to several targets of 14-3-3 proteins such as transcriptional regulators, metabolic enzymes and ion channels related to stomatal movement, and possibly the 14-3-3 proteins themselves (Simeunovic et al., 2016). They have also been shown to phosphorylate RBohD, a major source of ROS during PTI, and both RBohD and RBohF in order to bring about the required prolonged increase in Ca^{2+} levels required for the effector-triggered immunity response. It is quite possible that eIF4G, eIFiso4G and/or eIF4B could be phosphorylated for the purpose of summoning a particular regulatory protein to bind to the cap-binding complex to bring about a particular

translational response. We also have yet to explore the role of the phosphorylation of a 60S ribosomal protein by CDPKs.

My hypothesis was that this CDPK phosphorylates eIFiso4G at a specific sequence(s) in the regulation of translation. To prove this, in collaboration with the Biobrick Freshman Research Initiative (FRI) (The University of Texas at Austin), we are making wheat eIFiso4G phospho-null and –mimic mutants to confirm the phosphorylation sites and to test the activity of the eIFiso4G in *in vitro* translation assays We are also making Arabidopsis eIFiso4G1 phospho-null and –mimic constructs at sites comparable to the wheat eIFiso4G consensus sites to determine if these constructs can or cannot rescue the strong phenotype in mutant plants lacking eIFiso4G1 and eIFiso4G2. Any mutant that cannot rescue the phenotype would suggest that the mutation affects the activity of the protein *in vivo*.

CHAPTER 3: DO THE CAP-BINDING COMPLEXES USE THE DISULFIDE BOND TO REGULATE PROTEIN SYNTHESIS?

3.1 INTRODUCTION

The cap-binding protein and its plant-specific isoform, eIF4E and eIFiso4E, respectively, both contain a conserved pair of cysteines that are located near the m^7G cap-binding pocket that are not found in other eukaryotes (Figures 1.1 and 1.5) (Monzingo et al., 2007; Patrick & Browning, 2012). These cysteines have been shown to form a disulfide bond in wheat eIF4E under oxidizing conditions (Monzingo et al., 2007). Cysteines are considered highly sensitive targets of ROS-mediated post-translational modifications (Akter et al., 2015). This suggests that formation of these disulfide bonds may be affected by the redox state of the cell and we hypothesize may have a role in regulation of plant translation.

ROS are regularly produced in the plant from various sources, including the chloroplast during photosynthesis, the mitochondria, as by-products of the ETC (electron transport chain), by-products of peroxisomes and during pathogen response (Gao et al., 2014; Mittler, 2017; Moller et al., 2007; Xia et al., 2015). ROS are used in higher order plants in the regulation of differentiation, stress signaling, redox levels, development, systemic responses and cell death. The different types of ROS are singlet oxygen, $^1O_2$, hydrogen peroxide, H_2O_2, hydroxide radicals, OH-, and superoxide, O_2-, all having been formed from atmospheric oxygen through excitation, reduction, enzymatically by superoxide dismutase (SOD) or through free iron in the cell. Each ROS has a unique set of basic characteristics, such as half-life, migration distance, differing reactivities, production sites and different scavenging systems. The scavenging systems, enzyme systems for the removal of ROS, assist in keeping the ROS in cells at non-toxic levels.

Oxidative modifications to proteins are typically irreversible and damaging (Moller et al., 2007). However, due to the wide range of oxidation states that cysteine residues are able to have, many ROS modifications to cysteines, such as disulfide bond formation and ROS triggered S-glutathionation, are reversible, making them excellent indicators of the redox state of the cell and can be used in ROS signaling (Akter et al., 2015). Furthermore, intra- or intermolecular disulfide bond formation, which is primarily carried out by H_2O_2, can function as a regulatory mechanism (Akter et al., 2015; Moller et al., 2007). Proteomic studies have revealed disulfide reduction potentially regulates a multitude of proteins that take part in almost all aspects of plant life (Buchanan & Balmer, 2005; Meyer et al., 2012).

On average, proteins contain approximately 2% cysteine residues (Wall et al., 2012). Although the acid dissociation constant (pK_a) of the thiol group (-SH) on the side chain of cysteines is relatively low in proteins, not all cysteines are readily reactive due to the environment that a particular cysteine residue may be in (Akter et al., 2015; Wall et al., 2012). That is, some thiol groups may be buried too deep in the protein or sterically hindered in some way, preventing accessibility, while for other cysteines the local environment has increased the pK_a of the thiol group, making it less reactive. Features decreasing the pKa and therefore increasing the reactivity of the thiol group to ROS includes the amount of hydrogen bonding available to the cysteine sulfur and the presence of local dipoles and positively charged amino acids (Akter et al., 2015).

There are multiple oxidation states the cysteine thiols can undergo (Akter et al., 2015; Marinho et al., 2014; Moller et al., 2007). Sulfenic acid (Cys-S-OH) is formed by the reaction of a H_2O_2, forming the first and most stable oxidation state of a cysteine residue and is considered to be the point at which all other modifications arise from. In destabilizing conditions, the sulfenic acid can undergo additional oxidations by H_2O_2 to

form sulfinic acid (S-O_2H) and sulfonic acid (S-O_3H), both of which are irreversible and are therefore damaging to the protein. Alternatively, a sulfenic acid can be S-Glutathionylated, form a sulfenylamide by folding back on itself and reacting with a nitrogen on the backbone of the protein, or form a disulfide bond, all of which are reversible.

Disulfide bonds are reduced mostly by way of the oxidoreductases thioredoxin and glutaredoxin that use NADPH as their main reducing agent (Meyer et al., 2012). Plants have a high number of thioredoxins and glutaredoxins but they appear to have redundancy as most thioredoxin and glutaredoxin mutants fail to have any phenotypes.

This chapter explores the sensitivity of the cysteine residues in eIF4E to oxidation as compared with eIFiso4E. Previous work in our lab (Ching-Ying Tseng, unpublished) has revealed that the two conserved cysteines in Arabidopsis eIFiso4E are redox reactive in the presence of $CuCl_2$. My hypothesis, is that eIF4E may differ from eIFiso4E in its sensitivity to $CuCl_2$ and ability to form disulfide bonds *in vitro* or *in vivo* (Carberry et al., 1991; Gallie & Browning, 2001; Monzingo et al., 2007; Patrick et al., 2014; Rodriguez et al., 1998).

3.2 MATERIALS AND METHODS

3.2.1 *In Vitro* $CuCl_2$ Assay to Test for Redox Reactive Thiols

Methoxypolyethylene glycol maleimide-5000 (mPEG-Mal) (Sigma, molecular weight 5,000 Da) is a protein pegylation reagent that covalently attaches to sulfurs of cysteine residues that have been reduced (Peled-Zahavi et al., 2010). N-ethylmaleimide (NEM) (Sigma Aldrich, molecular weight 125 Da) is also used to block sulfides that are reduced. Two different reactions were carried out for each protein, one with $CuCl_2$, for oxidizing the cysteine residues to form a disulfide bond and one with DTT for reducing all cysteine residues (Figure 3.1).

Starting with 130 picomoles of the recombinant forms of *At* eIF4E and *At* eIFiso4E, one set of reactions with each protein contained 91μM $CuCl_2$, in order to fully oxidize all sulfides, with 50 mM Tris, pH 7.6, added to bring the reaction to a final volume of 200 μL. The second set of reactions with each protein contained 180 μM DTT, which ensures that all sulfides are reduced to start the reaction, also with the addition of 50 mM Tris, pH 7.6, added to a final volume of 200 μL. This reaction was incubated for 1 hour at 25˚C. The reactions were then added to 10,000 MW cut off 0.5 mL concentrators (Amicon Ultra), followed by 3, 400 μL washes with 200 mM NEM to block free cysteine residues in labeling buffer (6M urea, 200 mM Bis-Tris (pH 6.5), 0.5% (w/v) SDS, 10 mM EDTA), centrifuging the samples at 14,000 x g for 8 minutes at 4˚C with each wash. The samples were then transferred to a new Microfuge tube and incubated in the dark for 1 hour at 25˚C. At the end of the incubation, the samples were

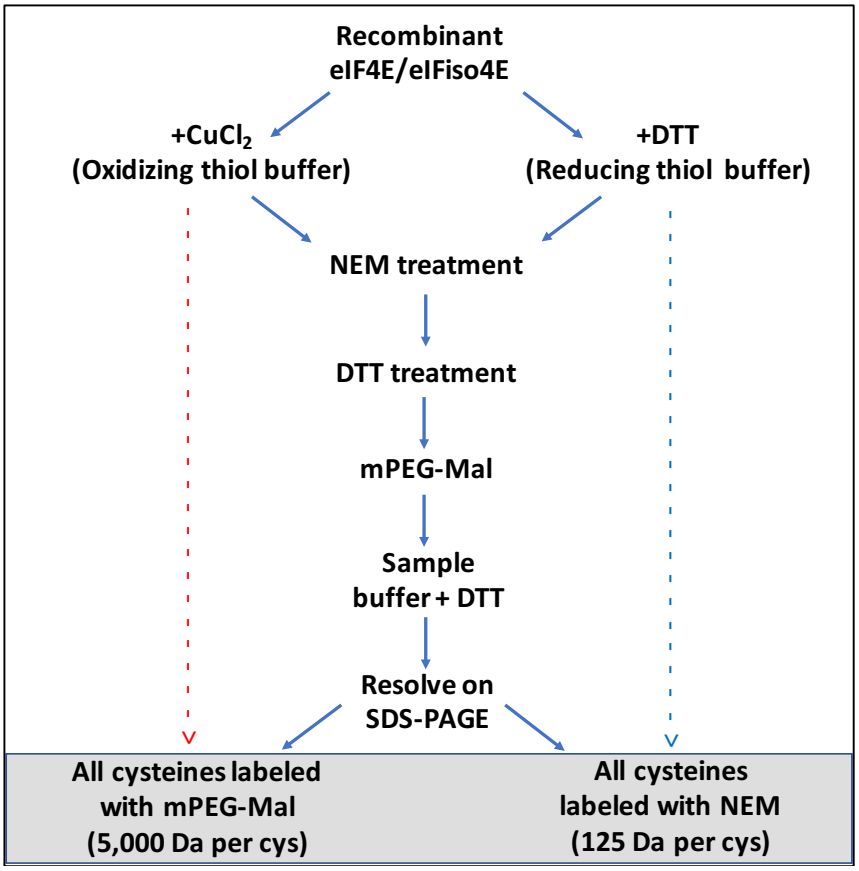


Figure 3.1 **Flow chart of *in vitro* $CuCl_2$ assay to test for redox reactive thiols.** Starting with two samples of either recombinant eIF4E or eIFiso4E, one sample is incubated in $CuCl_2$ to oxidize all the available cysteines while the other sample is incubated in DTT, to reduce all the cysteines. Both samples are then incubated in NEM (125 Da) to label all thiols (R-S-H). Following that, a DTT treatment then ensures that all available thiolates ($R\text{-}S^- H^+$) are reduced to form thiols. Thiols are then labeled in a treatment with mPEG-Mal (5,000 Da). Sample buffer and DTT is then added to the reactions and resolved on an SDS-PAGE gel. Possible outcomes (gray box): All cysteines that start the reaction oxidized are labeled with mPEG-Mal (red dotted line) and all the cysteines that started the reaction reduced are labeled with NEM (blue dotted line).

transferred to clean concentrators and washed using the same procedure as above but using 100 mM DTT in the labeling buffer instead of NEM for the 3 washes. The samples were then transferred to new Microfuge tubes and incubated for 1 hour at 25˚C. Again, the samples were transferred to new concentrators and washed as above except using labeling

buffer for the washes, and washing the sample to 100 μL with the third wash. The 100 μL of sample was added to a new Microfuge, with the addition of 100 μL of 50 mM of mPEG-Mal and then incubated in the dark for 1 hour at 25˚C. 2 μL of 1 M DTT was then added to the reaction to stop it and then precipitated overnight in 4 times the volume of cold acetone at -20˚C. The samples were then centrifuged in a microfuge at 14,000 rpm at 4˚C for 5 minutes, followed by a 1 mL acetone wash and centrifuged. The pellets were air dried 10 minutes then resuspended with 60 μL of 2X SDS sample buffer with 10 μL of each of the 4 reactions electrophoresed on two different 12.5% SDS-PAGE gels, one for a western for eIF4E and the other for eIFiso4E. Purified recombinant *At* eIF4E and purified recombinant *At* eIFiso4E were loaded as standards (Mayberry et al., 2007).

The proteins from each gel were transferred to Immobilon-P PVDF (polyvinylidene fluoride, Millipore Sigma) using the Bio-Rad Trans-Blot Turbo Transfer System. The membranes were blocked overnight at 4˚C with rotation in 5% non-fat dried milk, previously boiled, filtered and cooled, in HNAT buffer (10 mM Hepes-KOH, pH 7.6, 150 mM NaCl, 0.2% BSA, 0.2% Tween 20). The blocking solution was poured off and the membranes briefly rinsed in HNAT buffer followed by probing with rabbit polyclonal antibody raised against recombinant *At* eIF4E (1:5000) and *At* eIFiso4E (1:1000) in HNAT buffer at room temperature (RT) with rotation for 1 hour (Mayberry et al., 2011). The membranes were then washed 3 times with HNAT buffer, 5 minutes each wash. Goat-anti rabbit antibodies conjugated to horse radish peroxidase (Kirkegaarge Perry Laboratories) were then used to probe the membranes in a 1:20,000 dilution at RT with rotation for 1 hour. The membranes were then washed 3 times at 5 minutes each in HNAT and visualized using the LumiSensor Chemiluminescent HRP Substrate Kit (GenScript) and a Bio-Rad Chemidoc Imager.

3.3 RESULTS

3.3.1 *In Vitro* $CuCl_2$ Assay to Test for Redox Reactive Thiols

Previously, *At* eIFiso4E was shown to contain redox reactive thiols using mPEG-Mal labeling (Ching-Ying Tseng, unpublished). Here, to test *At* eIF4E for redox reactive thiols and to compare it with *At* eIFiso4E, recombinant *At* eIF4E and *At* eIFiso4E were treated in separate samples with either $CuCl_2$ or DTT in each sample in order to ensure the full oxidation or the full reduction, respectively, of all the cysteine sulfurs, *At* eIF4E (3) and *At* eIFiso4E (2) (Figure 3.1). All four samples were then treated with NEM in order to block all reduced cysteines followed by treatment with mPEG-Mal to label all oxidized cysteines, cysteines in a disulfide bond or otherwise blocked, adding 5,000 Da per cysteine labeled.

There was neither a complete oxidation or reduction for any of the samples as would have been expected. All three cysteines in *At* eIF4E were labeled in the +$CuCl_2$ sample as indicated by the light band seen at the 3 mPEG-M label in Figure 3.2, but not as completely and not as robustly as was seen with the labeling in *At* eIFiso4E. It appears that only one *At* eIF4E cysteine is preferentially labeled as opposed to having 2 or all 3 labeled, while there appears to be no real preference for one or two cysteines labeled for the *At* eIFiso4E +$CuCl_2$ sample. The *At* eIFiso4E sample had a large amount of 0 mPEG-Mal labeling. This is most likely due to a need for more $CuCl_2$ in the reaction, as the urea in the buffer would have insured the accessibility of the cysteines due to the protein's denaturation.

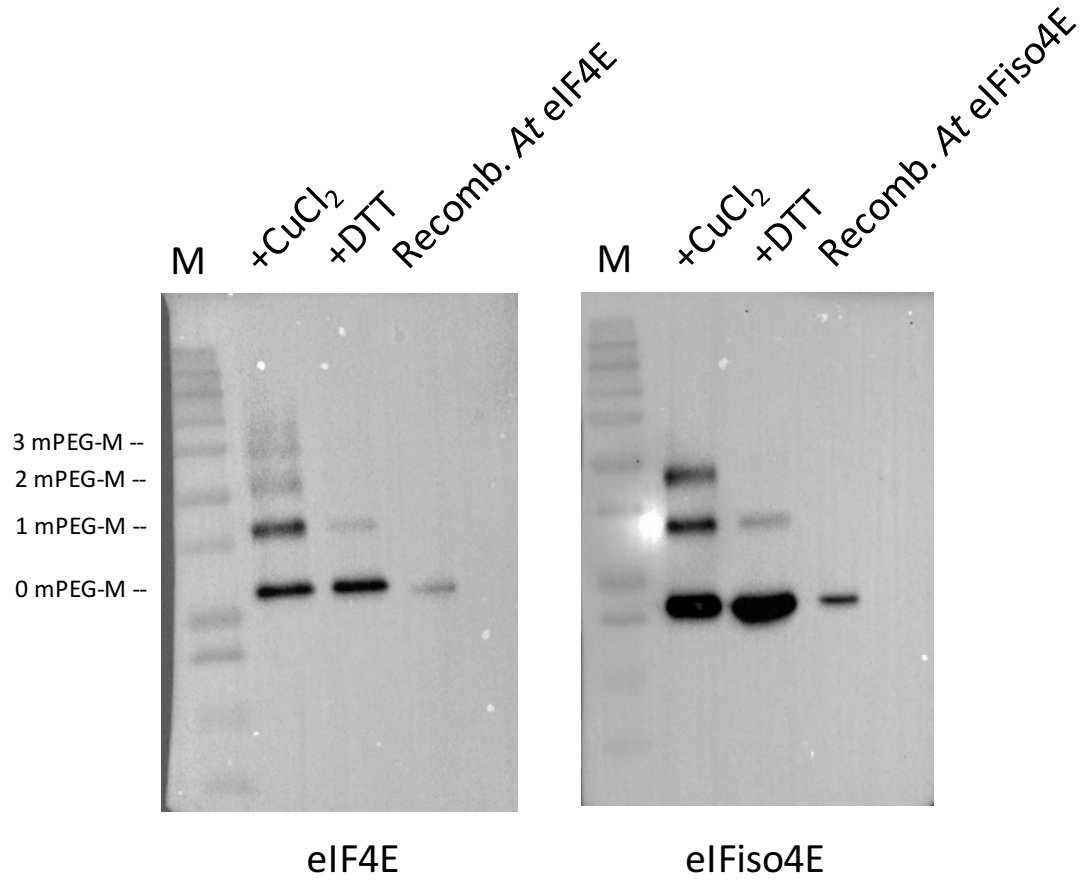


Figure 3.2 **Redox Reactive Thiols in *At* eIF4E and *At* eIFiso4E.** Reactions containing recombinant *At* eIF4E (left) or *At* eIFiso4E (right) were both incubated with $CuCl_2$ and DTT, followed by labeling with mPEG-Mal. The reactions were loaded on a PAGE gel and probed by Western Blot with their respective antibodies. "0 mPEG-M" marks the location indicating no mPEG-Mal label on the protein, "1 mPEG-M" indicates 1 cysteine has the mPEG-Mal label, "2 mPEG-M" indicates 2 cysteines and "3 mPEG-M", 3 cysteines are labeled.

3.4 DISCUSSION

Plants do not appear to use the typical eukaryotic methods for the regulation of translation (Browning & Bailey-Serres, 2015). There have been no 4E-BPs identified in plants and only one eIF2α kinase has been found in plants as opposed to the 4 different eIF2α kinases found in mammals. The apparent absence or near absence of these well conserved and well characterized methods of eukaryotic regulation in plants suggests that plants possibly use an alternative mechanism to regulate protein synthesis.

Plants possess a unique set of well conserved cysteines in the eIF4E-family of cap-binding proteins not found in other eukaryotes (Patrick & Browning, 2012). Their proximity and placement near the cap-binding pocket makes them appear to be perfectly poised for possible use in the regulation of translation (Figure 1.5) (Monzingo et al., 2007). Structural studies of wheat eIF4E and eIFiso4E have shown that wheat eIF4E can form a disulfide bond and both wheat eIF4E and eIFiso4E exhibit structural changes in oxidizing conditions that indicate participation by the cysteines in redox chemistry (Monzingo et al., 2007; O'Brien et al., 2013a). Proteomic studies on Arabidopsis and *Midicago truncatula* have revealed several initiation factors, including eIF4A, eIF5A1 and eIF5A2, to be targets of redox chemistry (Alkhalfioui et al., 2007; Yamazaki et al., 2004). Given that plants have a high level of ROS due to the added biochemistry of photosynthesis, that plants have a very high number of types of thioredoxins and glutaredoxins and that the list of proteins involved in ROS targeted regulation is growing, it stands to reason that these conserved disulfide bonds in plant cap-binding proteins could indeed be a mechanism of the regulation of the initiation of translation (Browning & Bailey-Serres, 2015; Meyer et al., 2012).

Not all cysteines are necessarily redox reactive due to their location and environment (Akter et al., 2015). In this study, we set out to determine the reactivity of

the cysteines in Arabidopsis eIF4E as compared with Arabidopsis eIFiso4E. eIF4E does not appear to be as redox reactive as eIFiso4E and it appears that both cysteines involved in the disulfide bond are not as reactive as the eIFiso4E cysteine residues (Figure 3.2). eIFiso4E shows a dark band at the 2 mPEG-M marker, indicating that both cysteines had been pegylated as opposed to a quite light signal seen at the 2 mPEG-M marker with eIF4E, with just one labeled cysteine being favored over 2. Had the two cysteines in eIF4E been equally reactive, a definitive signal at the 2 mPEG-M mark would have been expected.

More studies are needed to be done to substantiate this result and to explore these cysteines further. C/S mutations of both cysteines in eIF4E and eIFiso4E need to be made to determine which is more reactive in *in vitro* studies. Furthermore, *eIF4E* C/S plant lines need to be made to determine the effect, if any, of the mutation. While no obvious phenotype was observed for *eIFiso4E* C/S mutant lines, more careful phenotypic analysis under a variety of environmental stresses need to be carried out. In addition, a study of the redox state of eIF4E and eIFisoE in plants with various mutations that generate a more oxidizing environment *in vivo* will be informative. Preliminary data from a collaboration with Dr. Dixie Goss (Hunter College) to determine the binding constant for m^7GTP for oxidized eIF4E indicates that the oxidized form binds m^7GTP ~3-4X more tightly, whereas a mutant that cannot form the disulfide bonds does not change binding affinity under oxidizing conditions. Further experiments will need to be carried out with eIFiso4E to determine if there is a similar change upon oxidation. Increasing the binding affinity for m^7G capped mRNAs under oxidizing conditions could have two effects, one preventing the release of mRNAs and thus reducing translation or increasing affinity for mRNAs or certain mRNAs. Further analysis will be necessary to determine what effects oxidizing conditions *in vivo* of the cap-binding proteins may have on translation initiation.

CHAPTER 4: IDENTIFYING SPECIFIC POPULATIONS OF MRNA ASSOCIATED WITH PLANT PROTEIN SYNTHESIS INITIATION FACTORS EIF4F AND EIFISO4F

4.1 Overview

Several methods of regulation of translation in mammals and yeast are well characterized, such as eIF2α phosphorylation and 4E-BPs (Sonenberg & Hinnebusch, 2009). However, the regulation of translation in plants is less well understood (Browning & Bailey-Serres, 2015). Plants are unique from all other eukaryotes in that they contain a plant-specific isoform for the cap-binding complex eIF4F, eIFiso4F (Figure 1.1). Based on previous findings, we hypothesize that eIFiso4F may preferentially bind specific mRNAs (Gallie & Browning, 2001; Monzingo et al., 2007; O'Brien et al., 2013a; Patrick & Browning, 2012; Patrick et al., 2014; Rodriguez et al., 1998). With this in mind, we sought to determine if there are client mRNAs for either eIF4F or eIFiso4F using two different methods, RNA immunoprecipitation and ribosome profiling. RNA-IP from wild type plants should provide a list of mRNAs in direct contact with eIFiso4F, eIF4F or both forms. Ribosome profiling from plants lacking the large subunit of eIFiso4F or eIF4F compared to wild type plants should provide a glimpse of which mRNAs are no longer being translated in these mutants or if there are differences in ribosome loading and/or transit.

4.2 RNA Immunoprecipitation of Arabidopsis eIF4F and eIFiso4F

4.2.1 INTRODUCTION

RNA binding proteins are key to a number of cellular activities such as transcription, RNA editing, mRNA splicing, the exporting of RNA from the nucleus to the cytoplasm, and translation (Maquat & Carmichael, 2001). While approximately half of the RNA binding proteins in Arabidopsis are homologous to eukaryotic RNA binding proteins, the other half are plant-specific (Terzi & Simpson, 2009). How these plant-specific RNA binding proteins function may be novel RNA processing event(s) or carry out novel regulatory functions in RNA metabolism (translation, degradation, etc.) that are unique to plants.

An *in vivo* method of capturing the true targets of RNA binding proteins is necessary to accurately identify the RNA a given protein associates with (Zielinski et al., 2006). Problematic with some *in vivo* approaches, such as immuno- or affinity-purification methods, is the re-association of the RNA with non-target proteins when working with a cell lysate (Mili & Steitz, 2004). Cross-linking the RNA to the protein overcomes this problem. (Terzi & Simpson, 2009). There are important considerations in choosing a method of crosslinking, such as stability of the cross-link and the proximity of the linkage (Steen & Jensen, 2002). While UV cross-linking is a widely-used and acceptable method of crosslinking that fulfills these conditions in mammalian or yeast cells, it presents problems when working with plant cells due to the thick cell wall, along with the fact that UV radiation can induce damage to the proteins and nucleic acids (Terzi & Simpson, 2009; Zhang et al., 2016). Formaldehyde cross-linking creates a stable, covalent, yet reversible bond, at 2 Å, and leaves the molecules undamaged (Hoffman et al., 2015). RNA immunoprecipitation is an *in vivo* approach for determining the direct targets of an RNA binding protein that can be partnered with formaldehyde cross-linking (Niranjanakumari

et al., 2002; Peritz et al., 2006; Terzi & Simpson, 2009). This can be followed by a number of techniques such as RT-PCR, qRT-PCR, and high-throughput next-generation sequencing, to analyze the RNA that has been captured (Terzi & Simpson, 2009).

Using RNA immunoprecipitation, we set out the determine the role of the plant-specific cap-binding complex eIFiso4F in Arabidopsis (Figure 1.1). More specifically, to determine if there are any specific populations of mRNA that may associate with either eIF4F or eIFiso4F or both. Our hypothesis, based on previous findings, was that eIFiso4F preferentially binds specific mRNAs (Browning & Bailey-Serres, 2015; Gallie & Browning, 2001; Monzingo et al., 2007; O'Brien et al., 2013a; Patrick & Browning, 2012; Patrick et al., 2014). Our goal was to determine the identity of any mRNAs specifically linked to each isoform. To carry out this experiment, we used antibodies specific to eIF4G or eIFiso4G with 14-day old wild type Columbia Arabidopsis seedlings. Antibodies specific to eIF4G were used with *eIF4G* mutant seedlings and beads coated without any antibodies were used as negative controls. RNAs pulled down by this method were identified by high-throughput next-generation sequencing.

4.2.2 MATERIALS AND METHODS

In order to carry out the full RNA immunoprecipitation procedure, there are 2 preparatory protocols required to follow first. Cross-linking antibody to Protein A MagBeads (Genscript) and formaldehyde crosslinking the tissue (see Appendix II, Protocols 1 and 2, respectively). Both can be done several days in advance as the beads will be stored in sodium azide as a preservative and the tissue stored at -80°C. Details of the immunoprecipitation procedure are in Appendix II, protocol 3. The RNA IP samples were sequence at the GSAF (University of Texas at Austin) and analyzed using various methods including the entire Cufflinks package (Trapnell et al., 2012).

4.2.3 RESULTS AND DISCUSSION

Our results are considered inconclusive at this point with there being much noise, and no significant differences found between negative controls and RNA IP samples. Multiple errors inadvertently occurred at different stages of the experiment that may have compromised the experimental design.

Possibly contributing to the noise in the data is the large size of both eIF4G (~186,000 Da) and eIFiso4G (~86,000 Da) and the possibility of mRNA being cross-liked to eIF4G due to sheer proximity of the molecules to each other and not necessarily specificity. Other possible problems included the fact that we only had 2 replicates for each sample, contributing to noise (Sims et al., 2014). More replicates would have helped improve the data quality. Furthermore, in the preparation of the libraries by the sequencing facility, it was assumed that our mRNA had a 5' phosphate on them when indeed they should have had in theory a m^7G on the 5' end of the mRNA, assuming that the initiation factors were binding in proximity to the 5' cap group. Due to this, the messages that still

happened to have the m^7G on them would most likely not have been included in the library that was made for sequencing. While there could have been enough degradation of the mRNA that the caps had been removed, there is no way of determining that at this point. Finally, we most likely did not have the samples sequenced deeply enough as we only had a reading depth of 10 million reads for each sample (Conesa et al., 2016; Sims et al., 2014). Careful consideration of these details would need to be made to repeat this experiment. However, this data will be re-examined in light of the success in obtaining ribosome profiling data for wild type, the *eIF4G* mutant and *eIFiso4G* mutant (see next section). It is possible that despite the noise, there may be some correlations to be made and corroborated by smaller scale RNA-IP and qRT-PCR.

4.3 Ribosome Profiling

4.3.1 INTRODUCTION

The translational landscape of a cell is a reflection of the state of the cell at any point in time (Juntawong et al., 2015; Liu et al., 2013). That is, ribosomal occupancy is the ultimate response to the internal and external environments of the plant and the protein needs being created by those environments, in conjunction with the cell's response to those needs. How these changes at the translational level occur is a result of regulatory mechanisms acting upon the translational apparatus. While eukaryotic regulation of translation in mammals and yeast is well characterized, how plants regulate translation is still largely unknown (Browning & Bailey-Serres, 2015). Plants have a unique translational factor that other eukaryotes do not have, the complex eIFiso4F (Figure 1.1). Could some level of regulation be carried out by the cap-binding complex using some unique plant-specific method? Might there be mRNA specificity linked to the two cap-binding complexes, eIF4F and the plant specific isoform eIFiso4F?

Ribosome profiling (also known as Ribo-seq) is a snapshot in time of the exact mRNA sequences being translated at that moment and the position of the ribosomes on the mRNA (Brar & Weissman, 2015; Ingolia et al., 2012; Juntawong et al., 2015; Liu et al., 2013). It combines RNase footprinting, in which RNase is used to digest away all the mRNA that is not bound by the ribosomes, which have been flash frozen to instantly halt translation, followed by deep sequencing of the recovered fragments to determine the precise placement of translating ribosomes on an mRNA. This is in contrast to polysome profiling in which the entire message is isolated from the polysomes (Chasse et al., 2017).

Ribosome profiling can be used to determine alternatively spliced transcripts, the translation of uORFs or sORFs, novel start codons, the empirical start and stop sites of a protein, the rate of protein synthesis, and regulatory pauses (Brar & Weissman, 2015;

Ingolia et al., 2012; Juntawong et al., 2015). Interestingly, ribosome profiling has also revealed instances of the property of proportional synthesis in *Escherichia coli* and *Saccharomyces cerevisiae* (Brar & Weissman, 2015). That is, the subunits of a multiprotein complex being synthesized in amounts matching the stoichiometry of the complex. Another use for ribosome profiling has included the discovery of gene cluster functions found using multiple measurements of protein synthesis in budding yeast over a period of time (Brar & Weissman, 2015).

Strengths of ribosome profiling include sensitivity due to the depth of sampling, the measurement of protein synthesis at an instant in time and precise positional information (Brar & Weissman, 2015; Ingolia et al., 2012; Juntawong et al., 2015). Also, the integrity of the mRNA is also better preserved than with polysome profiling since only the ribosome protected fragment is needed as opposed to the full-length mRNA (Ingolia et al., 2012). Some weaknesses or things to consider in the interpretation of the data are systematic errors that can be associated with the method of halting the ribosome, such as the use of cyclohexamide (Brar & Weissman, 2015). While such inhibitors generally do not affect the global distribution of the density of ribosomes on mRNA, there can be local effects, particularly at the translational start site, in which the signal may be blurred, making it difficult to determine specific positions of the ribosomes. The use of flash freezing has been shown to minimize this issue. There can also be inaccuracies due to the mapping of repetitive sequences or very similar sequences. Computational methods exist to minimize this *in silico*. Furthermore, there can be contamination of the sample due to other RNA in the sample having the same size as a footprint, ~30 nucleotides. Again, computational methods exist to minimize these issues along with procedural methods to remove rRNA fragments from the sample. A final weakness includes the fact that ribosome profiling requires a much larger sample than RNA-seq, a transcriptome profiling method

that measures the levels of all the RNA in a cell at a specific point in time using deep sequencing (Wang et al., 2009a). This is due to the fact that the total amount of ribosome footprints in a cell is a much smaller amount than the total RNA levels, along with the extra step of isolating the ribosomes (Brar & Weissman, 2015).

RNA-seq is commonly used in conjunction with ribosome profiling as it provides a useful comparison of the messages being translated at any point in time with the complete set of messages in the cell at the same point in time (Brar & Weissman, 2015). Using this knowledge, the translational efficiency of a particular mRNA, the ratio of the footprints of that mRNA to the total abundance of that mRNA in the cell, can be determined (Juntawong et al., 2015). This is particularly helpful in understanding a disparity that oftentimes exists between mRNA levels and protein levels in a cell and can help to determine major points of translational regulation (Ingolia et al., 2012).

What actually gets translated in Arabidopsis is dependent upon what mRNA is bound by either of the cap-binding complexes and loaded to the ribosomes (Browning & Bailey-Serres, 2015). The lack of eIFiso4G has a profound phenotype suggesting some subset of mRNAs are not being expressed. Our hypothesis is that eIFiso4F preferentially binds and translates specific mRNAs specific to plant fitness (Gallie & Browning, 2001; Monzingo et al., 2007; O'Brien et al., 2013a; Patrick & Browning, 2012; Patrick et al., 2014). Using ribosome profiling of mutants lacking eIFiso4G or eIF4G in conjunction with RNA-Seq we seek to determine the client mRNAs for both isoforms.

4.3.2 MATERIALS AND METHODS

4.3.2.1 Ribosome Profiling

The ribosome profiling experiments were carried out in the laboratory of Dr. Julia Bailey-Serres (University of California Riverside) over a 4-week period last January where I learned how to do these types of experiments. To carry out ribosome profiling, we followed two different protocols making a few modifications. Starting with 5 mL of pulverized, packed and frozen tissue, we began by using a published protocol for the *Conventional Isolation of Polysomes by Use of Differential Centrifugation* found in (Juntawong et al., 2015). We first followed all the preparation steps required for that protocol as found in section 2, making only one adjustment. For the preparation of the sucrose gradients, section 2.4.3, we changed the volumes found in Table 2 for some of the layers. For the 60% sucrose layer, instead of 0.75 mL per layer, we used only 0.100 mL and for the 15% sucrose layer we used 1.300 mL. This was done because we are really only interested in the monosomal and disomal fractions and not interested in the polysomal fraction, which would be found in the 60% layer.

After section 2, we proceeded to section 3.1, *Conventional Isolation of Polysomes*. The only change we made in this protocol was in step number 16. Instead of centrifuging the polysome pellet for 3 hours at 170,000 x g (50,000 rpm) at 4°C, we centrifuged the pellet for 18 hours at 90,000 x g (35,000 rpm) at 4°C. This was done as more of a convenience factor. We then completed the entire protocol in section 3.2, *Generation of Ribosome-Protected Fragments,* with no changes.

Section 3.2 was followed by section 3.4, *Ribosome-Protected Fragment Size Selection*. The only change made to this protocol was in step 19. Instead of resuspending the pellet in 10 μL of RNase-free water, we resuspended the pellet in 10 μL of 10 mM Tris,

pH 8. For the remaining steps, we used the protocol found in *The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments* (Ingolia et al., 2012).

We started at step 28, the dephosphorylation reaction, and continued on until step 64, which included the dephosphorylation reaction and the linker ligation step, with the following changes. At step 36, we did an rRNA removal step not included in the protocol, using Ribo-Zero rRNA Removal Kit (Illumina) for plants. For this we followed the protocol that came with the kit which, in brief, is prepare the Ribo-Zero Microspheres (beads) as the protocol instructs, treat the RNA with Ribo-Zero rRNA Removal Solution, incubate the treated RNA with the beads followed by ethanol precipitation of the treated RNA. To ethanol precipitate the RNA, we followed step 26 except using 2 times the amount of GlycoBlue called for and adding 900 μL of isopropanol instead of 500 μL. We also added a 75% ethanol wash for the pelleted RNA as there was not one included in the protocol.

We then followed the protocol to step 64 to finish making the library, which included reverse transcription, circularization, another rRNA depletion step, PCR amplification and barcode addition, with one change at the PCR amplification and barcode addition step. For step 55, preparation of the PCR reaction, we used the KAPA HiFi PCR Kit (Kapa Biosystems) using the KAPA reagents along with Universal (Forward) Primer, found in Juntawong et al. (2015) and Reverse Indexed primers found in *BrAD-seq: Breath Adapter Directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction* (Townsley et al., 2015). The components and the amounts used in this reaction are given in Appendix II, Protocol 4.

For all of the steps in the Ingolia et al. (2012) protocol that included a gel extraction step, we chose option B, the overnight gel extraction method, in step 25, using 750 μL of

the RNA extraction buffer, given in the protocol, instead of 400 μL, incubating the samples at -80˚C instead of on dry ice.

The samples were then submitted to the DNA Technologies & Expression Analysis Cores at the University of California at Davis Genome Center to quantify the library, using the Agilent BioAnalyzer, and for sequencing using the Illumina HiSeq4000 system with a sequencing depth of approximately 40 million reads.

4.3.2.2 Total RNA-seq

To carry out RNA-seq, we used a combination of two protocols, the protocol used in *A Low-Cost Library Construction Protocol and Data Analysis Pipeline for Illumina-Based Strand-Specific Multiplex RNA-Seq* (Wang et al., 2011), modified with a few steps found in *BrAD-seq: Breath Adapter Directional sequencing; a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction* (Townsley et al., 2015).

Starting with 1 mL of pulverized, packed and frozen tissue per sample, split into 3 different RNase-free tubes, a Trizol reagent isolation of the RNA was carried out. The RNA isolated from the three tubes will be re-combined in the last step for each sample. Briefly, 1 mL of Trizol reagent was added to each of the 3 tubes for each sample, mixed and then incubated at RT for 5 min. 200 μL of chloroform was then added to each tube, mixed and incubated another 5 min. at RT. Centrifuge at max speed in a table top centrifuge, 10 min. Transfer the supernatants to a new tube and add 600 μL 100% isopropanol, 20 μL sodium acetate, pH 5.2, and 2 μL glycogen (5 mg/mL, Ambion AM9510) to each tube. Incubate at – 80˚ C for 30 min. Alternatively, this could be done at -20˚ C overnight. Centrifuge for 30 min. at max speed at 4˚ C. The pellets were washed with 75% EtOH, dried 5 min. at RT and dissolved in 20 μL RNase-free water. All 3 pellets were sequentially dissolved in the same 20 μL of water.

We then followed the steps found in Wang et al. (2011) starting with *Protocol 1: Preparation of mRNA fragments*, completing both the mRNA purification and the mRNA fragmentation steps. We then completed all of Protocol 2, cDNA synthesis which included reverse transcription and second strand synthesis. After that, we followed all of Protocol

3, which includes end-repair, dA-tailing and Y-shape adapter ligation, although we used the Y-shaped adapter found in Townsley et al. (2015).

Finally, Protocol 4 in Wang et al. (2011), was used for the dUTP excision step but for the PCR reaction, but we followed the method found in Townsley et al. (2015) using 8 bp barcodes with 15 cycles. These are standard Illumina 8 pb barcode primers, for which the sequences are found in Townsley et al. (2015). The final clean-up was done as in the Wang et al. (2011) protocol. The samples were then bioanalyzed and multiplexed with equal ng amounts of each sample and sent to be sequenced.

4.3.3 RESULTS AND DISCUSSION

The ribosome profiling experiment is an ongoing collaborative effort with the laboratory of Dr. Julia Bailey-Serres, University of California, Riverside. The Bailey-Serres Laboratory is very experienced in using this technique in plants (Juntawong et al., 2014; Juntawong et al., 2015). We used frozen, ground tissue from 10-day old wild type Arabidopsis (Columbia; Col-0), *eIFiso4G1* x *eIFiso4G2* (14B) and *eIF4G* (4G) seedlings harvested at two different time points, end of day, 10:45 pm, and extended night, 12:45 pm (Figure 4.3.1). Extended night, meaning that the plants were kept in the dark 6 hours beyond their typical start of daytime, 7:00 am. These two time points were chosen based on the results of a metabolic assay previously done by our laboratory in collaboration with Johannes Hanson (Utrecht University; unpublished data). Of the four different time points (end of day, before dawn, extended dark and 6 hr after dawn) used in the metabolic assay, end of day and extended dark appeared to have the biggest effect on the mutants versus the wild type plants.

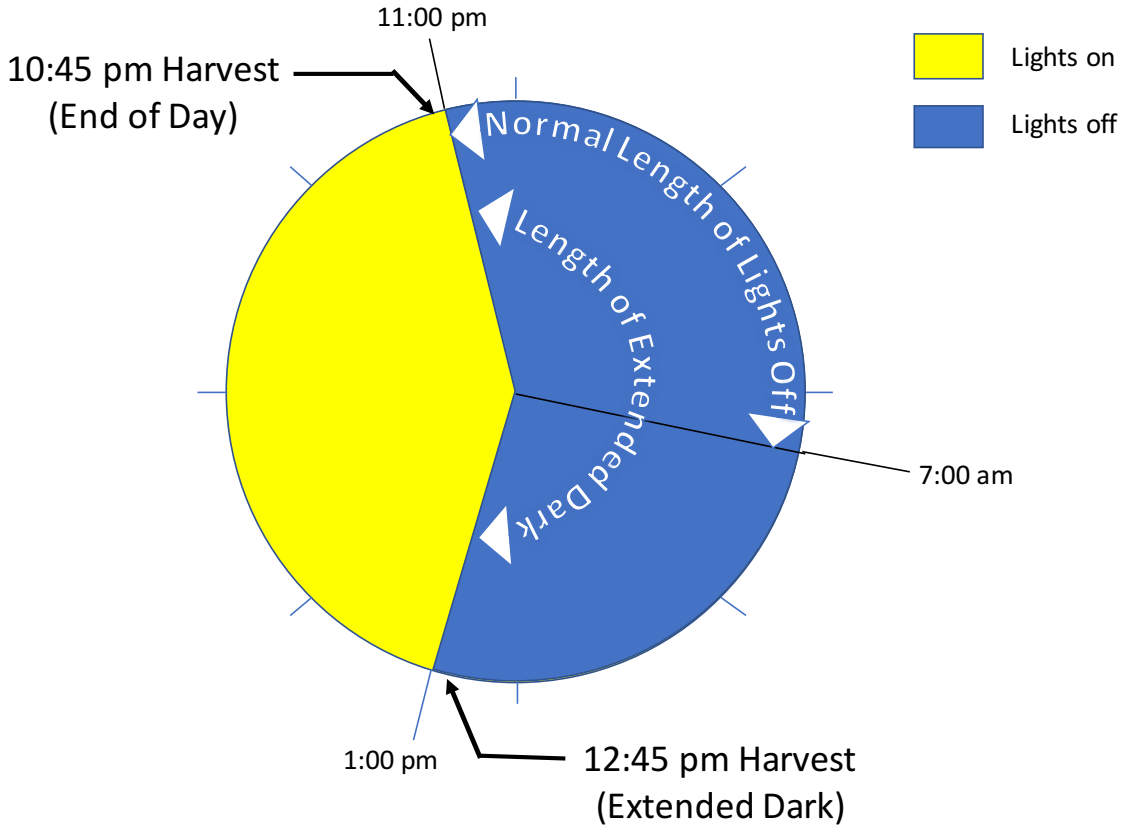
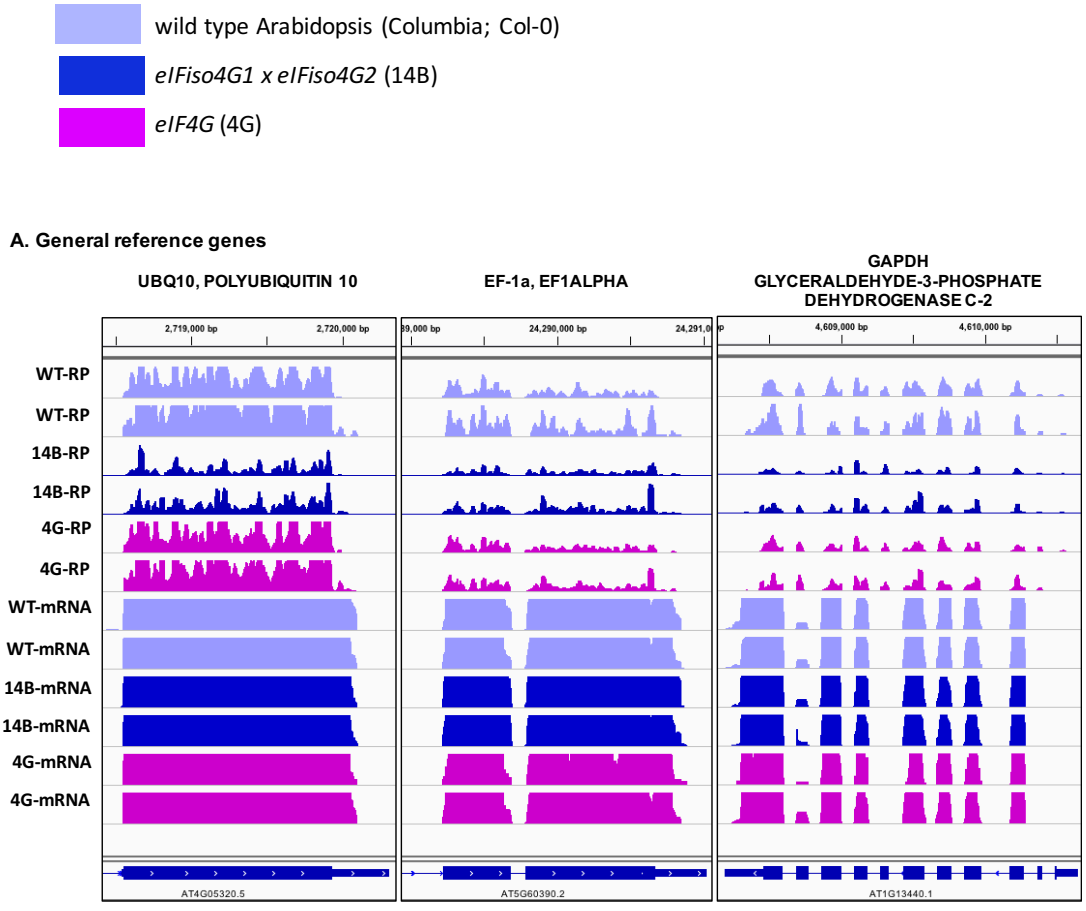


Figure 4.1 **Timing of tissue harvests for the ribosome profiling experiment.** Two harvests were carried out on each of the three plant types, wild type Arabidopsis (Columbia; Col-0), *eIFiso4G1* x *eIFiso4G2* and *eIF4G* 10-day old seedlings. The normal length of night for the plants is from 11:00 pm until 7:00 am, indicated by the white bidirectional arrow with the wording "Normal Length of Lights Off". For this experiment, the plants were kept in the dark for an extended length of time, that is, the normal length of lights off plus 6 hours past normal length of lights off, as indicated by the bidirectional arrow with the wording "Length of Extended Dark". The first harvest was carried out at the end of the day at 10:45, in the light. The second harvest was carried out in the dark at 12:45 pm.

Samples obtained from the ribosomal profiling experiment have been sequenced on the HiSeq 4000 System (Illumina) and are presently being analyzed using various computational analysis methods. Unfortunately, for many of the samples, not enough tissue was used therefore additional deep sequencing is needed, as well as repeating some

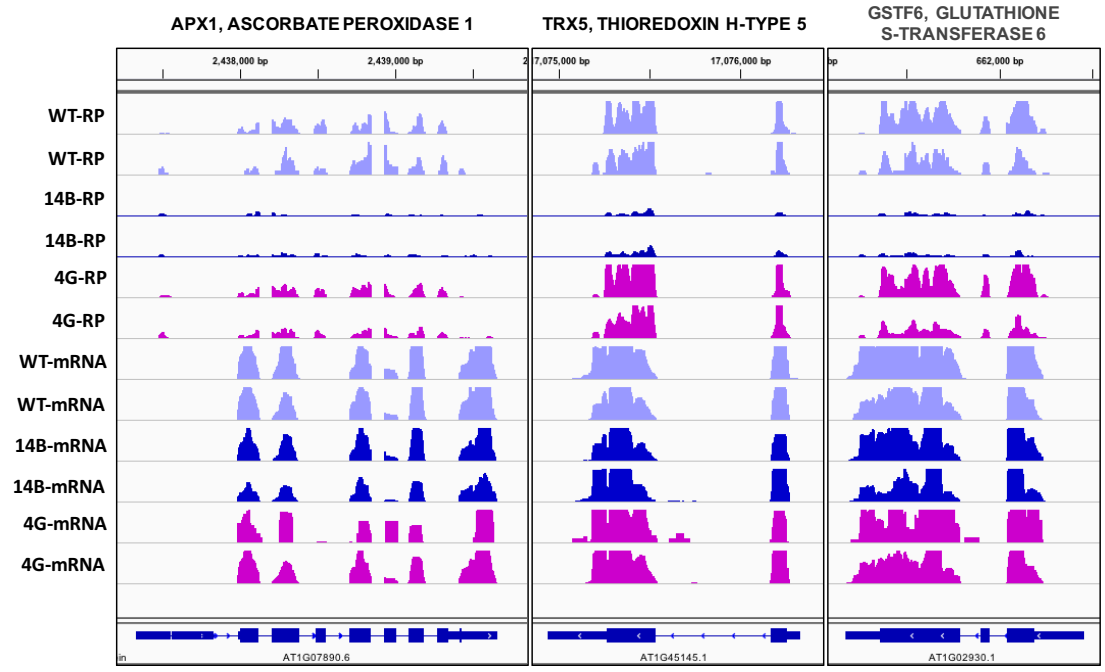
of the replicates, before complete data analysis commences. Early analysis of the data, using the Integrative Genomic Viewer (IGV) (Thorvaldsdottir et al., 2013), a visualization tool for analyzing large genomic data sets, with 2 of the biological replicates for each of the genotypes, wild type Arabidopsis (Columbia; Col-0), *eIFiso4G1* x *eIFiso4G2* and *eIF4G,* from the extended dark sample set only, indicates many interesting observations will come from this experiment. Most noticeably, across the entire genome, the RPKMs (Reads Per Kilobase of transcript per Million mapped reads) for the *eIFiso4G1 x eIFiso4G2* (14B) ribosome profiling (RP) reads are consistently low as compared to wild type RPKMs and typically, but not always, lower than the *eIF4G* RPKMs, in spite of normal or near normal levels of mRNA expression in 14B. This can be seen in some typical Arabidopsis reference genes (Figure 4.3.2, panel A) (Czechowski et al., 2005). Looking at genes related to the Gene Ontology (GO) term "cellular response to oxidative stress", 14B RP RPKMs are noticeably low or absent as compared to both wild type and *eIF4G* RP RPKMs (Figure 4.3.2, panels B, C and D) (Ashburner et al., 2000; Consortium, 2015). While overall 14B RPKMs are low as compared to wild type and 4G RP RPKMs, there are a few exceptions in which the 14B RP RPKMs are higher for a gene than 4G and even wild type in at least one gene, glyceraldehyde 3-phosphate dehydrogenase A subunit 2, GAPA-2, as shown in Figure 4.3.2, panel E. Investigating genes related to our unpublished proteomics data of 14B versus wild type Arabidopsis in which protein levels for 14B were low as compared to wild type Arabidopsis reveals that translation is indeed occurring at a lower rate in 14B as compared with wild type RPKMs, but not necessarily lower than 4G, for those genes (Figure 4.3.2, panel F). This is in contrast to western data that shows protein levels are consistently down in *eIFiso4G1 x eIFiso4G2* plants. Further data analysis will be needed to explain this result.

Figure 4.2 **Results of Integrative Genomic Viewer (IGV) analysis on preliminary ribosome profiling (RP) data**. For each panel, RPKMs (Reads Per Kilobase of transcript per Million mapped reads) are represented by the colored continuous bar map within 12 different tracks, purple for wild type Arabidopsis (Columbia; Col-0), dark blue for *eIFiso4G1 x eIFiso4G2* (14B) and magenta for *eIF4G* (4G). There are 2 replicates for each sample type. Samples are as follows: wild type Arabidopsis ribosome profiling samples, WT-RP; *eIFiso4G1 x eIFiso4G2* ribosome profiling samples, 14B-RP; *eIF4G* ribosome profiling samples, 4G-RP; wild type Arabidopsis RNA-seq samples, WT-mRNA; *eIFiso4G1 x eIFiso4G2* RNA-seq samples, 14B-mRNA; *eIF4G* RNA-seq samples, 4G-mRNA. Each individual gene analyzed is labeled above its representative data. A, General reference genes; B, C, and D, Genes related to the response to oxidative stress; E, genes in which 14B RPKMs are higher than 4G and are approximately the same, and in one case higher, than wild type RPKMs. F, several genes representative of our unpublished proteomics data in which 14B is low in protein expression in the genes shown. G, eIFiso4G1 and eIFiso4G2 mRNA expression levels. H, eIF4G mRNA expression levels.

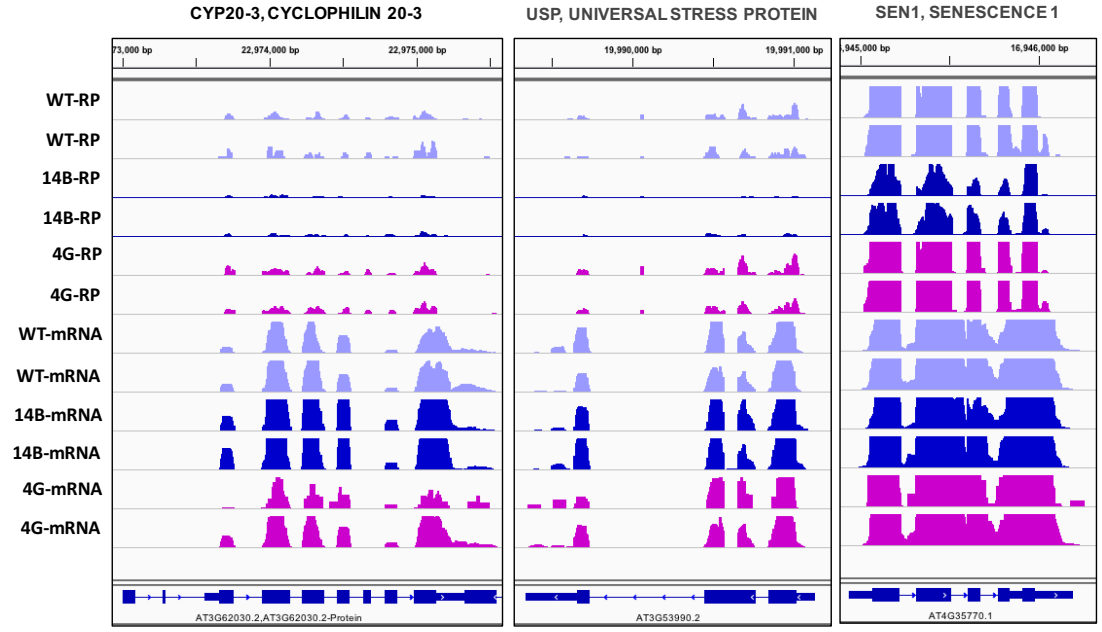


Continuation of Figure 4.2

B. Genes related to the cellular response to oxidative stress

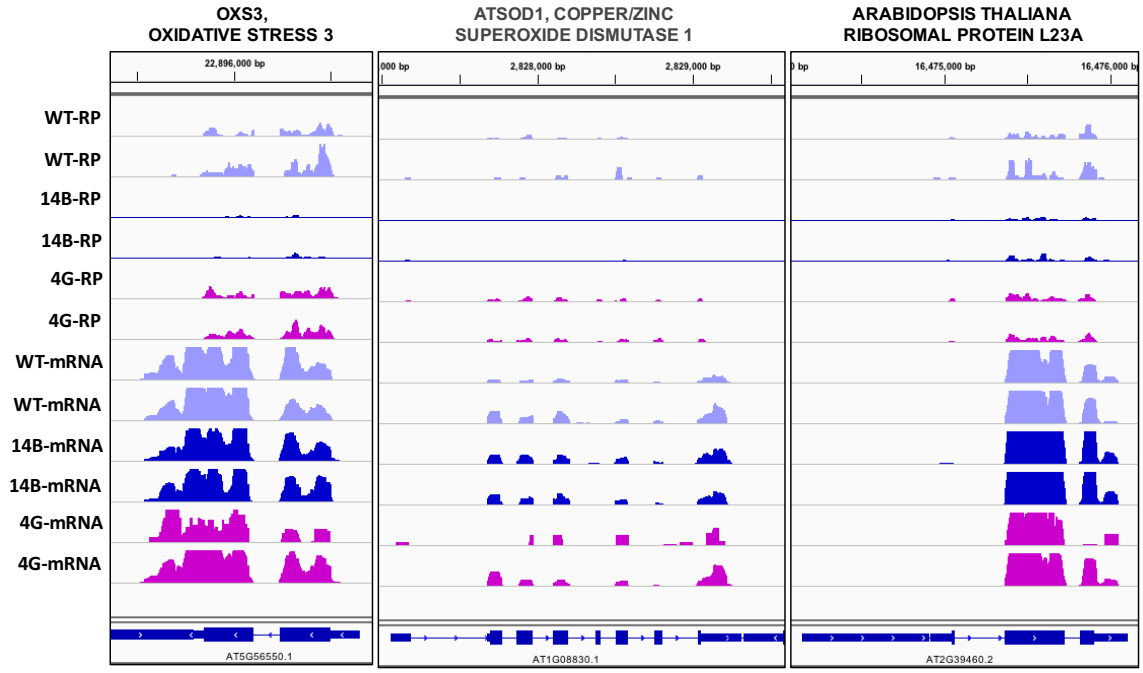


C. Genes related to the cellular response to oxidative stress

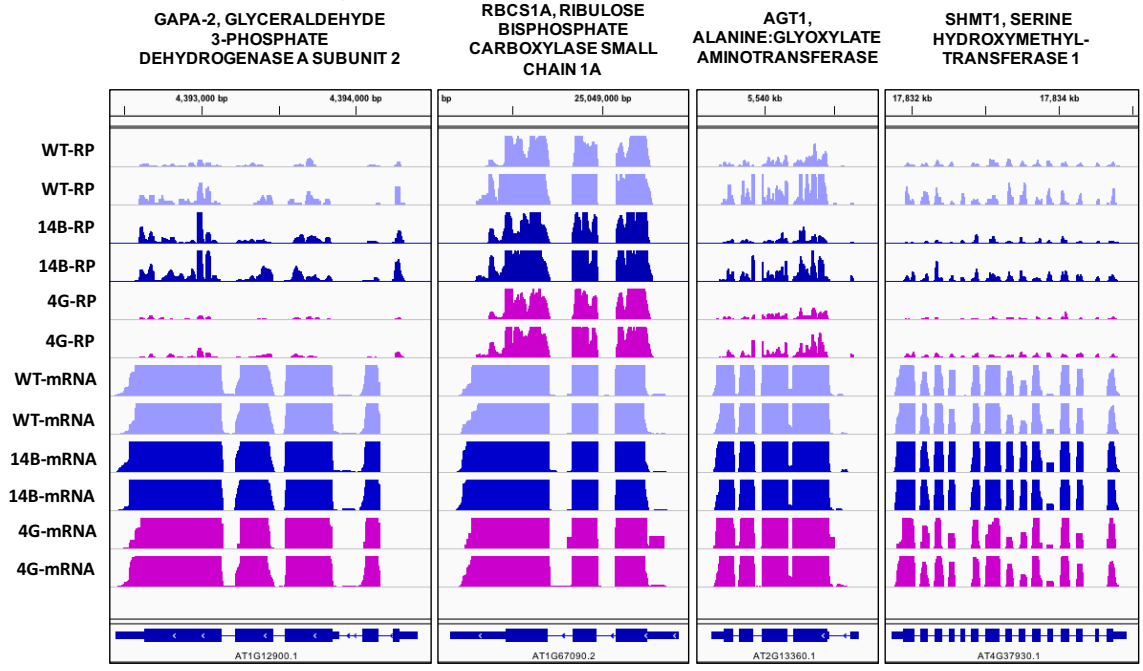


Continuation of Figure 4.2

D. Genes related to the cellular response to oxidative stress

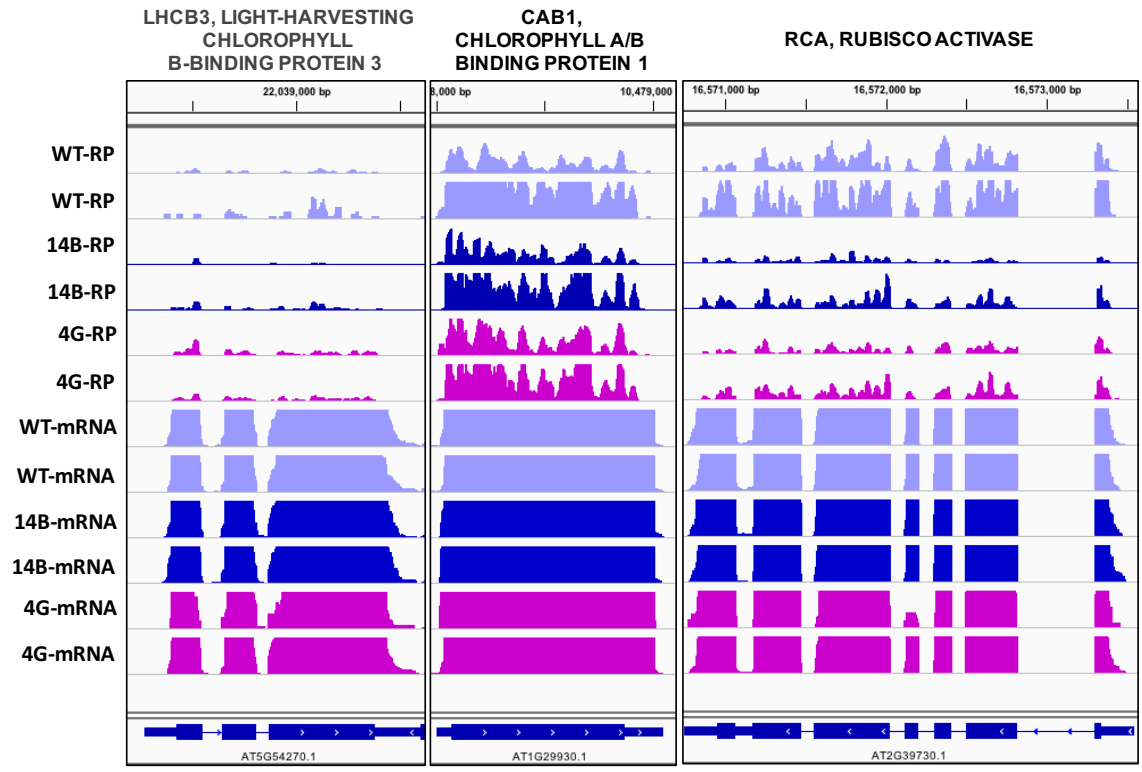


E. Genes in which 14B RPKMs are higher than 4G



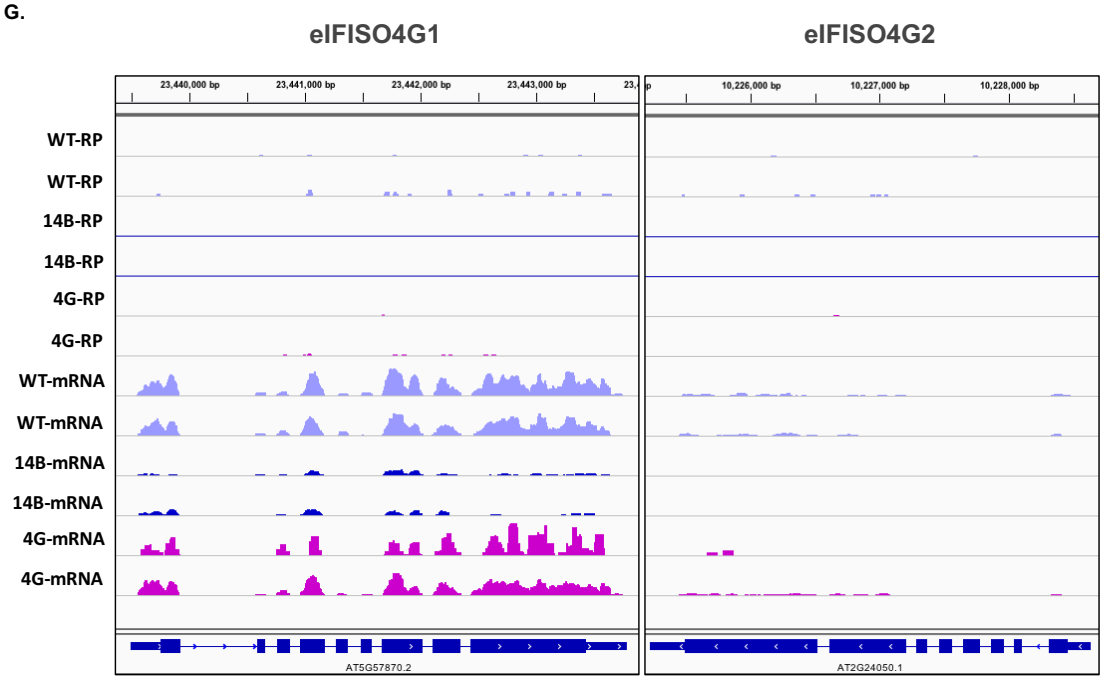
Continuation of Figure 4.2

F. Several genes representative of our proteomics data

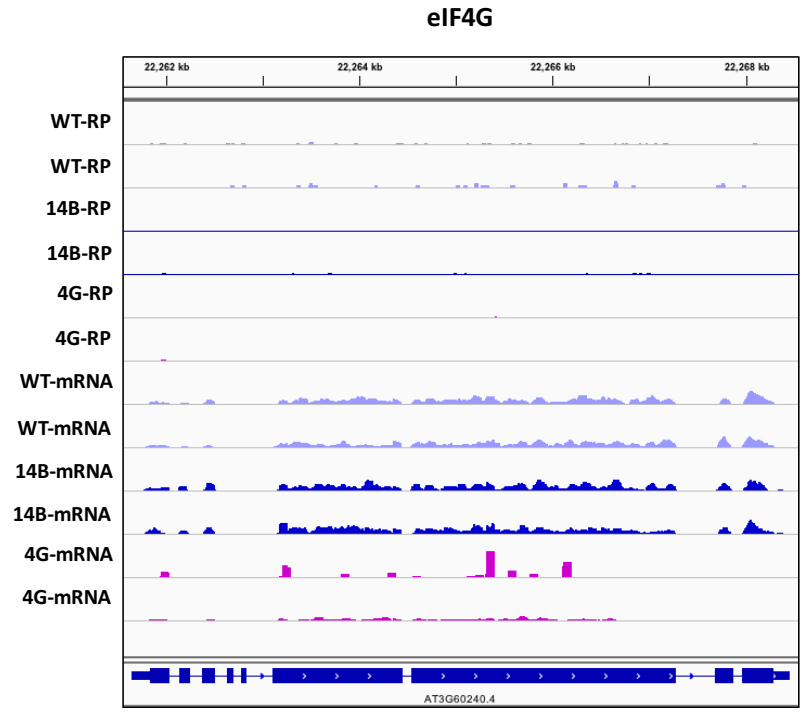


Continuation of Figure 4.2

G.



H.



Finally, looking at the mRNA expression levels in *eIFiso4G1 x eIFiso4G2* and *eIF4G*, the mRNA is either nonexistent, as seen with *eIFiso4G2* (Figure 4.3.2, panel G), or expressed in amounts that are very low, showing evidence that the gene is indeed knocked out. For *iso4G1*, the T-DNA insertion is located in exon 9 (Lellis et al., 2010). The 14B-mRNA RPKMs for that gene shows that there is a minimal amount of RPKMs in the 3' end of the gene (Figure 4.3.2, panel G). The T-DNA insertion for *eIF4G* (SALKseq_080031.0) is also located toward the 3'end of *eIF4G* where there is no expression of 4G mRNA seen (Figure 4.3.2, panel H) (Alonso et al., 2003).

Overall, based on this preliminary analysis using IGV, 14B appears to be hampered in its overall ability to efficiently carry out translation. This is in direct contrast to the mRNA levels in 14B, which are all seen to be approximately equal to wild type levels of mRNA (Figure 4.3.2, panels A – F and H). While 4G translation may also tend to be lower than wild type for many genes, it is not seen to be as consistently low as 14B at this time point.

Once the full data sets are complete, data analysis using various computational methods will be carried out (Backman & Girke, 2016). Comparisons and contrasts will be made for the different genotypes and times of harvest. Comparisons and contrasts will also be made of the mRNA levels versus the RP data. The unpublished proteomic data will also be re-examined along with a re-examination of the data obtained from the RIP-seq data. Further studies would need to be carried out to confirm and further investigate new findings. Such studies may include additional mass spectrometry of wild type versus mutant plants at the two different time points to confirm changes in protein levels and/or western analysis when antibodies are available. Our current western data and mass spectrometry from 2-D DIGE indicates several proteins are not being translated. Experiments were carried out to measure protein degradation and it did not appear to be

the cause of lower protein levels (H. Hellmann, unpublished data). Overall, very early examination of the data shows it to be a promising method to unravel the role of the plant specific translation factor, eIFiso4F and will likely raise many questions to be further investigated.

CHAPTER 5: SUMMARY

There is much yet to be learned about the regulation of plant protein synthesis (Sesma et al., 2017). In this study, we have used various methods to examine the role of the cap-binding complex in the regulation of translation in Arabidopsis. In chapter 2, we examined the role of a eIF4G kinase. The phosphorylation of proteins is a well-established regulatory mechanism used in eukaryotic regulation of translation (Sesma et al., 2017). We showed that this particular kinase is not only calcium dependent but also confirmed that the location of the CDPK site is in the C-terminus between the amino acids 511 and 624, most likely at one of two (or both) consensus sites. In chapter 3, we examined the thiol reactivity of the cysteines in eIF4E versus eIFiso4E. Our preliminary findings have revealed that there are differences between eIF4E and eIFiso4E. eIF4E does not appear to be as redox reactive as eIFiso4E and that both cysteines are not as reactive as the eIFiso4E cysteine residues. In chapter 4, I covered 2 methods we have used to determine if there are specific populations of mRNA that may associate with either eIF4F or eIFiso4F (or both), RNA immunoprecipitation and ribosome profiling. The data obtained by RNA immunoprecipitation were considered inconclusive at this point with there being much noise and no significant differences found between negative controls. However, there were experimental design errors that occurred at multiple stages of the experiment that may have compromised the results. However, this data when paired with the ribosome profiling may yet yield interesting and testable observations. The final data to be obtained from the ribosome profiling experiment that we are currently undertaking looks promising based on preliminary analyses.

Taken together, it is hoped that the work started here will lead to informative and important details about the role that the plant-specific cap-binding complex, eIFiso4F, plays in the regulation of translation.

There is a great need for understanding of the mechanisms and regulation of plant protein synthesis. Although food production has increased in the last half century, increasing demands such as a growing world population and climate change makes it imperative that we continue to find ways to protect and increase our world food supply (Godfray et al., 2010). Studying plant protein synthesis will contribute to our knowledge of how plants grow and develop (Browning & Bailey-Serres, 2015). This knowledge can lead improvements in crop health through finding new ways of preventing disease and preventing plant stress. These improvements will in turn lead to increases in biomass and quality of plant protein, which will serve to increase food production and food availability. Understanding of the mechanisms and regulation of plant protein synthesis can also lead to improved bioaccessibility of nutraceuticals through the development of foods with cellular structures that are more favorable to absorption by the gastrointestinal tract (McClements et al., 2015). Finally, an increased understanding of translation can lead to improvements in protein-based pharmaceuticals. Safer methods for recombinant pharmaceutical proteins are needed as systems using mammalian cells are fraught with problems including safety issues, such as transmittable viruses (Tekoah et al., 2015). The partial humanization of plants for the purpose of producing therapeutic proteins is already underway in Arabidopsis and in *Nicotiana benthamiana* (Strasser et al., 2004; Strasser et al., 2008). A thorough understanding of the regulation of plant protein synthesis in Arabidopsis will help to advance this technology.

APPENDIX I: SEQUENCES AND SEQUENCE ALIGNMENTS

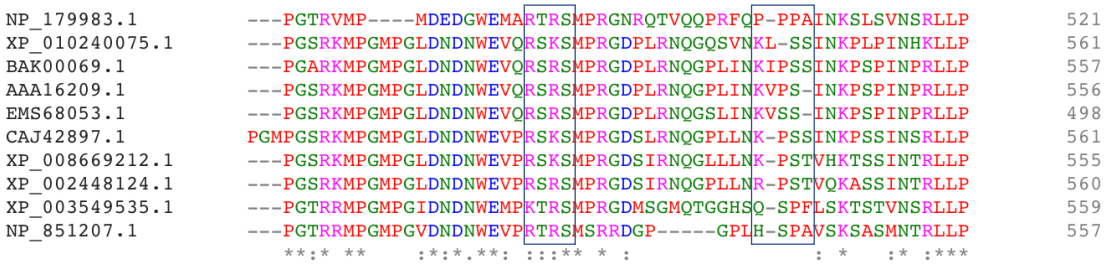
Sequence of calcium-dependent protein kinase [Triticum aestivum] cloned

Sequence ID: ABC59621.1

MQPDASGNAGGGANPRPKLPPVAAAPAPSGRPASVLPHKTANVRDHYRIGKKL
GQGQFGTTYLCVAKEDGGEFACKSIPKRKLLCREDYEDVWREIQIMHHLSEHPN
VVRIRGAYEDALFVHIVMELCAGGELFDRIVAKGHYTERAAAQLIRTIVGVVEAC
HSLGVMHRDLKPENFLFASTAEDAPLKTTDFGLSMFYKPGDKFSDVVGSPYYVA
PEVLQKCYGPEADVWSAGVILYILLCGVPPFWAETEAGIFRQILRGKLDFESEPW
PSISDSAKDLVRTMLCRDPTKRLSAHEVLCHPWIVDDAVAPDKPIDSAVLSRLKH
FSAMNKLKKMALRVIAESLSEEEIGGLRELFKMIDADNSGTITFDELKDGLKRVG
SELTEHEIQALMDAADIDNSGTIDYGEFIAATLHMNKLEREENLVSAFSFFDKDGS
GFITIDELSHACRKFGLDDVHLEDMIKDVDQNNDGQIDYSEFTAMMRKGNAGAT
GRRTMRNSLNLNLGDILNPSNS

Sequence Alignment of eIFiso4G

A portion of an alignment of 9 different plants, including both eIFiso4G1 and eIFiso4G2 from Arabidopsis, highlighting the possible CDPK phosphorylation sites. Both boxed in residues contain the conserved CDPK consensus phosphorylation site (basic-X-X-S/T) for most of the plants in the alignment with the box on the left having a higher level of conservation between the plants and most likely the site of CDPK phosphorylation. The alignment was carried out using Clustal Omega Multiple Sequence Alignment, EMBL-EBI (Sievers et al., 2011).



An * (asterisk) indicates positions in which the residue is fully conserved.
A : (colon) indicates conservation between residues that are strongly similar.
A . (period) indicates conservation between residues that are weakly similar.

NP_179983.1 MIF4G domain-containing protein / MA3 domain-containing protein [Arabidopsis thaliana] (eIFiso4G2)
XP_010240075.1 PREDICTED: eukaryotic translation initiation factor isoform 4G-1 [Brachypodium distachyon]
BAK00069.1 predicted protein [Hordeum vulgare subsp. vulgare]
AAA16209.1 initiation factor (iso)4f p82 subunit [Triticum aestivum]
EMS68053.1 Eukaryotic initiation factor iso-4F subunit p82-34 [Triticum urartu]
CAJ42897.1 putative eukaryotic translation initiation factor 4 gamma [Oryza sativa Indica Group]
XP_008669212.1 eukaryotic translation initiation factor [Zea mays]
XP_002448124.1 eukaryotic translation initiation factor [Sorghum bicolor]
XP_003549535.1 PREDICTED: eukaryotic translation initiation factor isoform 4G-1-like [Glycine max]
NP_851207.1 MIF4G domain-containing protein / MA3 domain-containing protein [Arabidopsis thaliana] (eIFiso4G1)

APPENDIX II: PROTOCOLS

Protocol 1: Cross-linking Antibody to Protein A MagBeads (Genscript) for use in RNA Immunoprecipitation

Materials:

Protein A MagBeads (Genscript)

Dimethyl pimelimidate dihydrochloride (DMP) (Sigma-Aldrich)

Antibodies

Buffers:

Binding/wash buffer (20 mM Na_2HPO_4, 0.15 M NaCl, pH 7.0)

TEA, 0.2 M, pH 8.2

PBS (0.127 M NaCl, 7 mM Na_2HPO_4-7 H_2O, 3 mM NaH_2HPO_4-H_2O)

Sodium Azide, 10%

Crosslinking Antibody to Protein A MagBeads (Genscript):

Following the instructions as given in the "Protein A MagBeads" Technical Manual No. TM0248, Version 06262010

1. Remove 240 μL resuspended beads from the Protein A MagBeads vial and place into a new tube.

The binding capacity of the beads is greater than 10 mg of rabbit IgG /mL of settled beads or 4 mLs of 25% slurry.

2. Place sample on a magnet and remove the supernatant.
3. Wash the beads with Binding/wash buffer 3 times.
4. Add IgG

   I added 0.11 mg/mL each of eIF4F and eIFiso4F antibodies for two different samples.

   For a negative control, I also prepared a set of beads having no antibodies to also use with the RIP protocol.
5. Incubate at RT for 30 – 60 min. with rotation.

   I cover the sample with foil as it rotates to protect the samples from any kind of damage from light.
6. Remove the supernatant. (Optional) Keep the supernatant in case the antibody does not bind for some reason.
7. Wash 4 times with Binding/wash Buffer.
8. Add 1 mL of 0.2 M TEA, pH 8.2.
9. Wash 2 times with TEA.
10. Resuspend the beads in 1 mL 0.2 M TEA, pH 8.2, and 5.4 mg DMP per /mL of buffer.
11. Rotate 30 min. at RT.
12. Remove the supernatant. (Optional) Keep the supernatant.
13. Resuspend beads with 1 mL 50 mM Tris, pH 7.5 to stop the reaction.
14. Rotate 15 min at RT.
15. Remove the supernatant.
16. Wash the beads 3 times with PBS, pH 7.4.

17. Store in PBS, pH 7.4, in 0.05% sodium azide (5 μL of 10% sodium azide per 1 mL of buffer).

Protocol 2: Formaldehyde Crosslinking the Tissue

Adapted from (Terzi & Simpson, 2009)

Materials:

Porcelain Büchner funnel

Vacuum flasks

Formaldehyde (FA)

Glycine, 1 M

Forceps

14-day-old seedlings

Before starting:

- Autoclave 1 beaker containing ~250 mL water (to harvest into), 297 mLs of water (this will be used for the formaldehyde solution), ~1 liter of water (for washes), spatula(s) for scraping and forceps for harvesting with, vacuum flask
- Chill all autoclaved water

Formaldehyde Crosslinking the Tissue:

1. Harvest 2-week-old seedlings into beaker of ice-cold water set in a bucket of ice.
2. Wash with ice-cold sterile water using a vacuum flask with a porcelain Büchner funnel.
3. Add 1% FA (add 3 mL FA to the 297 mL water) and place in a fresh vacuum flask.
4. Place washed seedlings inside the flask containing the FA solution and vacuum infiltrate 20 – 30 min.
5. Stop cross-linking with the addition of 1/10 volume of 1 M glycine and incubate 5 min. (add 30 mL of 1 M glycine to the 300 mL FA solution).

6. Wash fixed seedlings with cold, sterile water.
7. Air dry seedlings.
8. Grind seedlings in liquid N_2 to a fine powder.
9. Store at -80°C.

Protocol 3: RNA Immunoprecipitation

Materials:

RNase-away (Ambion)

Pierce Protease Inhibitor Mini Tablets, EDTA free (Thermo Fisher Scientific)

PMSF (200 mM)

Dounce Homogenizer

RNaseOUT (Invitrogen, 40 U/μL)

Proteinase K (Ambion, AM2546, 20 mg/mL)

DNase (Roche, DNase I, 10 U/μL)

Sodium azide, 10%

TRIzol Reagent (Invitrogen)

Glycogen (Ambion)

75% Ethanol (EtOH)

100% Isopropanol

Beads coated with antibody from protocol 1

Ground tissue from protocol 2

Buffers:

IP Buffer (50 mM HEPES, pH 7.6, 150 mM NaCl, 1 mM EDTA, 2.5 mM Mg/Cl_2,

10% glycerol, 0.1% Tween 20 [Sigma-Aldrich], RNaseOUT)

RNase-free PBS, pH 7 (0.127 M NaCl, 7 mM Na_2HPO_4-7 H_2O, 3 mM

NaH_2HPO_4-H_2O, RNase-free or DEPC-treated water)

Elution buffer (0.1 M glycine, pH 2-3)

Neutralization buffer (1 M Tris, pH 8.5)

Note: Use gloves at all times to protect your sample(s) from RNases.

Note: Keep the sample(s) as cold as possible as you move from step to step unless noted.

Before starting:

- The day before:
    - Make the IP buffer the day before without the protease inhibitors, PMSF and RNaseOUT.
    - Clean and RNase-away everything you will be using such as the dounce homogenizer, morter and pestle, spatula and the pipettes.
    - Be sure you have RNase-free pipette tips
- The day of:
    - Add protease inhibitors (1 tablet per 10 mLs of buffer), RNaseOUT (20U/mL final concentration) and PMSF (1 mM final concentration) to your IP buffer
    - Prepare the beads by washing them

RNA Immunoprecipitation (RNA-IP)

1. Homogenize required amount of cross-linked tissue (see Protocol 1 for cross-linking tissue) using a dounce homogenizer

    I used 30 grams of cross-linked tissue per set of beads (see Protocol 2).
2. Centrifuge supernatant in 50 mL conical tubes at 12,000 x g at 4˚C for 25 min.

    I used a JA14 rotor in a Beckman centrifuge.

3. Add RNaseOUT to the supernatant, saving a small aliquot to run on a gel as your input.
4. Split supernatant equally between 3 sets of washed, protein coated beads (see protocol 2 for protein coated beads).
5. Incubate the beads at 4˚C for 1 – 3 hours with rotation.

   The timing of the rotation did not appear to make a difference as to experimental outcome. The flexibility of the timing is more of a convenience factor.
6. Place on a magnet
7. Save supernatant into a new tube.

   Saving the supernatant is only for the purpose of safeguarding your sample in case the proteins did not bind to the beads for some reason.
8. Wash beads three times with RNase-free PBS, pH 7.
9. To carry out the elution, add elution buffer to the beads, incubate at RT for 5 minutes flicking the tube periodically, remove the supernatant to a fresh tube that is on ice and add neutralization buffer to the sample. Repeat all these steps 2 more times, adding each new eluate to the previous eluate.

   I added 240 µL of elution buffer each time followed by 60 µL of neutralization buffer each time. My final eluate per sample would then have a total of approximately 900 µL.
10. Add RNaseOUT to each sample.

    I used 1 µL RNase Out per 20 µL eluate.
11. De-cross-link using 5 M NaCl for a final concentration of 0.2 M.

12. Degrade any remaining proteins in the sample by adding Proteinase K to a final concentration of 0.4 mg/mL and incubating at 42˚C for 1 hour followed by 1 hour at 65˚C.
13. DNase treat the sample to remove any remaining DNA by adding to a final concentration of 10 U/50 µL

    I used approximately 19 µL DNase per tube.

    During this incubation is a good time to wash the beads you just used with PBS and store them in PBS with 0.05% sodium azide (5 µL/mL of 10% sodium azide) added to them.
14. Trizol isolation of the RNA, following the directions given with the reagent.

    In brief, 1 mL of Trizol reagent was added to the sample and kept at RT for 5 min. 200 µL of chloroform was added and the tube shaken for 15 sec. The sample was kept at RT for 3 min. and then centrifuged at 12,000 x g for 15 min. at 4˚C. The aqueous phase was removed and placed into a new tube with the addition of 2 µL of glycogen. 0.5 mL of 100% isopropanol was added and kept at RT for 10 min. The sample was centrifuged at 12,000 x g for 15 min. at 4˚C. The supernatant was removed and 1 mL of 75% EtOH was added.
15. Store the pellet at -20˚C overnight in 75% EtOH. (Optional) Or the sample can be placed at -80˚C for 1 hour and followed by the next step.

    The pellet can be stored at -80˚C at least a year in the 75% EtOH.
16. Lightly vortex then centrifuge the sample at 7,500 x g for 5 min. at 4˚C.
17. Air dry the pellet, being careful not to let it get too dry but ensuring all of the EtOH has evaporated.
18. Resuspend the RNA pellet in 20 µL of RNase free water.

19. Store your samples at -80˚C.

20. Submit for RNA seq analysis.

Protocol 4: Ribosome Profiling PCR Amplification and Barcode Addition

Component	**Amount per reaction (μL)**	**Final**
RNase-free water	4.3	
5X KAPA HiFi Buffer, GC	5.0	1X
10 μM KAPA dNTP Mix	0.75	0.3 mM each
Universal (Forward) Primer	0.75	0.3 μM
Reverse Indexed primers	0.75	0.3 μM
Template DNA	10	
1 U/μL KAPA HiFi DNA	0.5 μL	0.5 U

List of components used in the PCR reaction to amplify the circularized DNA and to add a different indexing primer to each sample (section 4.3.2.1). All reagents other than the Universal (Forward) Primer and the Reverse Indexed primers were from the KAPA HiFi PCR Kit (Kapa Biosystems). This was used in step 55 of Ingolia et al. (2012), using the Universal (Forward) Primer found in Juntawong et al. (2015) and Reverse Indexed primers found in Townsley et al., (2015).

REFERENCES

Aitken, C. E., & Lorsch, J. R. (2012). A mechanistic overview of translation initiation in eukaryotes. *Nat Struct Mol Biol, 19*(6), 568-576. doi:10.1038/nsmb.2303

Akter, S., Huang, J., Waszczak, C., Jacques, S., Gevaert, K., Van Breusegem, F., & Messens, J. (2015). Cysteines under ROS attack in plants: a proteomics view. *J Exp Bot, 66*(10), 2935-2944. doi:10.1093/jxb/erv044

Alkhalfioui, F., Renard, M., Vensel, W. H., Wong, J., Tanaka, C. K., Hurkman, W. J., . . . Montrichard, F. (2007). Thioredoxin-linked proteins are reduced during germination of Medicago truncatula seeds. *Plant Physiol, 144*(3), 1559-1579. doi:10.1104/pp.107.098103

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., . . . Ecker, J. R. (2003). Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science, 301*(5633), 653-657. doi:10.1126/science.1086391

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., . . . Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet, 25*(1), 25-29. doi:10.1038/75556

Backman, T. W. H., & Girke, T. (2016). systemPipeR: NGS workflow and report generation environment. *BMC Bioinformatics, 17*, 388. doi:10.1186/s12859-016-1241-0

Barber, J., & Andersson, B. (1992). Too much of a good thing: light can be bad for photosynthesis. *Trends Biochem Sci, 17*(2), 61-66.

Boex-Fontvieille, E., Daventure, M., Jossier, M., Zivy, M., Hodges, M., & Tcherkez, G. (2013). Photosynthetic control of Arabidopsis leaf cytoplasmic translation initiation by protein phosphorylation. *PLoS One, 8*(7), e70692. doi:10.1371/journal.pone.0070692

Boudsocq, M., Droillard, M. J., Regad, L., & Lauriere, C. (2012). Characterization of Arabidopsis calcium-dependent protein kinases: activated or not by calcium? *Biochem J, 447*(2), 291-299. doi:10.1042/bj20112072

Boudsocq, M., Willmann, M. R., McCormack, M., Lee, H., Shan, L., He, P., . . . Sheen, J. (2010). Differential innate immune signalling via Ca(2+) sensor protein kinases. *Nature, 464*(7287), 418-422. doi:10.1038/nature08794

Bouly, J. P., Gissot, L., Lessard, P., Kreis, M., & Thomas, M. (1999). Arabidopsis thaliana proteins related to the yeast SIP and SNF4 interact with AKINalpha1, an SNF1-like protein kinase. *Plant J, 18*(5), 541-550.

Brar, G. A., & Weissman, J. S. (2015). Ribosome profiling reveals the what, when, where and how of protein synthesis. *Nat Rev Mol Cell Biol, 16*(11), 651-664. doi:10.1038/nrm4069

Browning, K. S., & Bailey-Serres, J. (2015). Mechanism of cytoplasmic mRNA translation. *Arabidopsis Book, 13*, e0176. doi:10.1199/tab.0176

Browning, K. S., Lax, S. R., & Ravel, J. M. (1987). Identification of two messenger RNA cap binding proteins in wheat germ. Evidence that the 28-kDa subunit of eIF-4B and the 26-kDa subunit of eIF-4F are antigenically distinct polypeptides. *J Biol Chem, 262*(23), 11228-11232.

Browning, K. S., Webster, C., Roberts, J. K., & Ravel, J. M. (1992). Identification of an isozyme form of protein synthesis initiation factor 4F in plants. *J Biol Chem, 267*(14), 10096-10100.

Buchanan, B. B., & Balmer, Y. (2005). Redox regulation: a broadening horizon. *Annu Rev Plant Biol, 56*, 187-220. doi:10.1146/annurev.arplant.56.032604.144246

Callot, C., & Gallois, J. L. (2014). Pyramiding resistances based on translation initiation factors in Arabidopsis is impaired by male gametophyte lethality. *Plant Signal Behav, 9*(2), e27940.

Carberry, S. E., Darzynkiewicz, E., & Goss, D. J. (1991). A comparison of the binding of methylated cap analogues to wheat germ protein synthesis initiation factors 4F and (iso)4F. *Biochemistry, 30*(6), 1624-1627.

Chasse, H., Boulben, S., Costache, V., Cormier, P., & Morales, J. (2017). Analysis of translation using polysome profiling. *Nucleic Acids Res, 45*(3), e15. doi:10.1093/nar/gkw907

Cheng, S., & Gallie, D. R. (2013). Eukaryotic initiation factor 4B and the poly(A)-binding protein bind eIF4G competitively. *Translation (Austin), 1*(1), e24038. doi:10.4161/trla.24038

Cheng, S. H., Willmann, M. R., Chen, H. C., & Sheen, J. (2002). Calcium signaling through protein kinases. The Arabidopsis calcium-dependent protein kinase gene family. *Plant Physiol, 129*(2), 469-485. doi:10.1104/pp.005645

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., . . . Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol, 17*, 13. doi:10.1186/s13059-016-0881-8

Consortium. (2015). Gene Ontology Consortium: going forward. *Nucleic Acids Res, 43*(Database issue), D1049-1056. doi:10.1093/nar/gku1179

Curran, A., Chang, I. F., Chang, C. L., Garg, S., Miguel, R. M., Barron, Y. D., . . . Harper, J. F. (2011). Calcium-dependent protein kinases from Arabidopsis show substrate specificity differences in an analysis of 103 substrates. *Front Plant Sci, 2*, 36. doi:10.3389/fpls.2011.00036

Czechowski, T., Stitt, M., Altmann, T., Udvardi, M. K., & Scheible, W. R. (2005). Genome-wide identification and testing of superior reference genes for transcript normalization in Arabidopsis. *Plant Physiol, 139*(1), 5-17. doi:10.1104/pp.105.063743

de Souza, S. J., Long, M., Klein, R. J., Roy, S., Lin, S., & Gilbert, W. (1998). Toward a resolution of the introns early/late debate: only phase zero introns are correlated with the structure of ancient proteins. *Proc Natl Acad Sci U S A, 95*(9), 5094-5099.

Dennis, M. D., & Browning, K. S. (2009). Differential phosphorylation of plant translation initiation factors by Arabidopsis thaliana CK2 holoenzymes. *J Biol Chem, 284*(31), 20602-20614. doi:10.1074/jbc.M109.006692

Donnelly, N., Gorman, A. M., Gupta, S., & Samali, A. (2013). The eIF2alpha kinases: their structures and functions. *Cell Mol Life Sci, 70*(19), 3493-3511. doi:10.1007/s00018-012-1252-6

Ehsan, H., Ray, W. K., Phinney, B., Wang, X., Huber, S. C., & Clouse, S. D. (2005). Interaction of Arabidopsis BRASSINOSTEROID-INSENSITIVE 1 receptor kinase with a homolog of mammalian TGF-beta receptor interacting protein. *Plant J, 43*(2), 251-261. doi:10.1111/j.1365-313X.2005.02448.x

Emanuelle, S., Hossain, M. I., Moller, I. E., Pedersen, H. L., van de Meene, A. M., Doblin, M. S., . . . Stapleton, D. I. (2015). SnRK1 from Arabidopsis thaliana is an atypical AMPK. *Plant J, 82*(2), 183-192. doi:10.1111/tpj.12813

Gallie, D. R. (2016). Eukaryotic Initiation Factor eIFiso4G1 and eIFiso4G2 Are Isoforms Exhibiting Distinct Functional Differences in Supporting Translation in Arabidopsis. *J Biol Chem, 291*(3), 1501-1513. doi:10.1074/jbc.M115.692939

Gallie, D. R., & Browning, K. S. (2001). eIF4G functionally differs from eIFiso4G in promoting internal initiation, cap-independent translation, and translation of structured mRNAs. *J Biol Chem, 276*(40), 36951-36960. doi:10.1074/jbc.M103869200

Gao, X., Cox, K. L., Jr., & He, P. (2014). Functions of Calcium-Dependent Protein Kinases in Plant Innate Immunity. *Plants (Basel), 3*(1), 160-176. doi:10.3390/plants3010160

Godfray, H. C., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., . . . Toulmin, C. (2010). Food security: the challenge of feeding 9 billion people. *Science, 327*(5967), 812-818. doi:10.1126/science.1185383

Gupta, P., Rangan, L., Ramesh, T. V., & Gupta, M. (2016). Comparative analysis of contextual bias around the translation initiation sites in plant genomes. *J Theor Biol, 404*, 303-311. doi:10.1016/j.jtbi.2016.06.015

Hinnebusch, A. G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol, 59*, 407-450. doi:10.1146/annurev.micro.59.031805.133833

Hinnebusch, A. G., & Lorsch, J. R. (2012). The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol, 4*(10). doi:10.1101/cshperspect.a011544

Hoffman, E. A., Frey, B. L., Smith, L. M., & Auble, D. T. (2015). Formaldehyde crosslinking: a tool for the study of chromatin complexes. *J Biol Chem, 290*(44), 26404-26411. doi:10.1074/jbc.R115.651679

Holz, M. K., Ballif, B. A., Gygi, S. P., & Blenis, J. (2005). mTOR and S6K1 mediate assembly of the translation preinitiation complex through dynamic protein interchange and ordered phosphorylation events. *Cell, 123*(4), 569-580. doi:10.1016/j.cell.2005.10.024

Humphreys, J., Browning, K. S., & Ravel, J. M. (1988). Identification of a Kinase in Wheat Germ that Phosphorylates the Large Subunit of Initiation Factor 4F. *Plant Physiol, 88*(2), 483-486.

Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M., & Weissman, J. S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc, 7*(8), 1534-1550. doi:10.1038/nprot.2012.086

Jackson, R. J., Hellen, C. U., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol, 11*(2), 113-127. doi:10.1038/nrm2838

Jiang, J., & Laliberte, J. F. (2011). The genome-linked protein VPg of plant viruses-a protein with many partners. *Curr Opin Virol, 1*(5), 347-354. doi:10.1016/j.coviro.2011.09.010

Joshi, B., Lee, K., Maeder, D. L., & Jagus, R. (2005). Phylogenetic analysis of eIF4E-family members. *BMC Evol Biol, 5*, 48. doi:10.1186/1471-2148-5-48

Juntawong, P., Girke, T., Bazin, J., & Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis. *Proc Natl Acad Sci U S A, 111*(1), E203-212. doi:10.1073/pnas.1317811111

Juntawong, P., Hummel, M., Bazin, J., & Bailey-Serres, J. (2015). Ribosome profiling: a tool for quantitative evaluation of dynamics in mRNA translation. *Methods Mol Biol, 1284*, 139-173. doi:10.1007/978-1-4939-2444-8_7

Korneeva, N. L., Song, A., Gram, H., Edens, M. A., & Rhoads, R. E. (2016). Inhibition of Mitogen-activated Protein Kinase (MAPK)-interacting Kinase (MNK) Preferentially Affects Translation of mRNAs Containing Both a 5'-Terminal Cap and Hairpin. *J Biol Chem, 291*(7), 3455-3467. doi:10.1074/jbc.M115.694190

Lellis, A. D., Allen, M. L., Aertker, A. W., Tran, J. K., Hillis, D. M., Harbin, C. R., . . . Browning, K. S. (2010). Deletion of the eIFiso4G subunit of the Arabidopsis eIFiso4F translation initiation complex impairs health and viability. *Plant Mol Biol, 74*(3), 249-263. doi:10.1007/s11103-010-9670-z

Lellis, A. D., Kasschau, K. D., Whitham, S. A., & Carrington, J. C. (2002). Loss-of-susceptibility mutants of Arabidopsis thaliana reveal an essential role for eIF(iso)4E during potyvirus infection. *Curr Biol, 12*(12), 1046-1051.

Levy, Y. Y., & Dean, C. (1998). The transition to flowering. *Plant Cell, 10*(12), 1973-1990.

Li, Q., Shen, W., Zheng, Q., Tan, Y., Gao, J., Shen, J., . . . Zou, J. (2017). Effects of eIFiso4G1 mutation on seed oil biosynthesis. *Plant J*. doi:10.1111/tpj.13522

Liu, M. J., Wu, S. H., Wu, J. F., Lin, W. D., Wu, Y. C., Tsai, T. Y., . . . Wu, S. H. (2013). Translational landscape of photomorphogenic Arabidopsis. *Plant Cell, 25*(10), 3699-3710. doi:10.1105/tpc.113.114769

Maquat, L. E., & Carmichael, G. G. (2001). Quality control of mRNA function. *Cell, 104*(2), 173-176.

Marinho, H. S., Real, C., Cyrne, L., Soares, H., & Antunes, F. (2014). Hydrogen peroxide sensing, signaling and regulation of transcription factors. *Redox Biol, 2*, 535-562. doi:10.1016/j.redox.2014.02.006

Mayberry, L. K., Allen, M. L., Dennis, M. D., & Browning, K. S. (2009). Evidence for variation in the optimal translation initiation complex: plant eIF4B, eIF4F, and

eIF(iso)4F differentially promote translation of mRNAs. *Plant Physiol, 150*(4), 1844-1854. doi:10.1104/pp.109.138438

Mayberry, L. K., Allen, M. L., Nitka, K. R., Campbell, L., Murphy, P. A., & Browning, K. S. (2011). Plant cap-binding complexes eukaryotic initiation factors eIF4F and eIFISO4F: molecular specificity of subunit binding. *J Biol Chem, 286*(49), 42566-42574. doi:10.1074/jbc.M111.280099

Mayberry, L. K., Dennis, M. D., Leah Allen, M., Ruud Nitka, K., Murphy, P. A., Campbell, L., & Browning, K. S. (2007). Expression and purification of recombinant wheat translation initiation factors eIF1, eIF1A, eIF4A, eIF4B, eIF4F, eIF(iso)4F, and eIF5. *Methods Enzymol, 430*, 397-408. doi:10.1016/s0076-6879(07)30015-3

McClements, D. J., Li, F., & Xiao, H. (2015). The Nutraceutical Bioavailability Classification Scheme: Classifying Nutraceuticals According to Factors Limiting their Oral Bioavailability. *Annu Rev Food Sci Technol, 6*, 299-327. doi:10.1146/annurev-food-032814-014043

Menand, B., Desnos, T., Nussaume, L., Berger, F., Bouchez, D., Meyer, C., & Robaglia, C. (2002). Expression and disruption of the Arabidopsis TOR (target of rapamycin) gene. *Proc Natl Acad Sci U S A, 99*(9), 6422-6427. doi:10.1073/pnas.092141899

Metz, A. M., & Browning, K. S. (1996). Mutational analysis of the functional domains of the large subunit of the isozyme form of wheat initiation factor eIF4F. *J Biol Chem, 271*(49), 31033-31036.

Meyer, Y., Belin, C., Delorme-Hinoux, V., Reichheld, J. P., & Riondet, C. (2012). Thioredoxin and glutaredoxin systems in plants: molecular mechanisms, crosstalks, and functional significance. *Antioxid Redox Signal, 17*(8), 1124-1160. doi:10.1089/ars.2011.4327

Mili, S., & Steitz, J. A. (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA, 10*(11), 1692-1694. doi:10.1261/rna.7151404

Miras, M., Miller, W. A., Truniger, V., & Aranda, M. A. (2017). Non-canonical Translation in Plant RNA Viruses. *Front Plant Sci, 8*, 494. doi:10.3389/fpls.2017.00494

Mittler, R. (2017). ROS Are Good. *Trends Plant Sci, 22*(1), 11-19. doi:10.1016/j.tplants.2016.08.002

Moller, I. M., Jensen, P. E., & Hansson, A. (2007). Oxidative modifications to cellular components in plants. *Annu Rev Plant Biol, 58*, 459-481. doi:10.1146/annurev.arplant.58.032806.103946

Monzingo, A. F., Dhaliwal, S., Dutt-Chaudhuri, A., Lyon, A., Sadow, J. H., Hoffman, D. W., . . . Browning, K. S. (2007). The structure of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond. *Plant Physiol, 143*(4), 1504-1518. doi:10.1104/pp.106.093146

Muench, D. G., Zhang, C., & Dahodwala, M. (2012). Control of cytoplasmic translation in plants. *Wiley Interdiscip Rev RNA, 3*(2), 178-194. doi:10.1002/wrna.1104

Nicholson, B. L., Wu, B., Chevtchenko, I., & White, K. A. (2010). Tombusvirus recruitment of host translational machinery via the 3' UTR. *RNA, 16*(7), 1402-1419. doi:10.1261/rna.2135210

Nicholson, B. L., Zaslaver, O., Mayberry, L. K., Browning, K. S., & White, K. A. (2013). Tombusvirus Y-shaped translational enhancer forms a complex with eIF4F and can be functionally replaced by heterologous translational enhancers. *J Virol, 87*(3), 1872-1883. doi:10.1128/jvi.02711-12

Niranjanakumari, S., Lasda, E., Brazas, R., & Garcia-Blanco, M. A. (2002). Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods, 26*(2), 182-190. doi:10.1016/s1046-2023(02)00021-x

O'Brien, J. P., Mayberry, L. K., Murphy, P. A., Browning, K. S., & Brodbelt, J. S. (2013a). Evaluating the conformation and binding interface of cap-binding proteins and complexes via ultraviolet photodissociation mass spectrometry. *J Proteome Res, 12*(12), 5867-5877. doi:10.1021/pr400869u

O'Brien, J. P., Pruet, J. M., & Brodbelt, J. S. (2013b). Chromogenic chemical probe for protein structural characterization via ultraviolet photodissociation mass spectrometry. *Anal Chem, 85*(15), 7391-7397. doi:10.1021/ac401305f

Patrick, R. M., & Browning, K. S. (2012). The eIF4F and eIFiso4F Complexes of Plants: An Evolutionary Perspective. *Comp Funct Genomics, 2012*, 287814. doi:10.1155/2012/287814

Patrick, R. M., Mayberry, L. K., Choy, G., Woodard, L. E., Liu, J. S., White, A., . . . Browning, K. S. (2014). Two Arabidopsis loci encode novel eukaryotic initiation factor 4E isoforms that are functionally distinct from the conserved plant eukaryotic initiation factor 4E. *Plant Physiol, 164*(4), 1820-1830. doi:10.1104/pp.113.227785

Peled-Zahavi, H., Avital, S., & Danon, A. (2010). Methods of redox signaling by plant thioredoxins. *Methods in Redox Signaling, 34*(251-256).

Peritz, T., Zeng, F., Kannanayakal, T. J., Kilk, K., Eiriksdottir, E., Langel, U., & Eberwine, J. (2006). Immunoprecipitation of mRNA-protein complexes. *Nat Protoc, 1*(2), 577-580. doi:10.1038/nprot.2006.82

Pyronnet, S., Imataka, H., Gingras, A. C., Fukunaga, R., Hunter, T., & Sonenberg, N. (1999). Human eukaryotic translation initiation factor 4G (eIF4G) recruits mnk1 to phosphorylate eIF4E. *Embo j, 18*(1), 270-279. doi:10.1093/emboj/18.1.270

Rhee, S. Y., Beavis, W., Berardini, T. Z., Chen, G., Dixon, D., Doyle, A., . . . Zhang, P. (2003). The Arabidopsis Information Resource (TAIR): a model organism database providing a centralized, curated gateway to Arabidopsis biology, research materials and community. *Nucleic Acids Res, 31*(1), 224-228.

Rodriguez, C. M., Freire, M. A., Camilleri, C., & Robaglia, C. (1998). The Arabidopsis thaliana cDNAs coding for eIF4E and eIF(iso)4E are not functionally equivalent for yeast complementation and are differentially expressed during plant development. *Plant J, 13*(4), 465-473.

Roux, P. P., & Topisirovic, I. (2012). Regulation of mRNA translation by signaling pathways. *Cold Spring Harb Perspect Biol, 4*(11). doi:10.1101/cshperspect.a012252

Rozovsky, N., Butterworth, A. C., & Moore, M. J. (2008). Interactions between eIF4AI and its accessory factors eIF4B and eIF4H. *RNA, 14*(10), 2136-2148. doi:10.1261/rna.1049608

Schulz, P., Herde, M., & Romeis, T. (2013). Calcium-dependent protein kinases: hubs in plant stress signaling and development. *Plant Physiol, 163*(2), 523-530. doi:10.1104/pp.113.222539

Sesma, A., Castresana, C., & Castellano, M. M. (2017). Regulation of Translation by TOR, eIF4E and eIF2alpha in Plants: Current Knowledge, Challenges and Future Perspectives. *Front Plant Sci, 8*, 644. doi:10.3389/fpls.2017.00644

Shaikhin, S. M., Smailov, S. K., Lee, A. V., Kozhanov, E. V., & Iskakov, B. K. (1992). Interaction of wheat germ translation initiation factor 2 with GDP and GTP. *Biochimie, 74*(5), 447-454.

Shatkin, A. J. (1976). Capping of eucaryotic mRNAs. *Cell, 9*(4 PT 2), 645-653.

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., . . . Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol, 7*, 539. doi:10.1038/msb.2011.75

Simeunovic, A., Mair, A., Wurzinger, B., & Teige, M. (2016). Know where your clients are: subcellular localization and targets of calcium-dependent protein kinases. *J Exp Bot, 67*(13), 3855-3872. doi:10.1093/jxb/erw157

Simon, A. E., & Miller, W. A. (2013). 3' cap-independent translation enhancers of plant viruses. *Annu Rev Microbiol, 67*, 21-42. doi:10.1146/annurev-micro-092412-155609

Sims, D., Sudbery, I., Ilott, N. E., Heger, A., & Ponting, C. P. (2014). Sequencing depth and coverage: key considerations in genomic analyses. *Nat Rev Genet, 15*(2), 121-132. doi:10.1038/nrg3642

Sonenberg, N. (1994). mRNA translation: influence of the 5' and 3' untranslated regions. *Curr Opin Genet Dev, 4*(2), 310-315.

Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell, 136*(4), 731-745. doi:10.1016/j.cell.2009.01.042

Steen, H., & Jensen, O. N. (2002). Analysis of protein-nucleic acid interactions by photochemical cross-linking and mass spectrometry. *Mass Spectrom Rev, 21*(3), 163-182. doi:10.1002/mas.10024

Strasser, R., Altmann, F., Mach, L., Glossl, J., & Steinkellner, H. (2004). Generation of Arabidopsis thaliana plants with complex N-glycans lacking beta1,2-linked xylose and core alpha1,3-linked fucose. *FEBS Lett, 561*(1-3), 132-136. doi:10.1016/s0014-5793(04)00150-4

Strasser, R., Stadlmann, J., Schahs, M., Stiegler, G., Quendler, H., Mach, L., . . . Steinkellner, H. (2008). Generation of glyco-engineered Nicotiana benthamiana for the production of monoclonal antibodies with a homogeneous human-like N-glycan structure. *Plant Biotechnol J, 6*(4), 392-402. doi:10.1111/j.1467-7652.2008.00330.x

Tekoah, Y., Shulman, A., Kizhner, T., Ruderfer, I., Fux, L., Nataf, Y., . . . Shaaltiel, Y. (2015). Large-scale production of pharmaceutical proteins in plant cell culture-the Protalix experience. *Plant Biotechnol J, 13*(8), 1199-1208. doi:10.1111/pbi.12428

Terzi, L. C., & Simpson, G. G. (2009). Arabidopsis RNA immunoprecipitation. *Plant J, 59*(1), 163-168. doi:10.1111/j.1365-313X.2009.03859.x

Thorvaldsdottir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform, 14*(2), 178-192. doi:10.1093/bib/bbs017

Townsley, B. T., Covington, M. F., Ichihashi, Y., Zumstein, K., & Sinha, N. R. (2015). BrAD-seq: Breath Adapter Directional sequencing: a streamlined, ultra-simple and fast library preparation protocol for strand specific mRNA library construction. *Front Plant Sci, 6*, 366. doi:10.3389/fpls.2015.00366

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., . . . Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc, 7*(3), 562-578. doi:10.1038/nprot.2012.016

Treder, K., Kneller, E. L., Allen, E. M., Wang, Z., Browning, K. S., & Miller, W. A. (2008). The 3' cap-independent translation element of Barley yellow dwarf virus binds eIF4F via the eIF4G subunit to initiate translation. *RNA, 14*(1), 134-147. doi:10.1261/rna.777308

Truniger, V., & Aranda, M. A. (2009). Recessive resistance to plant viruses. *Adv Virus Res, 75*, 119-159. doi:10.1016/s0065-3527(09)07504-6

Vasicek, L., O'Brien, J. P., Browning, K. S., Tao, Z., Liu, H. W., & Brodbelt, J. S. (2012). Mapping protein surface accessibility via an electron transfer dissociation selectively cleavable hydrazone probe. *Mol Cell Proteomics, 11*(7), O111.015826. doi:10.1074/mcp.O111.015826

Wall, S. B., Oh, J. Y., Diers, A. R., & Landar, A. (2012). Oxidative modification of proteins: an emerging mechanism of cell signaling. *Front Physiol, 3*, 369. doi:10.3389/fphys.2012.00369

Wang, A., & Krishnaswamy, S. (2012). Eukaryotic translation initiation factor 4E-mediated recessive resistance to plant viruses and its utility in crop improvement. *Mol Plant Pathol, 13*(7), 795-803. doi:10.1111/j.1364-3703.2012.00791.x

Wang, L., Si, Y., Dedow, L. K., Shao, Y., Liu, P., & Brutnell, T. P. (2011). A low-cost library construction protocol and data analysis pipeline for Illumina-based strand-specific multiplex RNA-seq. *PLoS One, 6*(10), e26426. doi:10.1371/journal.pone.0026426

Wang, Z., Gerstein, M., & Snyder, M. (2009a). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet, 10*(1), 57-63. doi:10.1038/nrg2484

Wang, Z., Treder, K., & Miller, W. A. (2009b). Structure of a viral cap-independent translation element that functions via high affinity binding to the eIF4E subunit of eIF4F. *J Biol Chem, 284*(21), 14189-14202. doi:10.1074/jbc.M808841200

Xia, X. J., Zhou, Y. H., Shi, K., Zhou, J., Foyer, C. H., & Yu, J. Q. (2015). Interplay between reactive oxygen species and hormones in the control of plant development and stress tolerance. *J Exp Bot, 66*(10), 2839-2856. doi:10.1093/jxb/erv089

Yamazaki, D., Motohashi, K., Kasama, T., Hara, Y., & Hisabori, T. (2004). Target proteins of the cytosolic thioredoxins in Arabidopsis thaliana. *Plant Cell Physiol, 45*(1), 18-27.

Zhang, X. S., & Choi, J. H. (2001). Molecular evolution of calmodulin-like domain protein kinases (CDPKs) in plants and protists. *J Mol Evol, 53*(3), 214-224. doi:10.1007/s002390010211

Zhang, Y., Dickinson, J. R., Paul, M. J., & Halford, N. G. (2003). Molecular cloning of an arabidopsis homologue of GCN2, a protein kinase involved in co-ordinated response to amino acid starvation. *Planta, 217*(4), 668-675. doi:10.1007/s00425-003-1025-4

Zhang, Z., Boonen, K., Ferrari, P., Schoofs, L., Janssens, E., van Noort, V., . . . Geuten, K. (2016). UV crosslinked mRNA-binding proteins captured from leaf mesophyll protoplasts. *Plant Methods, 12*, 42. doi:10.1186/s13007-016-0142-6

Zielinski, J., Kilk, K., Peritz, T., Kannanayakal, T., Miyashiro, K. Y., Eiriksdottir, E., . . . Eberwine, J. (2006). In vivo identification of ribonucleoprotein-RNA interactions. *Proc Natl Acad Sci U S A, 103*(5), 1557-1562. doi:10.1073/pnas.0510611103