5-2023

Characterization of Arabidopsis Eukaryotic Translation Initiation Factor 2α (eIF2α) Mutants

Mark Edens
medens@vols.utk.edu

To the Graduate Council:

I am submitting herewith a thesis written by Mark Edens entitled "Characterization of Arabidopsis Eukaryotic Translation Initiation Factor 2α (eIF2α) Mutants." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Biochemistry and Cellular and Molecular Biology.

<div style="text-align: right;">Albrecht G. von Arnim, Major Professor</div>

We have read this thesis and recommend its acceptance:

Brad M. Binder, Barry D. Bruce

<div style="text-align: right;">Accepted for the Council:
Dixie L. Thompson</div>

<div style="text-align: right;">Vice Provost and Dean of the Graduate School</div>

(Original signatures are on file with official student records.)

Characterization of Arabidopsis *Eukaryotic Translation Initiation Factor 2α* (*eIF2α*) Mutants

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Mark Edens

May 2023

Acknowledgements

I would like to thank my advisor, Dr. Albrecht G. von Arnim for his guidance, mentorship and direction throughout this project. I would also like to specially thank him for his instruction in experimental design, his ability to turn a negative result into an impactful insight and with the ongoing process of learning in plant biology.

I thank my lab members Ricardo Urquidi Camacho and Anwesha Dasgupta. Thanks for being a pleasure to work with and excellent colleagues. Special thanks to former lab members Dr. Ansul Lokdarshi and Jeremiah Holt for their assistance and contributions to my project and related projects in the lab. I also thank my committee members Dr. Barry D. Bruce and Dr. Brad M. Binder for their insights and guidance of my project. I would like to make a special acknowledgement Dr. Tessa M. Burch-Smith who sparked my interest in plant science research and inspired me to pursue this career.

I thank my parents, Mark Edens Sr. and Jo Ann Edens and my brothers Thomas Mullins and Johnathan Muncy for their support throughout my graduate school experience.

Abstract

Plants are stationary organisms that are charged with overcoming a multitude of biotic and abiotic stresses. The eukaryotic translation initiation factor 2 (eIF2) is responsible for charging the P-site of the forming 80S ribosome with the initiator methionyl-tRNA. Thus, eIF2 is a protein of utmost consequence in the growth and development of organisms. The eIF2 protein is also a mediator of global translational regulation in the eukaryotic integrated stress response (ISR), where the α-subunit is phosphorylated by protein kinases, such as the kinase GCN2. Here, five different mutant *eIF2α* alleles were examined in Arabidopsis but only one was found to be deleterious to plant growth, development, and reproduction. Additionally, *eIF2α* RNA and protein were expressed at variable levels in *eIF2α* mutant seedlings and maturing rosette-stage plants. Plants overexpressing *eIF2α* wild-type and phospho-variant alleles had altered root lengths and displayed variable growth defects at the rosette and reproductive stages. Finally, recombinant forms of eIF2α, an alleged partner protein ABCF1, and the GCN2-kinase domain were expressed and purified from an *E. coli* cell line. These data demonstrate that the *eIF2α* gene plays a vital role in the growth and development of seedlings and mature Arabidopsis plants and overexpression of eIF2α disrupts the native development of plant structures.

Table of Contents

List of Tables

List of Figures

Chapter 1.
Introduction

Plants are sessile organisms that exist in ever-changing, dynamic environments and exhibit a wide range of stress responses to combat environmental challenges. Translation lies at the core of the central dogma of biology and the modulation of translation is of critical importance for plant survival in their environment. There are several ways that organisms can modify the rate and targets of translation. Inhibition of TOR results in reduced translation in response to energy deprivation. In plants, hypoxia triggers sequestering of translation factors in stress granules which results in inhibition of translation. In animals and yeast, many different stress triggers disrupt the function of the eukaryotic initiation factor 2 (eIF2) via phospho-regulation of its α-subunit, causing P-eIF2α to act as a competitive inhibitor of its own GEF, eIF2B. With eIF2B inhibited, the eIF2 complex is rendered inactive and cytosolic translation is repressed. The eIF2-eIF2B complex is a natural target for a translation modulation pathway as the abundance of unphosphorylated eIF2 is rate limiting for translation.

1.1 Translation Initiation

In eukaryotes, translation initiation is highly regulated by the coordination of several eukaryotic translation initiation factors. The initiation factors serve to recognize mature mRNA, recruit proteins and their co-factors to the 40S ribosome to form the 48S pre-initiation complex, and ultimately recruit the 60S subunit to form the 80S ribosome for elongation to occur.

Translation initiation begins with the recognition of the 5' m^7G cap of mRNA by eIF4F, a complex of the eIF4G scaffold and eIF4E cap-binding protein. The remaining eIF4 subunits, eIF4A dead-box helicase and its stimulating co-factor eIF4B bind the eIF4G scaffold and unwind the 5' end of the mRNA. The eIF2 protein is a trimer with a gamma subunit that binds and hydrolyzes GTP, a beta subunit that stabilizes bound initiator methionyl-tRNA (Met-tRNA^i), and an alpha subunit that can be phosphorylated by one of several *eIF2α* kinases [1-3]. The trimeric eIF2 protein, a GTP molecule and the Met-tRNA^i together form the ternary complex, which charges the P-site of the ribosome with Met-tRNA^i. Before the eIF4 complex associates with the preinitiation complex, the ternary complex binds the 40S, positioning the
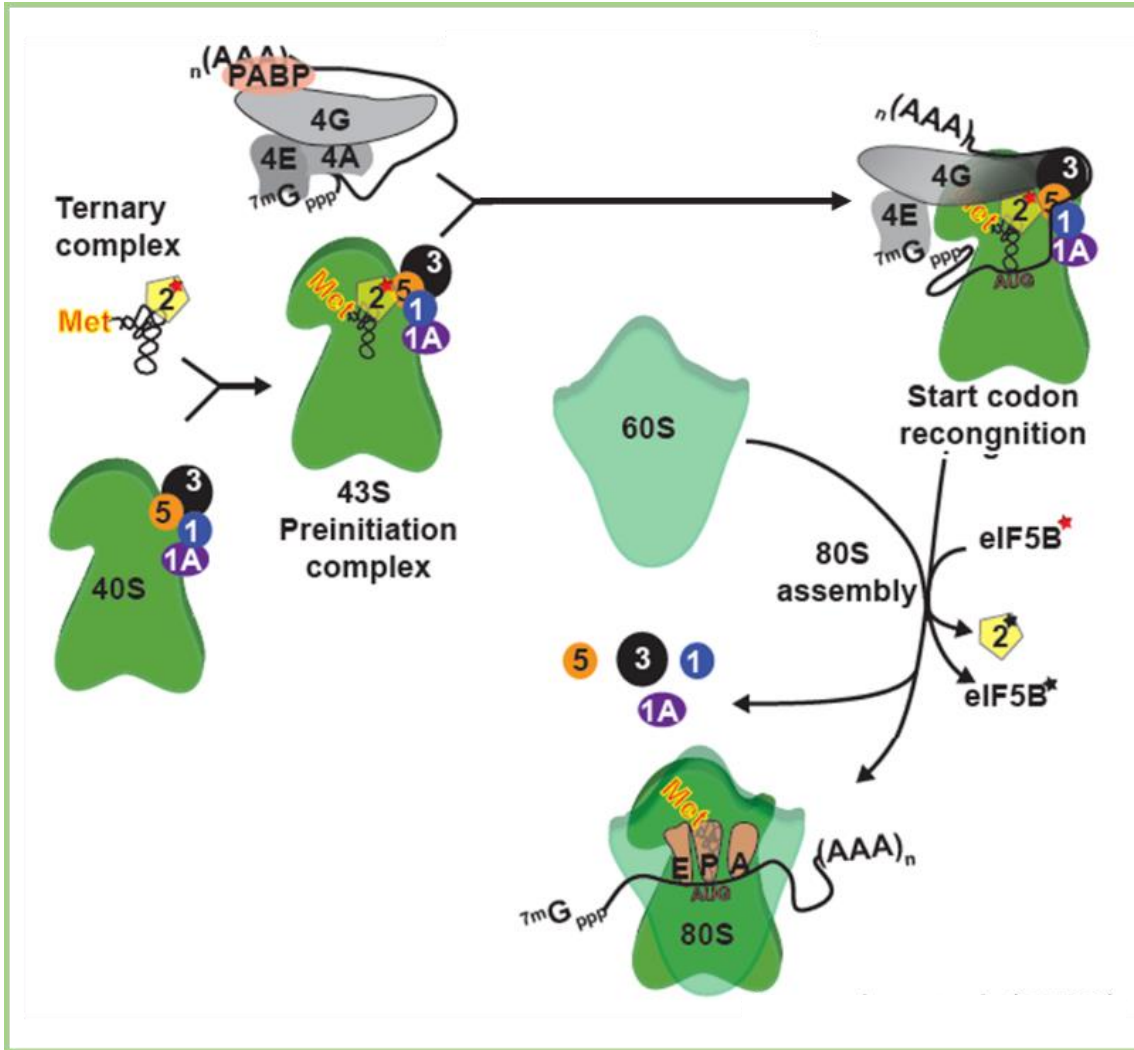
Figure 1. General Overview of Translation Initiation.

This figure is an adaptation of Figure 1 from Urquidi Camacho and Lokdarshi (2020) *"Translational gene regulation in plants: A green new deal*. Ribosomal subunits are represented as large, green triangular shapes, with sedimentation coefficients. Initiation factors are represented as colored shapes with the initiation factor number. (ex. 4G is eIF4G; 2 is eIF2). GTP is represented as a red star. GDP is represented as a blue star. Initiator methionine is represented as "Met" attached to the characteristic 3-hairpin loop tRNA fold. The mRNA strand is represented as a black ribbon with a 5' cap and 3'poly-A tail.

tRNA in the P-site of the small subunit [1, 2]. Once the ribosome has scanned down the mRNA and recognized a start codon, the associated eIF5 protein stimulates the hydrolysis of GTP on eIF2γ [2].

The eIF5 initiation factor is associated with the preinitiation complex-bound eIF3 scaffold which assembles eIF1, eIF1A and eIF5 to form the 43S preinitiation complex [2, 4]. The 43S-associated eIF3 scaffold then binds the mRNA-bound eIF4F scaffold to form the 48S preinitiation complex.

The 48S preinitiation complex scans the mRNA from 5' to 3' in search of an AUG start codon that initiates a protein coding region (ORF). Start codon recognition is tightly regulated in eukaryotes. It is estimated that 30% to 40% of plant genes contain upstream open reading frames (uORFs), short coding regions that start within the 5'UTR. The recognition of uORFs typically inhibits protein synthesis by causing premature collapse of the translating ribosome or the out-of-frame synthesis of a peptide destined for degradation [2, 5]. Environmental conditions heavily contribute to regulation of gene expression by uORFs and can result in differentially expressed genes in response to stress. When an AUG has been recognized by binding the eIF2-associated Met-tRNAi, the GTP on eIF2γ is hydrolyzed, stimulated by eIF5 [2]. The translation initiation factors dissociate from the small ribosomal subunit and eIF5B mediates the binding of the 60S subunit— concluding the initiation phase and releasing the remaining initiation factors to form the 80S ribosome [2].

1.2 Translation initiation factor eIF2

There is a very high sequence similarity between species for the subunits of eIF2 which underscores the importance of the complex. The presence and role of eIF2 is ubiquitous throughout eukaryotes since it is rate limiting for translation and is the target of a powerful method of translation regulation by kinase activity. Mammals and yeast have one copy of *eIF2α* in each respective genome and as a result *eIF2α* mutants are lethal [6]. In plants, there are two copies of *eIF2α* on chromosome 2 and 5. The eIF2 initiation factor is a heterotrimer that performs the essential role of charging the P-site of the ribosome with the Met-tRNAi. The complex consists of eIF2α and eIF2β which both associate with the eIF2γ subunit. In eukaryotes, the eIF2γ subunit is responsible for binding GTP and the Met-tRNAi. The eIF2β subunit

3

contains a binding domain for eIF5 and eIF2B, the GEF that recycles GDP following P-site loading of the Met-tRNAi [1, 4]. The α-subunit of eIF2 contains the phospho-serine in Eukaryotes, which is serine 56 in Arabidopsis. The phospho-serine is responsible for translational control via the ISR when phosphorylated by its sole plant kinase general control non-derepressible 2 (GCN2).

1.3 Translation modulation and eIF2α

In humans and yeast, translation modulation in response to inhibition of eIF2B by P-eIF2α is known as the integrated stress response (ISR) and general amino acid control pathway (GAAC), respectively. The ISR is a conserved signaling pathway by which at least some eukaryotes adapt and respond to environmental stress by modulating translation via phospho-regulation of eIF2α protein. When the phosphoserine is targeted by an eIF2α kinase, there is a repression of global protein synthesis by inhibition of a functional ternary complex. In animals and fungi, the phosphorylation of P-eIF2α results in the inability of eIF2B to perform its GEF function. The decameric GEF eIF2B binds eIF2 on all three subunits according to CRYO-EM and crosslinking studies in budding yeast and humans, but binding profiles are different between studies and the exact mechanism of binding is still under debate [7-9]. P-eIF2α inhibits eIF2B function by increasing the binding affinity of GDP to eIF2 [1]. The consequence of eIF2B inhibition is the lack of GTP-charged ternary complexes which leads to translational repression.

In eukaryotes, there are four eIF2α kinases that are activated in response to stress: PKR is activated in response to viral infection, PERK in response to ER stress, HRI in response to heme deprivation and heat shock, and GCN2 in response to amino acid deprivation and a unique variety of other stressors in plants. Other eukaryotes such as humans and mice possess all 4 eIF2α kinases. Plants and budding yeast have only one eIF2α kinase— GCN2. In Arabidopsis, GCN2 is activated by UV irradiation, nutrient deprivation, herbicides, high light conditions, pathogenic infection, oxidative and cold stress [10-14].
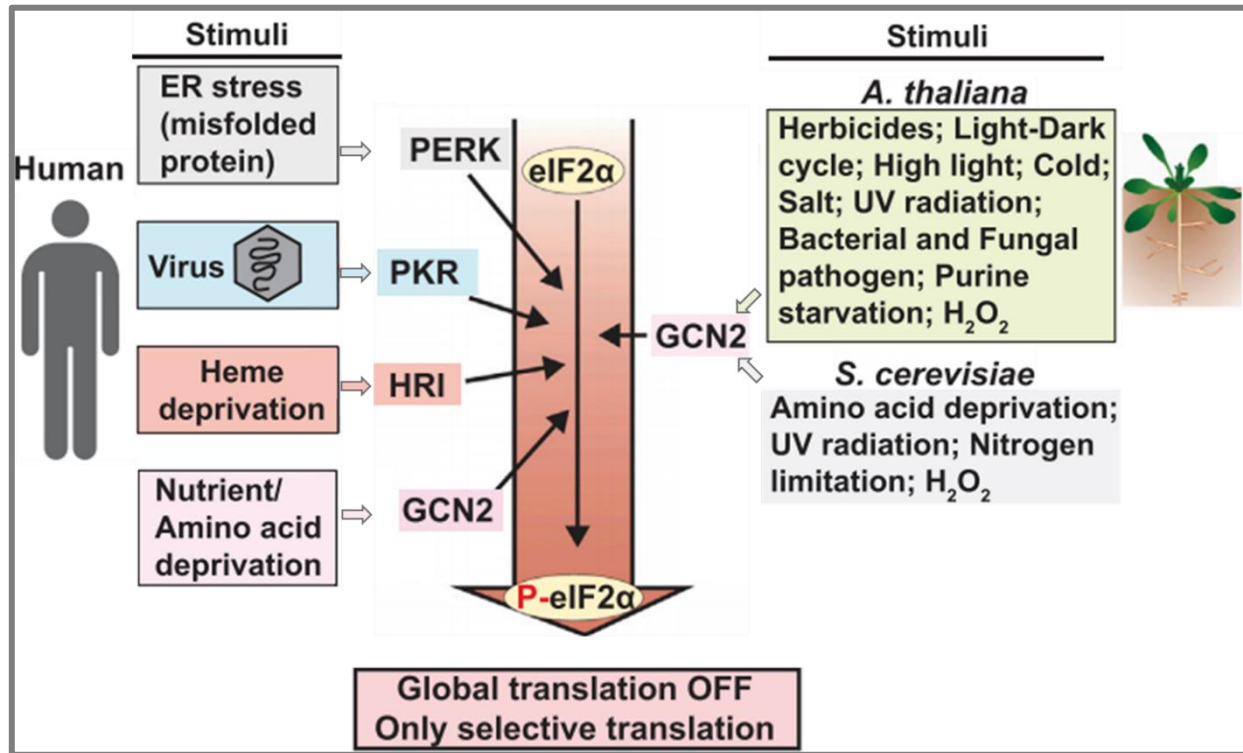
Figure. 2 – Diagram Outlining the ISR in Mammals, Plants and Budding Yeast.

This figure is an adaptation of Figure 3 from Lokdarshi and von Arnim (2022) *"Review: Emerging roles of the signaling network of the protein kinase GCN2 in the plant stress response"* modeling the integrated stress response pathway involving eIF2-kinases in mammals, plants, and budding yeast [15]. There are 4 kinases for eIF2α in humans, however, in budding yeast and plants, GCN2 is the sole kinase for the eIF2α protein.

In budding yeast and humans, select survival-response mRNAs escape translational repression by the ISR. For example, GCN4 is a transcription factor that targets amino acid biosynthesis genes in yeast. Its mRNA avoids the translational repression by uORF regulation during amino acid starvation [5, 16]. An experiment surveying differentially translated genes in response to herbicide treatment demonstrates that there are no GCN4-like transcripts that escape translational repression following GCN2 activation in Arabidopsis [10].

1.3.1 Cofactors for GCN2 kinase: GCN1 and GCN20

GCN1 is required for the phosphorylation of eIF2α in the mammalian and budding yeast models [17]. The C-terminal region of GCN1 contains a GCN2 binding domain and several ribosomal binding sites spanning the length of the protein [18, 19]. When *S. cerevisiae* is starved of amino acids, elongating ribosomes will stall when a deacylated tRNA is incorporated into the A-site, which results in ribosome collisions [20]. It has been demonstrated that GCN1 can recognize and bind the colliding disome [20]. Therefore, GCN1 may acta as a sensor for amino acid starvation and serve as a recruiting scaffold for GCN2. In plants, the precise role of GCN1 is not as clear, but GCN1 is required for GCN2 kinase activity [21, 22]. It plays a role not only in the ISR, but also in stomatal aperture modulation to combat pathogens, chloroplast biogenesis and in embryogenesis [23, 24].

GCN20, belonging to the soluble ATP binding cassette protein-F (ABCF) family of proteins, is a co-factor for GCN1 function in the ISR. In budding yeast and humans, *gcn20* [-] mutants have a reduced, but not abolished, eIF2α phosphorylation by GCN2 [25]. In Arabidopsis, there are 5 ABCF genes with close sequence homology to *S. cerevisiae* GCN20, yet none are individually required for GCN2 activity in the ISR. ABCF3 has been recognized to be homologous to *S. cerevisiae* GCN20 because it is functionally linked to GCN1 in chloroplast biogenesis, stomatal closure, as well as mirroring transcriptomic profiles when assaying *abcf3* and *gcn1* mutants respectively [26]. However, in this thesis, *ABCF1* is cloned and purified due to a previously misannotated entry in TAIR which identified AT5G60790 (*ABCF1*) as the Arabidopsis GCN1 homologue.

1.3.2 GCN2

In mammals and yeast, GCN2 is comprised of five functional domains which are an N-terminal GCN1/GCN20 co-factor binding domain, a pseudokinase Ψ domain, an eIF2α kinase domain, Histidyl-tRNA synthetase-like (His-RS) regulatory region, and a C-terminal ribosome binding domain [19, 27]. The N-terminal region binds GCN2 co-factors GCN1 and GCN20, also known as ILITHYIA (AtGCN1) and ABCF3 (AtGCN20) respectively in Arabidopsis [21, 22]. The eIF2α kinase domain is responsible for the phosphorylation of eIF2α on the serine 52 (56 in Arabidopsis) and the pseudokinase Ψ domain contains conserved residues that bolster the kinase function of the eIF2α kinase domain [19]. The His-RS domain binds uncharged t-RNA, which is thought to induce a conformational change that stimulates kinase activity [28-30]. Although it has been demonstrated that in addition to deacylated tRNAs, free purified ribosomes and the P-stalk region of the ribosome alone can activate GCN2 *in-vitro* [31].

In comparison with the budding yeast model, Arabidopsis GCN2 shares homology with the GCN1/GCN20 binding domain, the eIF2α kinase domain and the His-RS domain. However, Arabidopsis GCN2 has a large, undefined region where the budding yeast, human and drosophila pseudokinase Ψ domain exists. Arabidopsis GCN2 lacks the conserved residue motif for pseudokinase function.

In plants, it has been demonstrated that ROS is a signaling molecule of great importance to GCN2 activity. Chlorosulfuron (CSF) is an herbicide that inhibits acetolactate synthase (ALS), an enzyme necessary for branched-chain amino acid (BCAA) biogenesis. When Arabidopsis plants are treated with CSF in light, eIF2α is phosphorylated and a buildup of ROS can be measured by DAB staining and H2DCFDA staining. In the dark, eIF2α phosphorylation is absent during CSF treatment. However, when plants are treated with hydrogen peroxide directly in the dark, there is a robust phosphorylation of eIF2α, and this effect can be abrogated by treatment with the photosynthetic inhibitors DCMU and DBMIB. Taken together, these data demonstrate that GCN2 activity is regulated by ROS from the chloroplast during stress conditions [10].

GCN2 has also been implicated in plant defense against *P. syringae*. The presence of *P. syringae* results in the phosphorylation of eIF2α via GCN2 kinase activity, which is abrogated in *gcn2* Arabidopsis

mutants. In Arabidopsis, the heat shock-like transcription factor TBF1, which is responsible for modulating transcription of pathogen survival transcripts is regulated by uORFs in similar fashion to ATF4 and GCN4. TBF1 is more highly translated after one hour following *P. syringae* inoculation, however the increased translation rate depends on GCN2 mutants. Abscisic acid (ABA) participates in a sensitive signaling pathway with other plant hormones, including salicylic acid (SA) and jasmonic acid (JA), to manage pathogenic defense. *P. syringae* uses the type-III secretion system and coronatine, an active JA analog, to benefit from ABA-pathway crosstalk by inducing an overaccumulation of ABA. In addition, to the TBF1 data, it has been demonstrated that ABA levels can influence plant immunity against Pseudomonas at various stages of infection in a GCN2-dependent manner. Wild-type plants are more susceptible to Pseudomonas infection than the *gcn2* mutants which show lower levels of ABA accumulation during infection [14, 15].

Additionally, the eIF2α-GCN2 paradigm has been implicated in a hypoxia stress response in conjunction with the ethylene signaling pathway, which reduces polysome loading on mRNAs [13]. It was demonstrated that eIF2α is phosphorylated following the treatment of wild-type plants with aminocyclopropane-1-carboxylic acid (ACC), an ethylene precursor [32]. Under conditions of hypoxia, induced by the submergence of Arabidopsis seedlings, there is a collapse of polysomes in wild-type plants, which was reduced in *gcn2* mutant plants and nearly abrogated in *ein2* and *etr1* mutants. This result shows that GCN2 is activated through the ethylene signaling pathway in response to hypoxia stress [13].
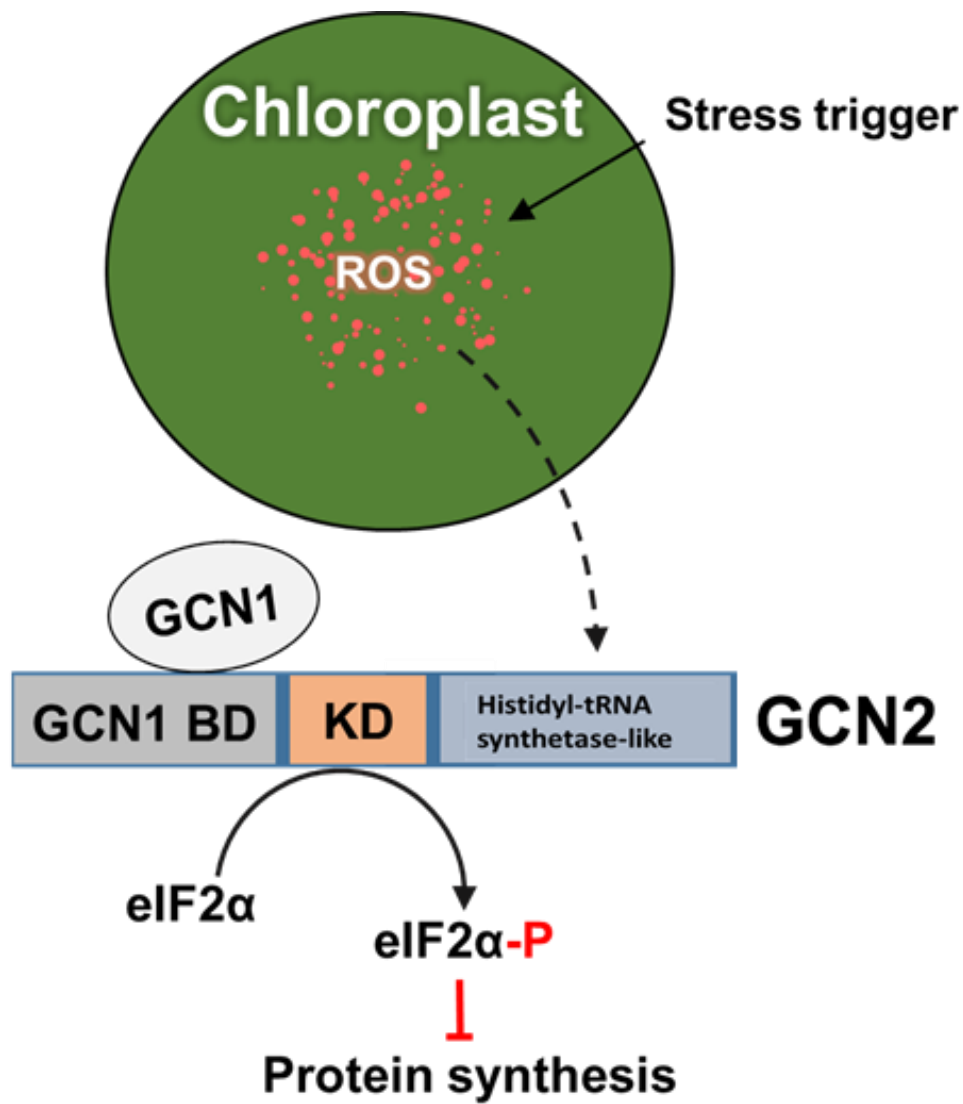
Figure. 3 – Working model for GCN2 Signal Transduction Pathway.

A stress trigger results in ROS buildup in the chloroplast and subsequent escape from the chloroplast. ROS directly

or indirectly activates GCN2 and phosphorylates eIF2α. This leads to a repression in global protein synthesis.

1.4 Addressed in this work

This project is focused on the characterization of Arabidopsis *eIF2α* allele mutants, a molecular biology approach to create a transgenic plant pipeline for dissecting the role of *eIF2α* and its phosphorylation using biochemistry and reverse genetics, First, Arabidopsis *eIF2α* mutant seedlings from ABRC were germinated on MS media, transferred to soil, and genotyped by PCR to achieve stable homozygous mutant seed stocks. The seedlings and mature plants were assayed for seedling root growth, rosette area, eIF2α protein abundance by western blot, and eIF2α mRNA transcript abundance was measured by RT-qPCR. Transgenes coding for *eIF2α* wild-type, phospho-mimetic, and phospho-null alleles were transformed into Arabidopsis by *A. tumefaciens* floral dips, genotyped, and growth defects catalogued. The project also included expression and purification of recombinant eIF2α, GCN2 and ABCF1 protein from *E. coli.* The recombinant *GCN2-kinase domain (GCN2-KD)*, *ABCF1* and *eIF2α* wild-type, phospho-mimetic (S56D), and phospho-null (S56A) alleles were cloned into protein expression vectors in a BL21STAR cell line, expressed and purified. This work serves as a contribution to the greater Integrated Stress Response project in the von Arnim lab.

Chapter 2.
Materials and Methods

2.1. Arabidopsis eIF2α allele mutants

Seed stocks for *eIF2α* allele mutants were ordered from ABRC. The seed stock IDs for AT2G40290 [*eIF2α(2g)*] are as follows: SALK 0936717 [*eif2α(2g)-1*] and SALK 025426C [*eif2α(2g)-2*].The seed stock IDs for AT5G05470 [*eIF2α(5g)*] are as follows: CS822591 [*eif2α(5g)-1*], SALK 065838 [*eif2α(5g)-2*], and SALK 083377 [*eif2α(5g)-3*]. One should note that the SALK 025426C *eIF2α(2g)-2* and SALK 065838 *eif2α(5g)-2* lines in this thesis are identical to the *eif2α(2g)* and *eif2α(5g)* in the Cho, H.Y. et al (2022) referenced in this thesis.

2.2. Allele mutant PCR genotyping

Homozygous *eIF2α* allele mutants were identified by analytical polymerase chain reaction (PCR) using genomic DNA extracted from rosette leaves. The positively identified mutant plants were propagated. The *eif2α(2g)-1* allele was examined by a wild-type probe using AT2-5'INT-F [2p-F](5'-GGACTGCCAAATGTGGATTCGAGC-3') + AT2-eIF2α-mid-R [2p-R](5'-CAAACTGCTAACACTTCGGATCC-3') primers and a pROK2 T-DNA probe using LBP1.3ext(5'-ATTTTGCCGATTTCGGAACCACC-3') + AT2-eIF2α-mid-R [2p-R](5'-CAAACTGCTAACACTTCGGATCC-3') primers. The *eif2α(2g)-2* allele was examined by a *Wild-type* probe using gAT2-eIF2α-intron5-F [2i6-F](5'-TGTGTAAATGCTATGTTTTGTCCCATTTTA-3') + gAT2-eIF2α-3'UTR-R [2i6-R](5'-GTAACCACCCCCTATGGGTAAAA-3') primers and a pROK2 T-DNA probe using gAT2-eIF2α-intron5-F [2i6-F](5'-TGTGTAAATGCTATGTTTTGTCCCATTTTA-3') + LBP1.3ext(5'-ATTTTGCCGATTTCGGAACCACC-3') primers. The *eif2α(5g)-1* allele was examined by a wild-type probe using AT5-5'INTeIF2α-F [5p-F](5'-CCGTGAGAATGTGAGCGACTC-3') + AT5-eIF2α-mid-R2 [5p-R](5'-TGTTGTTACCTCCACACCATCAGGT-3') primers and a pDAP101 T-DNA probe using LB3(5'-TAGCATCTGAATTTCATAACCAATCTCGATACAC-3') + AT5-eIF2α-mid-R2 [5p-R](5'-TGTTGTTACCTCCACACCATCAGGT-3') primers. The *eif2α(5g)-2* allele was examined by

a wild-type probe using SK06_RP [5E4-F] (5'-AAGCATGGACATGCTTTTGAG-3') + SK06_LP [5E4-R](5'-TCCCTCATCACCACTCATTTC-3') primers and a pROK2 T-DNA probe using SK06_RP [5E4-F](5'-AAGCATGGACATGCTTTTGAG-3') + LBP1.3ext (5'-ATTTTGCCGATTTCGGAACCACC-3') primers. The *eif2α(5g)-3* allele was examined by a wild-type probe using AT5-eIF2α-mid-F2 [5i6-F](5'-TGAAGAATTGAAAGATGCATTTTTGAAGGACA-3') + AT5-eIF2α-end'CDS-R [5i6-R](5'-GCTTGTACTACTAACCATGTTTGGGGT-3') primers and a pROK2 T-DNA probe using LBP1.3ext (5'-ATTTTGCCGATTTCGGAACCACC-3') + AT5-eIF2α-end'CDS-R [5i6-R](5'-GCTTGTACTACTAACCATGTTTGGGGT-3') primers.

The PCR experiment was performed using 1μL of 50ng/μL total genomic DNA according to the manufacturer's protocol for the PrimeSTAR Max DNA Polymerase kit (TakaRa, Cat # R045) or 2ul of 50ng/μL total genomic DNA when using the protocol for DreamTaq Green DNA Polymerase kit (EP0711). A single 20 μL reaction for the DreamTaq protocol was performed as follows: 2 μL – 10x DreamTaq Buffer, 0.55 μM Forward primer, 0.55 μM Reverse primer, dNTP mix diluted to 0.2mM of each nucleotide, 0.1 μL (0.5U) of DreamTaq Polymerase, 2 μL of gDNA template, raise to 20 μL with sterile water PCR-grade water. The amplicons were amplified by PCR according to the following protocol: Initial denaturing 95 °C for 2 minutes, cycle denaturing 95 °C for 30 seconds, annealing 57 °C for 30 seconds, cycle extension 72 °C for 30 seconds, final extension 72 °C for 5 minutes; cycle denaturing, annealing and extension were performed 33 times.

Final PCR products were visualized on a standard 1% agarose gel using UV-fluorescence of ethidium bromide.

2.3. Arabidopsis Growth and Harvest

2.3.1 Seedling preparation for germination

Seeds were sterilized in a 70% ethanol wash followed by 15-minute incubation in a solution of 50% bleach + 0.01% Tween-20. Following sterilization, seeds were washed with sterile water until detergent

micelles and bleach odor were no longer detectable. Seeds were stratified in the dark at 4 °C for 48 hours in sterile water.

After stratification, seeds were germinated on ½-concentration Murashige-Skoog (MS) plant media (MP Biomedicals, cat# 2633024) with 0.5-1% sucrose and 0.8% Phytoagar (Bioworld, cat # 40100072-2) in a 16 hr. light/ 8 hr. dark cycle, at 25 °C in 100uEin $m^{-2}s^{-1}$ light intensity. Three days after germination, seedlings can be transferred and arranged for a root length assay, or, between 3- and 7-days post-germination, seedlings can be transferred to soil for growth.

2.3.2 Arabidopsis soil growth

Arabidopsis seedlings were transferred to soil between 3- and 7- days of growth on ½-MS media with 1% sucrose. Seedling roots were planted into the depth of pre-dampened soil, in a small pot, to the base of the hypocotyl. After planting, seedlings were placed in the large growth chambers in the basement of Hesler Biology Building, covered by a transparent plastic lid taking care to avoid suffocating the seedlings. The lid will protect the seedlings from turbulent air in the growth chamber and ensure high humidity inside the lid area. The lied was removed after 1-2 days of adaptation to the new environment.

2.3.3 Arabidopsis DNA extraction

Extraction of genomic DNA was done by phenyl-chloroform method using phenyl:chloroform:isoamyl alcohol (25:24:1 v/v) (Thermofisher, Cat # 15593031). To extract DNA, a plant sample was ground in 300 µL of DNA extraction buffer solution (Composition: 0.1M Tris-HCl pH 8.0, 0.25M NaCl, 10mM EDTA, 0.1% SDS) inside of an Eppendorf 1.7 mL tube using a plastic pestle that snugly fits into the tube (Thermofisher, Cat # 12-141-364). An equal volume (300 µL) of phenol:chloroform:isoamyl alcohol solution was added to the tube and vortexed at medium-high speed for 30 seconds. The tubes were allowed to incubate at room temperature for at least 2 full minutes following the vortex step. The samples were centrifuged at 20,000x g for 5 minutes. After centrifugation, the upper aqueous layer (250 µL) was transferred into a new, pre-labeled tube (it is better to pipette 230 µL to avoid

transferring contaminants from other phases) and the old tube was discarded in the appropriate disposal vessel. Approximately 1/10th volume (25 μL) of 3M NaOAc pH 5.2 was added to the tube and mixed by inversion. Two volumes (500 μL) of ice cold 100% ethanol were then added to the tube and mixed by several inversions. The DNA was precipitated by incubating at -20 °C for at least 30 minutes (optimal precipitation: 24 hours at -80 C). Following precipitation, centrifuge for 10 minutes at 20,000x g and discard the supernatant. The pellet (may not be visible) was then washed with 700 μL of 70% ethanol stored on ice and centrifuged for 2 minutes at 20,000x g. The supernatant was discarded, and the pellet was washed again with 700 μL of 70% ethanol. Again, the supernatant was discarded, and the pellet was dried thoroughly (2hrs) in the chemical hood. Finally, the pellet was resuspended by light flicking in 100 μL of sterile deionized water for 1 minute and stored at -20 °C.

2.3.4 Arabidopsis protein extraction and quantification

Arabidopsis seedlings and rosette leaves were ground in a protein extraction buffer (Composition: 25mM Tris HCl pH 7.5, 75mM NaCl, 5% glycerol v/v, 0.05% NP-40 v/v, 0.5mM EDTA, 2mM DTT, 1x protease inhibitor [Halt Protease Inhibitor Cocktail, Cat# 78430]) in a 1.7 ml Eppendorf microfuge tube using a pestle. Ground plant tissue was centrifuged at 20,000x g for 10 minutes. The clarified supernatant was transferred to a new tube and the pellet was discarded.

Protein was quantified by Bradford assay. In a transparent 96-well plate for use with the ThermoFisher Multiskan MCC/340 plate reader, 48 μL of water, 2 μL of protein extract sample, and 150 μL of Pierce Coomassie Protein Assay Kit (Thermofisher, Cat # 23200) were combined in each well. Each protein sample was measured in triplicate or quadruplicate at A_{620} in the ThermoFisher Multiskan MCC/340 and concentrations were calculated against a bovine serum albumin standard curve.

Fresh extract was stored for no longer than 3 days at 4 °C or 3 months at -80 °C with SDS-PAGE sample buffer added to 1X concentration (1X composition: 50 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 1% β-ME, 12.5 mM EDTA, 0.02 % bromophenol blue).

2.3.5 Arabidopsis RNA extraction and purification

Arabidopsis seedlings and rosette leaves were harvested with liquid nitrogen and stored in autoclaved RNase-free microcentrifuge tubes inside the von Arnim Lab -80 ℃ until use. Tubes containing frozen tissue were placed in a wall-less tube rack resting in a shallow pool of liquid nitrogen. The frozen tissue was then ground in the autoclaved RNase-free microcentrifuge tube. The Invitrogen PureLink Plant RNA Reagent (Thermofisher, Cat # 12322012) manufacturer *small-scale* RNA isolation protocol was used to extract nucleic acids with DNase/RNase free materials.

DNase I treatment of RNA was performed by digesting DNA using the Invitrogen Turbo DNA-*free* kit (Thermofisher, Cat# AM1907). In brief, 10ug of RNA (can be scaled), was purified by combining 5 μL of 10x Turbo DNase buffer, 1 μL of Turbo-DNase I, and 10 μg of isolated RNA, raised to a total volume to 50 μL with DEPC-treated deionized water in a 1.7 mL tube. The mixture was inverted several times and then briefly centrifuged at low speed to recombine tube contents. The tube was then incubated for 25-minutes at 37 ℃. After incubation, the tube contained isolated RNA and digested DNA which was then either stored at -80 ℃ or the RNA was immediately precipitated as per the following instructions.

Prior to experimental use, DNase I was removed by creating a mixture of saturated phenol (pH 4.3; Cat# Fisher Scientific BP1751I-400):chloroform (Cat# Fisher Scientific C606-1) (1:1 v/v) wherein RNA was isolated in the aqueous phase of the mixture. The RNA-containing aqueous phase was transferred to an autoclaved 0.7mL tube band subjected to a standard ethanol precipitation procedure. The precipitated RNA was resuspended to the desired concentration with DEPC-treated deionized water.

2.4. Arabidopsis seedling root length assay

Seedlings that have been vertically germinated on ½-MS plates were transferred to a square ½-MS plate at 2-3 days post-germination. All seedling genotypes to be compared were present on each plate to control for plate-to-plate variation of the ½-MS media. All images of the seedlings were captured by mounted camera using a standard lens focal length, magnification, and distance from the plates at

predetermined time-points. Root lengths were measured by using the FIJI (imagej.net/software/fiji) platform free-hand measurement tool.

2.5. Arabidopsis Rosette area measurements

Rosette areas were measured using the Rosette Tracker plugin [33] in FIJI (ImageJ). The exact settings to use will be variable from experiment-to-experiment. The goal of adjusting the settings was to allow the software to discriminate between plant matter and soil. To describe the image analysis protocol briefly: upload an image of a rosette-stage plant and calibrate the software under "general" settings by specifying the number of plants to be measured, the orientation of the plants, and inputting a pixel-surface area scale. Next, choose the desired measurement and output options. Use the "calibrate colour" function in the "Day vs. Night" tab and choose the desired exclusion settings, save the settings, and click "ok". Finally, click "run" to measure rosette area and automatically export measurements.

2.6. Arabidopsis protein expression measured by western blot

2.6.1 SDS-PAGE separation and western blot transfer

10 µg of total protein, 2 µL of Precision Plus Protein Dual Color Standard (BioRad, Cat# 1610374), and 50 pg of purified eIF2α protein from Arabidopsis thaliana were loaded onto a 10mm long stacking and 55mm long resolving polyacrylamide SDS gel (4% stacking, 12% resolving) that was secured in a BioRad Mini-PROTEAN electrode assembly (BioRad, Cat# 1658037). Before loading, the electrode assembly was submerged in an SDS-PAGE buffer tank filled with SDS-PAGE running buffer (composition: 192mM glycine, 25mM tris, 4mM SDS). Proteins were stacked at 100V for approximately 15 minutes and resolved at 180V for approximately 60 minutes. SDS-PAGE -separated protein was electrophoresed into a methanol-charged polyvinylidene fluoride (PVDF) membrane using a BioRad transblot wet-blot apparatus (BioRad cat#, 1703812). The transfer was performed at 100V for 50 minutes in transfer buffer (composition: 192mM glycine, 25mM tris).

2.6.2 Western blot and rabbit anti-eIF2α antibody chemiluminescence

The antibody used to measure eIF2α expression was an aliquot generously provided from the lab of Karen Browning, University of Texas, Austin [34]. The primary HNAT (HEPES-NaCl-Tween 20) wash buffer composition was: 10mM HEPES-KOH pH 7.6, 150mM NaCl, and 0.2% Tween 20 (v/v). The blocking buffer composition was: 0.2% BSA (w/v) and 5% non-fat dried milk dissolved in HNAT buffer.

The PVDF membrane was blocked in the blocking buffer with agitation on a rocking table for 1 hour at room temperature in a square petri dish. Alternatively, the blocking step can be done for 12-24 hours at 4°C. The blocking buffer was rinsed from the PVDF membrane with 10-15 mL of wash buffer. Next, the primary antibody, rabbit anti-eIF2α, was suspended in blocking buffer in a 1:6000 ratio to a volume of 6 mL and was poured over the blot, the petri dish was sealed with parafilm, and incubated for 1 hour at room temperature with agitation. Alternatively, the primary antibody incubation can be done from 12-72 hours in 4°C. Following the primary antibody incubation, the antibody is typically reused up to 5 times in a 10-day period by transferring the used antibody solution into a 15mL tube with sodium azide to a concentration of 0.02% (v/v) and stored at -20°C. The primary antibody was washed from the blot using the HNAT buffer in a 5-minute wash repeated 10 times. The PVDF membrane was then incubated for 1 hour with agitation in the secondary antibody, goat anti-rabbit IgG Antibody H+L HRPeroxidase (PI-1000-1), diluted to a factor of 1:20000 in HNAT buffer. The secondary antibody was washed from the blot using the HNAT buffer in a 5-minute wash repeated 10 times.

Chemiluminescence was induced by using SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermofisher, Cat # 34580) reagent according to the manufacturer's protocol. Chemiluminescence was measured on the BioRad ChemiDoc MP using the sequential capture feature on the left side of the screen. The first capture settings should be as follows: chemiluminescent blot, 647SP, no light, with the "auto optimal" exposure set to the highest resolution. The second capture settings should be as follows: Colorimetric blot, 590/110, white epi, with the "auto optimal" exposure set to the highest resolution. These settings should be entered before using the chemiluminescent substrate. The substrate-

treated PVDF membrane was shaken off to remove excess substrate buffer before being placed in the center of the black tray. The image was captured using the green capture button on the bottom right of the display. The exposure times were adjusted, and the image recaptured if necessary. Images were exported to a flash drive in the BioRad analytical format and compressed JPEG format.

2.7. Measuring mRNA expression by RT-qPCR in Arabidopsis RNA extract

The reverse transcription quantitative PCR (RT-qPCR) experiments were done using the BioRad iTaq Universal SYBR Green One-step kit (BioRad, Cat #1725151) in the QuantStudio 5 qPCR thermocycling system. Components for a single reaction were assembled as follows: 50ng of DNase-treated Arabidopsis RNA, 5 μL 2x iTaq buffer, 3 μM forward primer, 3 μM reverse primer, 0.125 μL iScript Reverse Transcriptase, and DEPC-treated deionized water to bring the total reaction volume to 10 μL (more or less RNA can be used as long as it is a consistent, equal mass). The PCR cycle was programmed into QuantStudio v1.5.2 as follows: RT-reaction (50 °C for 2 minutes), initial denaturing stage (95 °C for 1 minute), PCR stage (95 °C for 15 seconds, 60 °C for 1 minute) for 40 cycles with data capture on the 60 °C step, Melt Curve stage (95 °C for 15 seconds, 65 °C for 1 minute, 95 °C for 1 second) with 5 second hold and data capture in 1.6 °C increments on the 65 °C to 95 °C temperature ramp. RT-qPCR-calculated cycle time (CT) values and DNA melt curve data were exported from QuantStudio v1.5.2 for statistical analysis.

Wild-type, *eif2α(2g)-1*, *eif2α(2g)-2*, *eif2α(5g)-1*, *eif2α(5g)-2*, and *eif2α(5g)-3* RNA samples were each probed on a 96-well plate using primers targeting mRNA expression of *eIF2α(2g), eIF2α(5g)*, and total *eIF2α,* with *EF1α* and *UBQ10* expression scored as the reference. The 96-well plate was arranged so that the wild-type and each allele mutant samples would be probed in triplicate at the seedling and rosette stages. The CT data were collected by the QuantStudio 5 cycler and exported to a QuantStudio v1.5.2 file. The technical replicate triplicate CT values for each sample-target combination were averaged together to form a biological replicate average CT value. A total of 3 biological replicates were analyzed for each experiment. The relative expression was measured by the PFAFFL method as described in *https://doi.org/10.1093/nar/29.9.e45* [35] which is a method of calculating fold expression while

18

controlling for primer efficiency. The standard error was calculated from the individual biological replicate fold induction values. A student's t-test assuming for unequal variance was performed on the canonical ΔCT values ($CT_{sample} - CT_{reference}$) using the 3 biological replicate values for each sample as an array.

2.7.1 RT-qPCR primers and primer efficiency experiment

The primers for the *eif2α(2g)* paralog mRNA were qPCR-AT2-CDSeIF2-F2(5'-GATTGAAATTCTAAACAAAGCCATAGCAGCA-3') [Chr2-F] and qPCR-AT2-CDSeIF2-R2(5'-AAGACGTAGCTTAGCCATGTGTTCTGTC-3') [Chr2-R]. Primers for *eIF2α(5g)* paralog mRNA were qPCR-AT5eIF2-end'CDS-F(5'-GGTAGCGGGATAATTGAATGAACAAAAGC-3') [Chr5-F] and qPCR-AT5eIF2-end'CDS-R(5'-GCTTGTACTACTAACCATGTTTGGGGT-3') [Chr5-R]. The primer design for total *eIF2α* expression targeted a region of both *eIF2α* paralogs that shared high sequence similarity. In the forward primer, there is one base pair mismatch against the *eIF2α(2g)* paralog sequence and one base pair mismatch against the *eIF2α(5g)* paralog sequence; thus, the same is true for the reverse primer set. The primers used for measuring total eIF2α mRNA levels were qPCR-eIF2-F(5'-GTGAGTGAG$_{(2g)}$CGTGATGAA$_{(5g)}$A-3') [eIF2α-F] and qPCR-eIF2α-R(5'-TCATCG$_{(5g)}$CCA$_{(2g)}$CTCATTTCTTCATT-3') [eIF2α-R] with a sub-denotation of the paralog for which the base pair is mismatched. The reference mRNAs for the RT-qPCR were translation elongation factor alpha (EF1α AT5G60390), detected with primers EF1α-F(5'-TCTCCGAGTACCCACCTTTG-3') [EF1α-F] and EF1α-R(5'-CTCCAGTTGGGTCCTTCTTG-3') [EF1α-R], and ubiquitin 10 (UBQ10 AT4G05320), detected with UBQ10-F(5'-ACCCTAACGGGAAAGACGAT-3') [UBQ10-F] and UBQ10-R(5'-GAGTTCTGCCATCCTCCAAC-3') [UBQ10-R].

The primer efficiency experiment was performed by creating a standard curve of CT values of wild-type RNA samples using primers for a particular *eIF2α* paralog target. 1μg of wild-type RNA was loaded in the first well of a PCR reaction plate and four more RNA samples were added in concentrations descending by 1 order of magnitude in each well. This primer efficiency experimental setup was repeated in triplicate for each primer pair used to target total *eIF2α*, *eIF2α(2g)*, *eIF2α(5g)*, *EF1α* and *UBQ-10* cDNA.

The wild-type RNA samples were examined by RT-qPCR following the PCR protocol from section 2.7 and a standard curve was plotted for [Log_{10} fold-dilution of sample] vs. [CT value] for each primer pair. The slope of the standard curve for each primer pair is used to determine the primer efficiency. The primer efficiency equation used is: $PE = 10^{\left(\frac{1}{SLOPE}\right)} - 1$.

2.8 Generation of transgenic plants expressing mutant alleles of eIF2α

2.8.1 Cloning of tagged AteIF2α transgene into Arabidopsis plants

The complete coding sequences (CDS) of Arabidopsis *eIF2α(2g)* and *eIF2α(5g)* were amplified from Arabidopsis cDNA using primers nested with restriction enzyme sequences. Phospho-deficient (S56A) and phospho-mimetic (S56D) alleles were created by site-directed mutagenesis. This was done by amplifying the 5' and 3' side of the gene in relation to the S56 codon with primers that overlap the S56 codon, re-coding it as an alanine or aspartate. In subsequent PCR reactions, the codon-56 region serves as an internal annealing site for the 5' and 3' *eIF2α* fragments. The restriction-digested *eIF2α* gene was ligated into a TOPO vector pCR4. The TOPO vector insert was then ligated by LR reaction into a plant gene expression vector under a constitutive 35S promoter in pEG100, or the native *eIF2α(2g)* 5'UTR promoter in the pEG301 vector, and an OCS transcriptional terminator. Both vectors also harbor a basta herbicide resistance gene (BAR^R) as a selectable marker. The pEG vectors were transformed into *Agrobacterium tumefaciens* GV3101 and maintained on LB medium by selection with kanamycin, rifampicin and gentamycin for Arabidopsis floral dip transgene transformation. Wild-type, *gcn2-1,* and *eif2α(5g)-2* were subjected to GV3101 floral dips and their genotypes are described in Table 3.

2.8.2 Genotyping and media selection of transgenic plants

T$_1$ generation seeds were screened for the transgene on ½ MS medium with 1% sucrose and 10 μg/mL BASTA$^+$. Resistant plants were transferred to soil for genotyping and propagation after 6 days of growth.

Plants were genotyped to verify the presence of endogenous *eIF2α(2g)* and *eIF2α(5g)* using gene-specific primers and the presence of the transgenic *eIF2α* using a vector-specific reverse primer, in addition to endogenous GCN2, *gcn2-1* and *eif2α(5g)-2* alleles when appropriate. The endogenous *eIF2α(2g)* allele was examined by a *Wild-type* probe using gAT2-eIF2α-intron5-F(5'-GGACTGCCAAATGTGGATTCGAGC-3') + qPCR-AT2-5'UTR-R(5'-CGGAAAATGAGTCGACTTGCG-3') primers. The endogenous *eIF2α(5g)* allele was examined by a wild-type probe using AT5-5'INTeIF2α-F(5'-CCGTGAGAATGTGAGCGACTC-3') + AT5-eIF2α-mid-R2(5'-TGTTGTTACCTCCACACCATCAGGT-3') primers and, in the case of a transgene in an *eif2α(5g)-2* background, a pROK2 T-DNA probe using AT5-eIF2α-mid-F2(5'-TGAAGAATTGAAAGATGCATTTTTGAAGGACA-3') + LBP1.3ext(5'-ATTTTGCCGATTTCGGAACCACC-3') primers. The transgenic *eIF2α(2g)* CDS alleles were examined by a transgene probe using qPCR-AT2-CDSeIF2α-F(5'-AAGACGTAGCTTAGCCATGTGTTCTGTC-3') + JucMarkSEQ-R(5'-TATCATGCGATCATAGGCGTCTCGCA-3'). Transgenic native promoter alleles and *eIF2α(5g)* CDS alleles were examined by a transgene probe using AT5-eIF2α-mid-F2 (5'-TGAAGAATTGAAAGATGCATTTTTGAAGGACA-3') + JucMarkSEQ-R(5'-TATCATGCGATCATAGGCGTCTCGCA-3'). The *wild-type* GCN2 allele probe utilized SKC-049(5'-CGAAGGTTTTGTTAGTGCTGG-3') forward + SKC-048(5'-AAAATATTCCCATGCCTGGTCC-3') reverse primers and the *gcn2-1* allele probe utilized SKC-049(5'-CGAAGGTTTTGTTAGTGCTGG-3') forward + DS3-2(5'-CGATTACCGTATTTATCCCGTTC-3') reverse primers.

2.9. Cloning, expression, and purification of recombinant proteins

2.9.1 eIF2α

Arabidopsis RNA was reverse transcribed to create a cDNA library. Restriction enzyme-encoded primers for eIF2α(2g) cDNA, NdeI-eIF2α(2g)-F(5'-ATT<u>CATATG</u>GCGAGTCAAAC-3') and EcoRI-eIF2α(2g)-R(5'-ATT<u>GAATTC</u>TTACTCGATGATTCC-3'), were used to transcribe the eIF2α(2g) CDS

with restriction enzyme overhangs to facilitate cloning. The *eIF2α(2g)* S56A and S56D alleles were created by site-directed mutagenesis using internal gene primers substituting the 56th serine for alanine or aspartate.

The S56A allele was created using NdeI-eIF2α(2g)-F + eIF2aChr2-S56A-R(5'-GACGACGAGCGAGCTCGGAGAAA-3') and eIF2aChr2-S56A-F(5'-TTTCTCCGAGCTCGCTCGTCGTC-3') + EcoRI-eIF2α(2g)-R. The S56A specific primers have the property of being reverse complemented to each other and thus the NdeI-eIF2α(2g)-F + eIF2aChr2-S56A-R and eIF2aChr2-S56A-F + EcoRI-eIF2α(2g)-R amplicon products serve as internal primers for the full length eIF2α(2g) gene. A final PCR using each of the two amplicons as both primer and template was performed to amplify a 1050bp NdeI-CDSeIF2α(2g)S56A-EcoRI amplicon. The target amplicon was excised from the gel, sequenced, digested and stored for cloning.

The NdeI-CDSeIF2α(2g)S56D-EcoRI cloning insert was created in the same manner using eIF2aChr2-S56D-F(5'-TTTCTCCGAGCTCGATCGTCGTC-3') and eIF2aChr2-S56D-R(5'-GACGACGATCGAGCTCGGAGAAA-3'). In total, three NdeI-CDSeIF2α(2g)-EcoRI constructs were cloned with either a Wild-type, S56A, or S56D allele. Each of the three digested NdeI-CDSeIF2α(2g)-EcoRI alleles were ligated into respective TOPO vectors then transformed into, and propagated by, a Top10 *E. coli* BL21STAR . The propagated TOPO-eIF2α(2g) inserts were then digested and ligated into pET28a expression plasmids and transformed into *E. coli* BL21STAR expression cell line.

The eIF2α protein harboring a polyhistidine tag at the N-terminus was expressed in the *E. coli* BL21Star cell line. The culture was grown in an LB broth (25g/L; BP1426-2) culture at 37°C, with constant shaking at 250RPM, until the culture reached an OD_{600} of 0.6. The cell culture was then grown for 18 hours at 22°C following 1mM IPTG induction, with constant shaking at 250RPM. Following protein expression, cultures were pelleted by centrifugation for 10 minutes at 5,000x *g*. The cell pellet was resuspended in 25 mL lysis buffer [composition: 30mM tris HCl pH 7.5, 300mM NaCl (w/v), 0.1mM DTT, 10mM imidazole, 10mM $MgCl_2$, 1% (v/v) triton X100, 0.5% (v/v) glycerol] and lysed by microfluidizer (Microfluidics

LM10) at 22,000 psi in 3 passes, so that cell culture turbidity was reduced and was clarified. The lysed cell culture was then centrifuged at 18,000x *g* for 30 minutes to isolate the soluble protein fraction.

His-tagged eIF2α was purified using GE Healthcare Ni Sepharose HP (Sigma-Aldrich; GE17-5268-01). The 45mL soluble protein fraction was batch equilibrated in wash buffer 1[composition: 30mM tris HCl pH 7.5, 300mM NaCl (w/v), 0.1mM DTT, 0.5% (v/v) triton X100, 0.5% (v/v) glycerol] with 1 mL of Ni Sepharose HP resin in a 50mL falcon tube, under constant rotation at 4 °C for 2 hours. The equilibrated resin was centrifuged at 100x *g* for 5 minutes. All pelleted resin was transferred to a purification column. The resin was washed in 40 column volumes of wash buffer 1, 10 column volumes of Wash buffer 2 [composition: 30mM tris HCl pH 7.5, 50mM NaCl (w/v), 0.1mM DTT, 60mM imidazole, 0.5% (v/v) glycerol], and eluted in 1 mL fractions of elution buffer [composition: 30mM tris HCl pH 7.5, 0.1mM DTT, 300mM imidazole, 0.025% (v/v) glycerol]. The elution step was performed by pipetting 0.5mL of elution buffer onto the resin directly, taking care to avoid touching the walls of the column. The protein was eluted until the buffer meniscus was level with the resin and an additional 0.5 mL of elution buffer was added. A 2ul aliquot from each elution fraction was distributed into PCR strip tubes containing Bradford reagent and fraction collection was stopped when low levels of protein were being eluted. Fractions containing protein were combined into a sealed dialysis membrane with 10 kDa molecular weight cut off and dialyzed for 24 hours in dialysis buffer [composition: 20mM Tris HCl pH 7.5, 50mM NaCl, 0.5mM DTT, 20% (v/v) glycerol]. The protein concentration was calculated by Bradford assay.

2.9.2 ABCF1

Arabidopsis RNA was reverse transcribed to create a cDNA library. Restriction enzyme-encoded primers for *ABCF1* cDNA were used to transcribe the *ABCF1* CDS. The *ABCF1* gene underwent a silent mutation by site-directed mutagenesis as described in section 2.9.1 to remove an internal restriction site that would interfere with the cloning of *ABCF1* using the NcoI-ABCF1-F(5'-ATT<u>CCATGG</u>GGCGAGCAAGAAGAAGG-3') + ABCF1-R1(5'-TGCATCCATAGCGTAACCTCTCATAG-3'), and ABCF1-F1(5'-

23

TGAGAGGTTAGACCAGAAACCGCTGA-3') + BamHI-ABCF1-R(5'-ATT<u>GGATCC</u>TCAAAAGTCCGGCTTTA-3') primer pairs. The ABCF1-F1 and ABCF1-R1 primers were the reverse complement of each other. When the amplification of *ABCF1* was completed, the full-length amplicon was digested and ligated into a TOPO vector and was transformed into, and propagated by, a Top10 *E. coli* cell line. The insert was then digested and ligated into the pDEST15 expression plasmid and transformed into BL21Star *E. coli* expression cell line.

The *E. coli* cell line harboring pDEST15-*ABCF1* was grown in an LB broth (25g/L; BP1426-2) culture at 37°C, with constant shaking at 250RPM, until the culture reached an OD_{600} of 0.6. The cell culture was grown for 18 hours at 22°C following 1mM IPTG induction, with constant shaking at 250RPM. Following protein expression, cultures were pelleted by centrifugation for 10 minutes at 5,000x *g*. The cell pellet was resuspended in 25 mL lysis buffer [composition: 30mM tris HCl pH 7.5, 300mM NaCl (w/v), 0.1mM DTT, 10mM imidazole, 10mM $MgCl_2$, 1% (v/v) triton X100, 0.5% (v/v) glycerol] and lysed by combination of lysozyme (0.05g/25mL lysis buffer; Sigma L-6876) and sonication (Fisher Scientific Sonic Dismembrator 550). The sample should be kept on ice during lysozyme incubation for 1 hour, then proceed to sonication. The sonication interval was 20 seconds of sonication at level 5 intensity followed by a 40 second rest repeated 9 times with the microtip, so that cell culture turbidity was reduced and was clarified. During sonication, the sample vessel was suspended in an ice bath to dissipate the heat which is generated from sonication. The lysed cell culture was then centrifugated at 18,000x *g* for 30 minutes to isolate the soluble protein fraction.

The 25mL of soluble cell lysate fraction was raised to 50mL of volume using glutathione equilibration/wash buffer [composition: 50mM Tris, 150mM NaCl, pH 8.0], and batch equilibrated a 50mL tube containing 0.75 mL of glutathione resin (ThermoFisher Cat#16101), in 4°C during constant, gentle inversion for 1 hour. The glutathione resin was gently pelleted by centrifugation at less than 100 x *g* and then transferred to the purification column by transfer pipette. The column was washed with 20 column volumes of glutathione equilibration/wash buffer. Fractions were collected after treatment the elution buffer

[composition: 50mM Tris, 150mM NaCl, pH 8.0] containing 10mM reduced glutathione as the eluent. Eluent was added and fractions collected in 1mL increments. The appropriate fractions were combined, dialyzed buffer [composition: 20mM Tris HCl pH 7.5, 50mM NaCl, 0.5mM DTT, 20% (v/v) glycerol], quantified by Bradford assay, and the GST-ABCF1 protein was stored at -80°C. Fractions containing protein were combined into a sealed dialysis membrane with 10 kDa molecular weight cut off and dialyzed for 24 hours in dialysis The protein concentration was calculated by Bradford assay.

2.9.3 GCN2-KD

Arabidopsis RNA was reverse transcribed to create a cDNA library. The sequence of the kinase domain was determined by the predictive domain analysis software InterPro (https://www.ebi.ac.uk/interpro/). Restriction enzyme-encoded primers for GCN2-KD cDNA, NheI-GCN2(KD)-F(5'-ATT<u>GCTAGC</u>AGCTTCTCAATGATTTT-3') + NotI-GCN2(KD)-R(5'-ATTG<u>CGGCCG</u>CTTGAAGTTCTGTGGC-3') , were used to transcribe the GCN2-Kinase domain CDS. The amplicon was excised from an agarose gel and digested by NheI and NotI. The digested amplicon was ligated into a TOPO vector and was transformed into, and propagated by, a Top10 *E. coli* cell line. The insert was then digested and ligated into the pET28a expression plasmid and transformed into BL21Star *E. coli* expression cell line.

The *E. coli* cell line harboring pET28a-GCN2-KD was grown in an LB broth (25g/L; BP1426-2) culture at 37°C, with constant shaking at 250RPM, until the culture reached an OD_{600} of 0.6. The cell culture was grown for 18 hours at 22 °C following 0.5 mM IPTG induction, with constant shaking at 250RPM. Following protein expression, cultures were pelleted by centrifugation for 10 minutes at 5,000x *g*. The cell pellet was resuspended in 25 mL lysis buffer [composition: 30mM tris HCl pH 7.5, 300mM NaCl (w/v), 0.1mM DTT, 10mM imidazole, 10mM $MgCl_2$, 1% (v/v) triton X100, 0.5% (v/v) glycerol in 1X PBS] and lysed by FRENCH Press (ThermoFisher FA-078A) at 1300p.s.i. using a pre-refrigerated pressure cell. The FRENCH Press lysis was repeated 2 more times so that cell culture turbidity was reduced and was clarified.

The lysed cell culture was then centrifugated at 18,000x *g* for 30 minutes to isolate the soluble protein fraction.

The His-tagged GCN2-KD was purified using GE Healthcare Ni Sepharose HP (Sigma-Aldrich; GE17-5268-01). The 25mL soluble protein fraction was raised to a volume of 50mL by addition of GCN2-KD wash buffer 1 [composition: 30mM tris HCl pH 7.5, 300mM NaCl (w/v), 0.1mM DTT, 0.5% (v/v) triton X100, 0.5% (v/v) glycerol in 1X PBS] with 1 mL of Ni Sepharose HP resin in a 50mL falcon tube, under constant rotation at 4 °C for 2 hours. The equilibrated resin was centrifuged at 100x *g* for 5 minutes. All pelleted resin was transferred to a purification column using a transfer pipette. The resin was washed in 40 column volumes of GCN2-KD wash buffer 1, 10 column volumes of Wash buffer 2 [composition: 30mM tris HCl pH 7.5, 50mM NaCl (w/v), 0.1mM DTT, 60mM imidazole, 0.5% (v/v) glycerol in 1X PBS] and eluted in 1 mL fractions of elution buffer [composition: 30mM tris HCl pH 7.5, 0.05mM DTT, 300mM imidazole, 0.025% (v/v) glycerol in 1X PBS]. Fractions containing protein were combined into a sealed dialysis membrane with 10 kDa molecular weight cut off and dialyzed for 24 hours in dialysis buffer [composition: 20mM Tris HCl pH 7.5, 50mM NaCl, 0.5mM DTT, 20% (v/v) glycerol in 1X PBS]. The protein concentration was calculated by Bradford assay.

Chapter 3.
Results

The goal of this project was to establish functional *eIF2α* allele mutants in Arabidopsis, determine the role of *eIF2α* in the growth and development of the plant, and to express recombinant proteins involved in the ISR for biochemical analysis. A preliminary account of results was presented in an Honor's thesis by Jeremiah Holt in 2016 at the University of Tennessee which described the phenotypic characterization of two Arabidopsis mutants and recombinant protein expression.

3.1 Genotypes of eIF2α allele mutants were verified by PCR

Arabidopsis strains with T-DNA insertions in the two paralogous genes for *eIF2α* were procured from the Arabidopsis Biological Resource Center at Ohio State University. The strains were examined to determine their status as *eIF2α* mutants and elucidate the role of eIF2α in the growth and development of Arabidopsis. There were two *eIF2α(2g*) mutant alleles and three eIF2α(5g) mutant alleles [Figure 4a]. The *eIF2α(2g)* mutant alleles have T-DNA insertions in the upstream 5' upstream promoter region near the transcription start site and the 6th intron, named *eif2α(2g)-1* and *eif21α(2g)-2*, respectively. The *eIF2α(5g)* mutant alleles have T-DNA insertions in the upstream 5'UTR promoter region, the 4th exon and the 6th intron, named *eif2α(5g)-1*, *eif2α(5g)-2* and *eif2α(5g)-3*, respectively [Figure 4a]. Analytical PCR was done to determine the genotype of the mutants and confirm that T-DNA insertions were present and WT alleles were absent in the mutant lines [Figure 4b; Table 1]. Heterozygous and homozygous mutant lines were confirmed in all five mutant alleles and were planted for phenotypic analysis and propagation.

3.1.1 Seedling growth and development

In a seedling root length assay experiment, it was observed that *eif2α(2g)-2* was severely stunted in growth and delayed germination. The *eif2α(2g)-2* mutants in Figure 5a were germinated 5 days earlier than the other seedlings in the assay and the timeline of delayed germination and development is demonstrated in Figure 5c. The *eif2α(2g)-2* mutant also showed an increased expression of anthocyanin. At the seedling stage *eif2α(2g)-1* and all *eIF2α(5g)* mutant alleles appeared to develop as WT [Figure 5a,b].
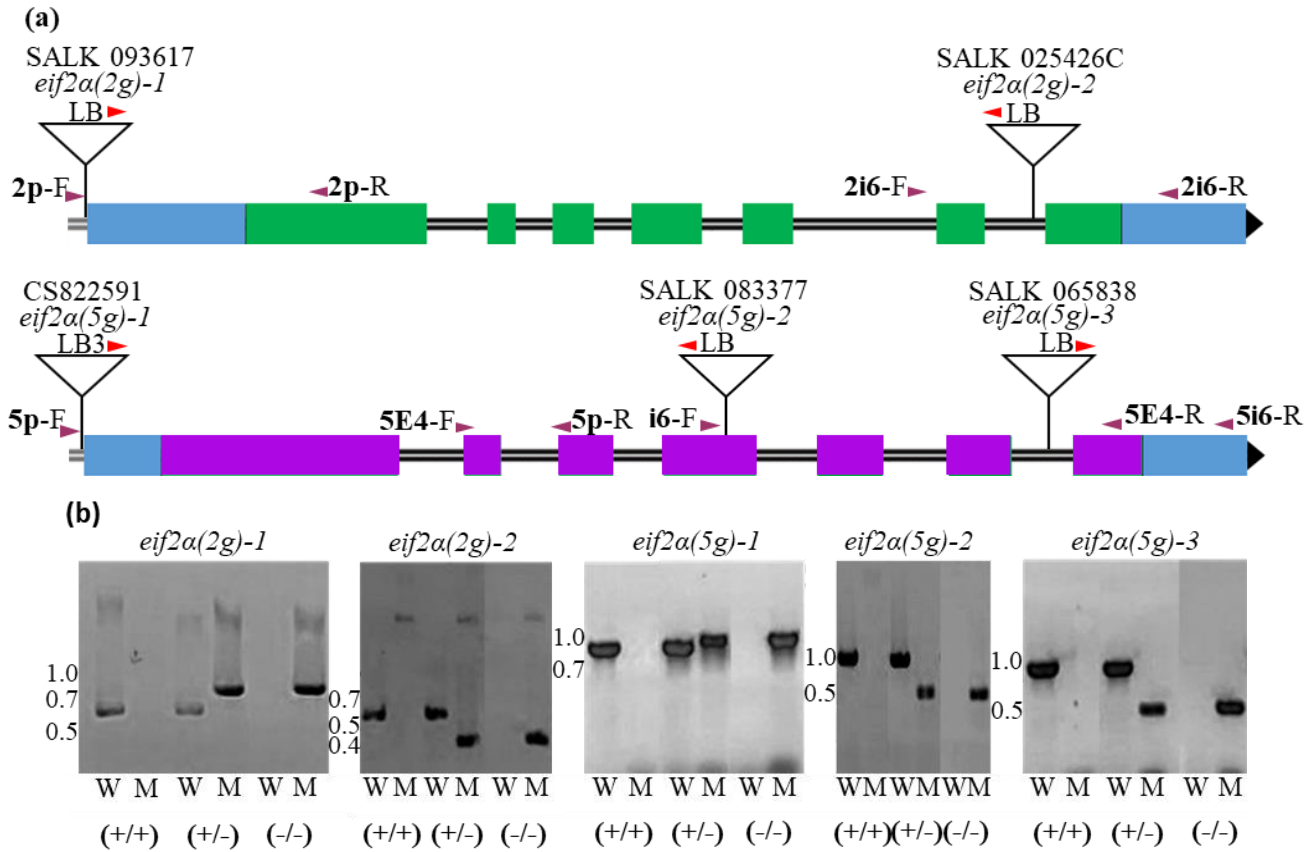
Figure 4 – Genotyping of Arabidopsis Single-allele eIF2α Mutants

Panel (a) Shows a gene model of Chromosome 2 eIF2, *eIF2α(2g)*, and Chromosome 5 eIF2α, *eIF2α(5g)*. The UTRs are green, the introns are black, the CDS of *eIF2α(2g)* is green and the CDS of *eIF2α(5g)* is magenta. T-DNA insertions are notated at their precise location. The name of mutant allele lines and their SALK or SAIL line numbers are notated above their corresponding T-DNA insertion. The approximate location of primers used for genotyping are shown in the gene model and noted in Table 1. (b) 1% Agarose gels showing results of analytical PCR to determine the genotype of the eIF2α mutant alleles. The "W" label signifies that product of the wild-type eIF2α allele reaction of the relevant eIF2α paralog is displayed in that lane. The "M" label signifies that the products of the reaction probing for the relevant eIF2α mutant allele is displayed in that lane and is described further in Table 1. Segregating heterozygous seed stocks from ABRC were used to grow a population of rosette stage plants for genotyping. The population of plants for each mutant line were genotyped to differentiate between plants that segregated into wild-type (+/+), heterozygous (+/-) and homozygous (-/-) mutant plants.

28

Table 1 – Arabidopsis eIF2α Allele Mutant Genotyping Primer Table

Table 1. Shows the primer pairs used for the analytical PCR in Figure 4. Each row demonstrates the primer pairs used to amplify the appropriate genomic DNA sequence for each mutant referenced in Figure 4a, and amplicon length of PCR reactions shown in Figure 4b for *eIF2α* allele mutants.

Mutant Line	*Wild-type* (W) Allele Primer Pairs	W Amplicon length (bp)	Mutant (M) Allele Primer Pairs	M Amplicon length (bp)
eif2α(2g)-1 (SALK_093617)	2p-F + 2p-R	600	LBP1.3 + 2p-R	780
eif2α(2g)-2 (SALK_025426C)	2i6-F +2i6-R	600	2i6-F + LBP1.3	430
eif2α(5g)-1 (CS_822591)	5p-F + 5p-R	930	LB3 + 5p-R	990
eif2α(5g)-2 (SALK_065838)	5E4-F + 5E4-R	1000	5E4-F + LBP1.3	530
eif2α(5g)-3 (SALK_083377)	5i6-F + 5i6R	900	LBP1.3 + 5i6R	520

29

3.1.2 Rosette growth and development

In rosette plants, *eif2α(2g)-1* and all *eIF2α(5g)* mutant alleles, again, appeared to develop as WT. The *eif2α(2g)-2* rosette leaves appear chlorotic or present very high levels of anthocyanin [Figure 5d]. Additionally, the *eif2α(2g)-2* mutant did not produce mature siliques and was infertile [Figure 5d]. Figure 5d demonstrates the morphology of maturing *eIF2α* mutant allele plants. As in seedlings, the mature *eif2α(2g)-2* mutant was also very severely stunted in growth [Figure 5d,e,f].

3.2 Arabidopsis *eIF2α* mutants have altered mRNA and protein expression

Having observed the effect of the loss of *eIF2α* in the *eIF2α(2g)* and *eIF2α(5g)* mutant alleles, the next step was to examine the expression *eIF2α* at the mRNA and protein level for each allele. The *eIF2α* paralogs have very similar protein sequences, but at the DNA level the paralogs differ somewhat in sequence. Expression data were collected using whole seedlings and mature rosette leaves from rosette-stage plants. To measure the mRNA expression, RT-qPCR was performed on each mutant allele and compared to WT mRNA expression of *eIF2α*. The experiment was performed using RNA from whole seedling extract and from mature rosette leaves.

Figure 5 – Effects of eIF2α Loss of Function in the Growth and Development of Arabidopsis Plants

Figure 5. (a) Arabidopsis seedlings germinated on ½ MS + 1% sucrose and used for measuring root lengths. (b) This panel shows the quantification of root lengths 6 days after germination (13 days after germination for *eif2α(2g)-2* mutants). The *eif2α(2g)-2* mutant roots were not large enough to measure until 6 days post germination. Error bars represent standard error. (c) This panel shows *eif2α(2g)-2* mutant seedling development at 0, 3, 5, 7 and 11 days after germination. (d) Demonstrates representative images of mature eIF2α allele mutants. The top panel scale bar represents 5mm and the bottom panel represents 15mm. (e) This panel shows an overhead view of rosette-stage Arabidopsis eIF2α mutants that were used to (f) determine eIF2α mutant rosette diameters using the ImageJ Rosette Tracker tool. Error bars represent standard error.

(a) WT *Col*   *eif2α(2g)-1*   *eif2α(2g)-2*   *eif2α(5g)-1*   *eif2α(5g)-2*   *eif2α(5g)-3*

(b) Root Length Assay

(c) G0 G3 G11 1.5 mm 1.5 mm G5 G7 2.5 mm 2.5 mm 2.5 mm

(d) WT *(Col)*   *eif2α(2g)-1 (-/-)*   *eif2α(2g)-2 (-/-)*   *eif2α(5g)-1 (-/-)*   *eif2α(5g)-2 (-/-)*   *eif2α(5g)-3 (-/-)*

(e) WT   *eif2α(2g)-1*   *eif2α(2g)-2*   *eif2α(5g)-1*   *eif2α(5g)-2*   *eif2α(5g)-3*

*P < 0.01; **P<0.001; ***P<0.0001

(f) Rosette Area
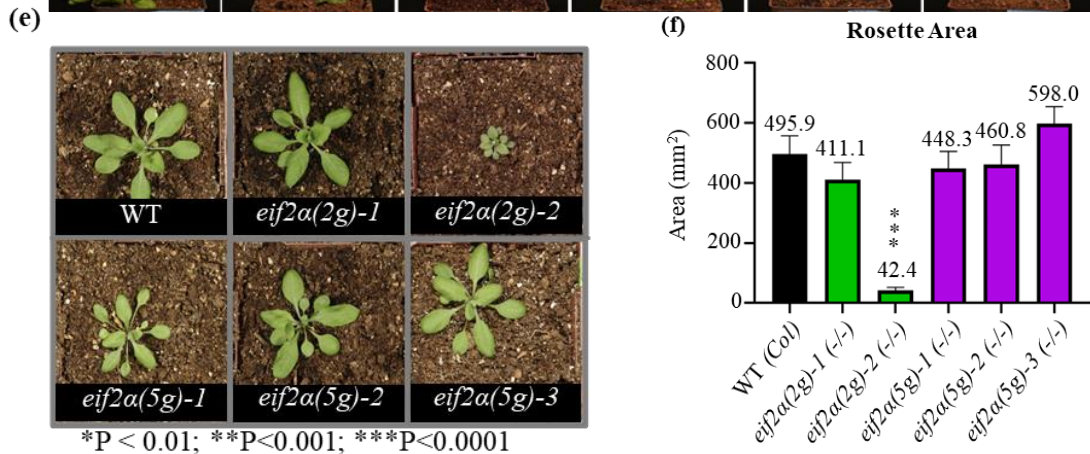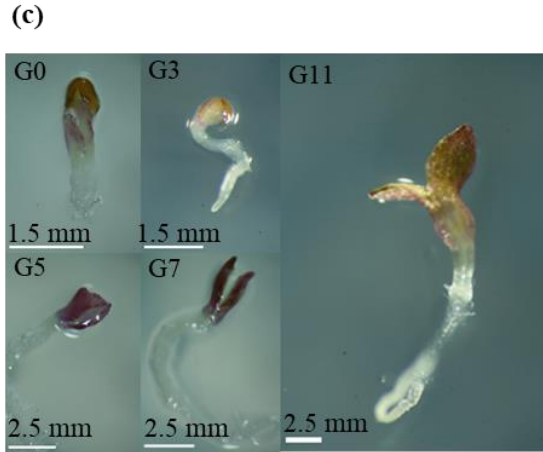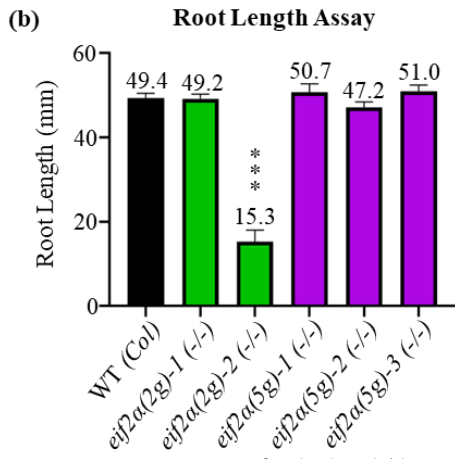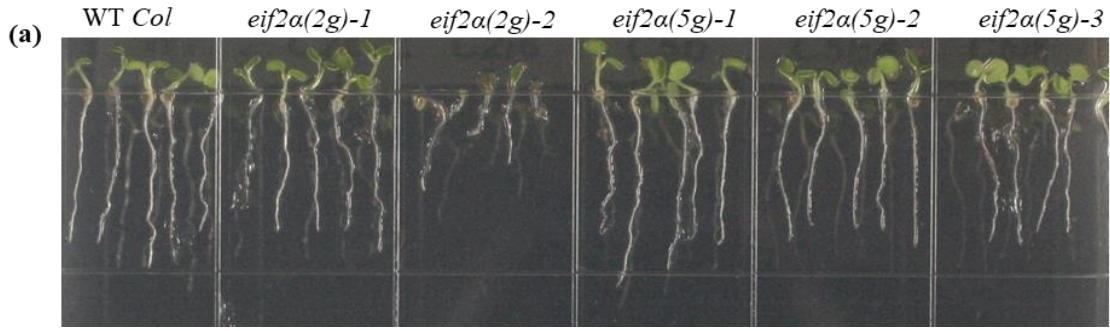
Three primer pairs were designed to measure *eIF2α* expressed specifically from *eIF2α(2g)*, *eIF2α(5g)* as well as both paralogs together in each mutant allele [Table 2; Figure 6a]. Since T-DNA insertions are known to disrupt transcription, primers were designed to anneal to sequences downstream from the TDNA insertion points[Figure 6a; Table 2]. Gene-specific primer pairs were selected to anneal to the 3'UTR of *eIF2α(5g)* and an *eIF2α(2g)*-specific sequence flanking the *eif2α(2g)-2* T-DNA insertion. When comparing the *eIF2α* paralog sequences, the coding region sequence is generally more similar than the UTRs and thus was the natural target for a primer pair to detect total *eIF2α* expression. However, the coding region DNA sequence is still quite different between the paralogs, therefore, it was impossible to design a primer that perfectly annealed to both paralogs in the coding region. Therefore, the forward and reverse primers were designed to anneal to both paralogs with similar efficiency by including one mismatched-base pair in the forward and the reverse primer, respectively, in each paralog [see materials and methods 2.7.1].

To measure the total protein expression, extracts from whole seedlings and mature rosette leaves were used. Western blotting was performed to probe protein extract with an anti-eIF2α antibody to assay total *eIF2α* expression compared to WT. Since *eIF2α* is paralogous in Arabidopsis, there is not a way to differentiate between *eIF2α(2g)* and *eIF2α(5g)* as they have nearly identical amino acid sequence and presumably have very similar epitopes.

Figure 6 – MRNA Expression Profiles for *eIF2α* Paralogs in Arabidopsis *eIF2α* Mutant Seedlings

Figure 6. (a) Shows a cDNA gene model of *eIF2α* UTRs in blue, *eIF2α(2g)* CDS in green and *eIF2α(5g)* CDS in magenta. Primers referenced in Table 2 are shown on the gene model with arrows indicating the primer target position and 5' to 3' orientation. (b) Here, the relative expression of *eif2α(5g)* mRNA in Arabidopsis seedlings for each mutant allele is shown. (c) In this panel, the relative expression of *eIF2α(2g)* mRNA in seedlings is shown. (d) The relative expression of total *eIF2α* mRNA is displayed in this panel. In all cases, either *EF1α* or *UBQ-10* expression was used as a reference to compare relative *eIF2α* expression and an unequal variance t-test was used. *P < 0.01; **P<0.001; ***P<0.0001.

(a)

Chr2-F ▶ ◀ Chr2-R

eIF2α(2g)

*eIF2α-F ▶ ◀ *eIF2α-R

Chr5-F ▶ ◀ Chr5-R

eIF2α(5g)

*eIF2α-F ▶ ◀ *eIF2α-R

(b) Seedling *eIF2α(2g)* Relative Expression (*EF1α*)

1.000 1.021 *** 1.23 1.114 1.190
0.001

(c) Seedling *eIF2α(5g)* Relative Expression (*EF1α*)

1.000 0.850 0.776 1.116 ** ***
0.058 0.104

(d) Seedling Total *eIF2α* Relative Expression (*EF1α*)

1.000 0.907 *** 1.319 1.034 1.038
0.041

(e) Seedling *eIF2α(2g)* Relative Expression (*UBQ-10*)

1.000 0.910 *** 1.017 1.351 0.929
0.001

(f) Seedling *eIF2α(5g)* Relative Expression (*UBQ-10*)

1.000 0.870 1.208 0.743 *** **
0.069 0.083

(g) Seedling Total *eIF2α* Relative Expression (*UBQ-10*)

1.000 1.006 * 0.914 1.344 0.868
0.068

WT (Col), eif2α(2g)-1 (-/-), eif2α(2g)-2 (-/-), eif2α(5g)-1 (-/-), eif2α(5g)-2 (-/-), eif2α(5`g)-3 (-/-)

33

Table 2 – RT-qPCR Primers Used for Calculating mRNA Expression of Arabidopsis *eIF2α*

Mutants

Table 2. Shows the primer pairs and relevant metrics for the primer targets used in real-time quantitative PCR in Figures 6 and 8 for determining expression using the PFAFFL method. Primer efficiency was calculated from the slope of a linear regression in an RNA dilution series for each primer target. The calculation is described in more detail in section 2.7.1.

Primer Target	Primer pairs	Primer Efficiency	Amplification factor	Amplicon length (bp)
eIF2α(2g)	Chr2-F + Chr2-R	0.93	1.93	145
eIF2α(5g)	Chr5-F + Chr5-R	1.07	2.07	135
Total *eIF2α*	eIF2α-F + eIF2α-R	1.06	2.06	80
EF1α	eEF1-F + eEF1-R	1.01	2.01	102
UBQ10	UBQ10-F + UBQ10-R	0.87	1.87	148

3.2.1 Primer efficiency

Before quantifying Arabidopsis mRNA expression via RT-qPCR, the primer efficiency, $Primer\ Efficiency = \left(10^{\frac{-1}{slope}} - 1\right)$, and amplification factor, $Amplification\ Factor = 1 + Primer\ Efficiency$, were determined for each primer pair using WT RNA extract. Briefly, the primer efficiency is determined from the slope of a standard curve which is generated by plotting the measured CT values of RNA samples at increasing concentrations (see methods 2.7.1). Shown in Table 2, all primer efficiencies were within the optimal 0.9-1.1 range except in the case of the UBQ10 target primers, which had a slightly reduced but still sufficient efficiency of 0.87. This result means that the PCR amplification factors for the primer targets are 2 ± 0.1, which is close to the theoretical optimum of 2.0 per cycle of amplification [Table 2]. Since the traditional ΔΔCT method of determining differences in mRNA levels assumes a primer pair efficiency of exactly 1.0, the RT-qPCR data were analyzed by the PFAFFL method which uses the experimentally determined primer pair efficiencies.

3.2.2 Seedling mRNA expression

Expression from the *eIF2α(2g)* paralog was measured using the Chr2-F and Chr2-R primers which anneal to the 3'UTR of the *eIF2α(2g)* cDNA [Figure 6a; Table 2]. The seedling RT-qPCR experiment revealed that the *eif2α(2g)-2* mutant was expressing virtually no *eIF2α(2g)* mRNA indicating a severe loss of function allele [Figure 6b]. The promoter mutant allele, *eif2α(2g)-1*, expresses *eIF2α(2g)* at the same level as wild type [Figure 6b], which is coherent with the absence of a growth phenotype in Figure 5. Plants with mutant *eIF2α(5g)* alleles expressed normal levels of *eIF2α(2g)* mRNA [Figure 6b] and vice versa [Figure 6c].

Using the Chr5-F and Chr5-R primers targeting the 3'UTR of *eIF2α(5g)*, the *eif2α(5g)-1* allele with a T-DNA in the 5' region of the gene had normal gene expression. However, *eif2α(5g)-2* and *eif2α(5g)-3* were identified as loss of function alleles expressing *eIF2α(5g)* mRNA at 6.0% and 10.4% of wild-type levels [Figure 6c]. Considering the near wild-type growth of the mutants [Figure 5], the *eif2α(5g)-2* and

eif2α(5g)-3 mutants demonstrate that the *eIF2α(5g)* gene is essentially dispensable under these growth conditions.

Total *eIF2α* expression was measured using the eIF2α-F and eIF2α-R primers which anneal at a section of the *eIF2α(2g)* and *eIF2α(5g)* CDS which share nearly identical DNA sequence [Figure 6a]. Total eIF2α mRNA expression in the null allele *eif2α(2g)-2* is 4.1% of the wild-type expression. Given that *eif2α(2g)-2* is a null allele, we conclude that the minor paralog *eIF2α(5g)* is expressed at 24-fold lower levels than *eIF2α(2g)* [Figure 6d]. The total *eIF2α* mRNA expression data is congruent with the plant phenotypic data in Figure 5 in demonstrating that the *eIF2α(2g)* gene is the more important of the two *eIF2α* paralogs. These data were gathered using *EF1α* as a reference gene and the results were mimicked when using *UBQ-10* as a reference gene [Figure 6e,f,g]

3.2.3 Seedling protein expression

Total eIF2α protein expression in seedlings was measured by western blot utilizing a rabbit anti-eIF2α antibody (see methods 2.6.2). In this experiment, we saw that the mRNA-null allele *eif2α(2g)-2* had significantly reduced eIF2α protein expression compared to WT while the other alleles displayed WT expression of *eIF2α* in seedlings [Figure 7a,b]. This result, taken together with the mRNA expression data suggests that the *eIF2α(2g)* paralog is responsible for the majority of total eIF2α protein being expressed in the plant. This conclusion is further supported by results shown in Figure 5 wherein we observe that the *eif2α(2g)-2* allele, and no others, have severely stunted development. Additionally, the *eif2α(2g)-2* protein extract was loaded at 2.5x higher concentration with respect to WT to evenly load rubisco [Figure 7a].

3.2.4 Rosette mRNA expression

When targeting the *eIF2α(5g)* paralog expression in rosette leaves, there was no observed change in mRNA expression of *eIF2α* in the *eIF2α(2g)* mutant alleles and *eif2α(5g)-1* allele. In *eif2α(5g)-2* and *eif2α(5g)-3*, the mRNA expression of *eIF2α* from the Chromosome 5 paralog mirrors the seedling expression data and levels of *eIF2α(5g)* mRNA are reduced to 0.6% and 4.8% compared to WT expression,

respectively [Figure 8]. These data show that *eif2α(5g)-2* and *eif2α(5g)-3* are loss of function alleles for the Chromosome 5 *eIF2α* paralog in mature, rosette-stage plants. The *eIF2α(2g)* targeted reaction again mirrors that of the seedling expression data, in that there is a marked decrease in *eIF2α* expression in the *eif2α(2g)-2* mutant and no others [Figure 8b].

The reaction targeting total *eIF2α* mRNA in the plant demonstrates that the *eIF2α(2g)* paralog is responsible for the majority of *eIF2α* mRNA expression [Figure 8d]. In the rosette stage plant, as in seedlings, the alleles *eif2α(2g)-2, eif2α(5g)-2, eif2α(5g)-3* mutant alleles are loss of function mutations [Figure 8b,c]. The EF1α-controlled results were mimicked in the UBQ-10 controlled experiment as well [Figure 8e,f,g].
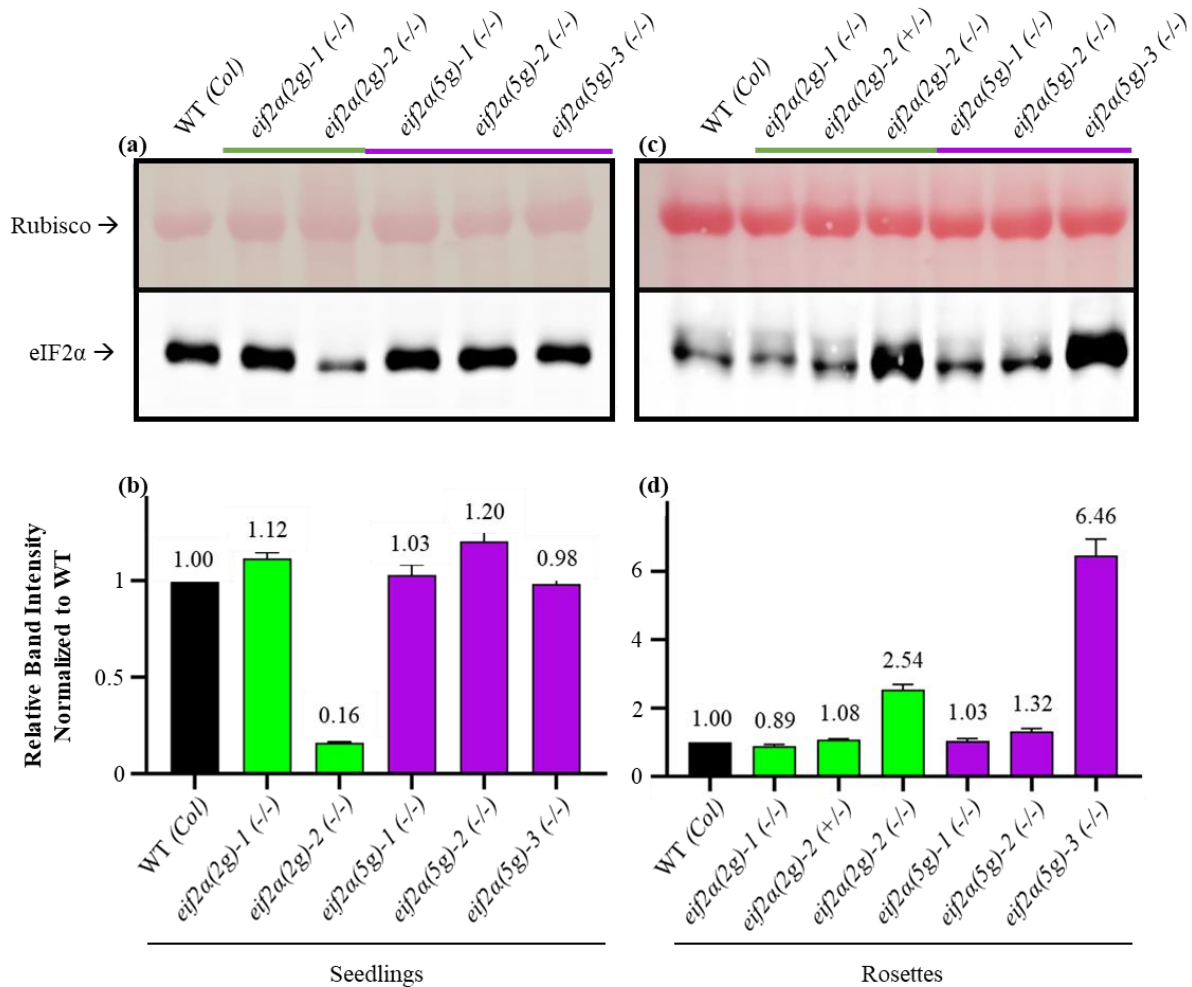
Figure 7 – Quantification of Total Relative *eIF2α* Protein Expression in Arabidopsis *eIF2α* Mutant Seedlings

Figure 7. (a) Arabidopsis seedling protein extracts were separated on a 12% SDS gel. Protein immunoblot was performed with a rabbit anti-eIF2α antibody from the lab of Karen Browning on PVDF membrane with total protein extract from WT (Col), *eif2α(2g)-1 (-/-), eif2α(2g)-2 (-/-), eif2α(5g)-1 (-/-), eif2α(5g)-2 (-/-) and eif2α(5g)-3 (-/-).* The loading control is ponceau-stained rubisco large subunit. The seedling *eif2α(2g)-2* mutant protein extract was loaded at 2.5x total protein concentration with respect to the other samples to load an equal amount of rubisco. The quantified band intensity for *eif2α(2g)-2* was divided by 2.5 to control for the adjusted sample loading. (b) Relative quantification of band intensity normalized by WT. Error bars are Standard error. The (c) Arabidopsis rosette leaf protein extracts were separated on a 12% SDS gel. Protein immunoblot was performed as described in Figure 7a.
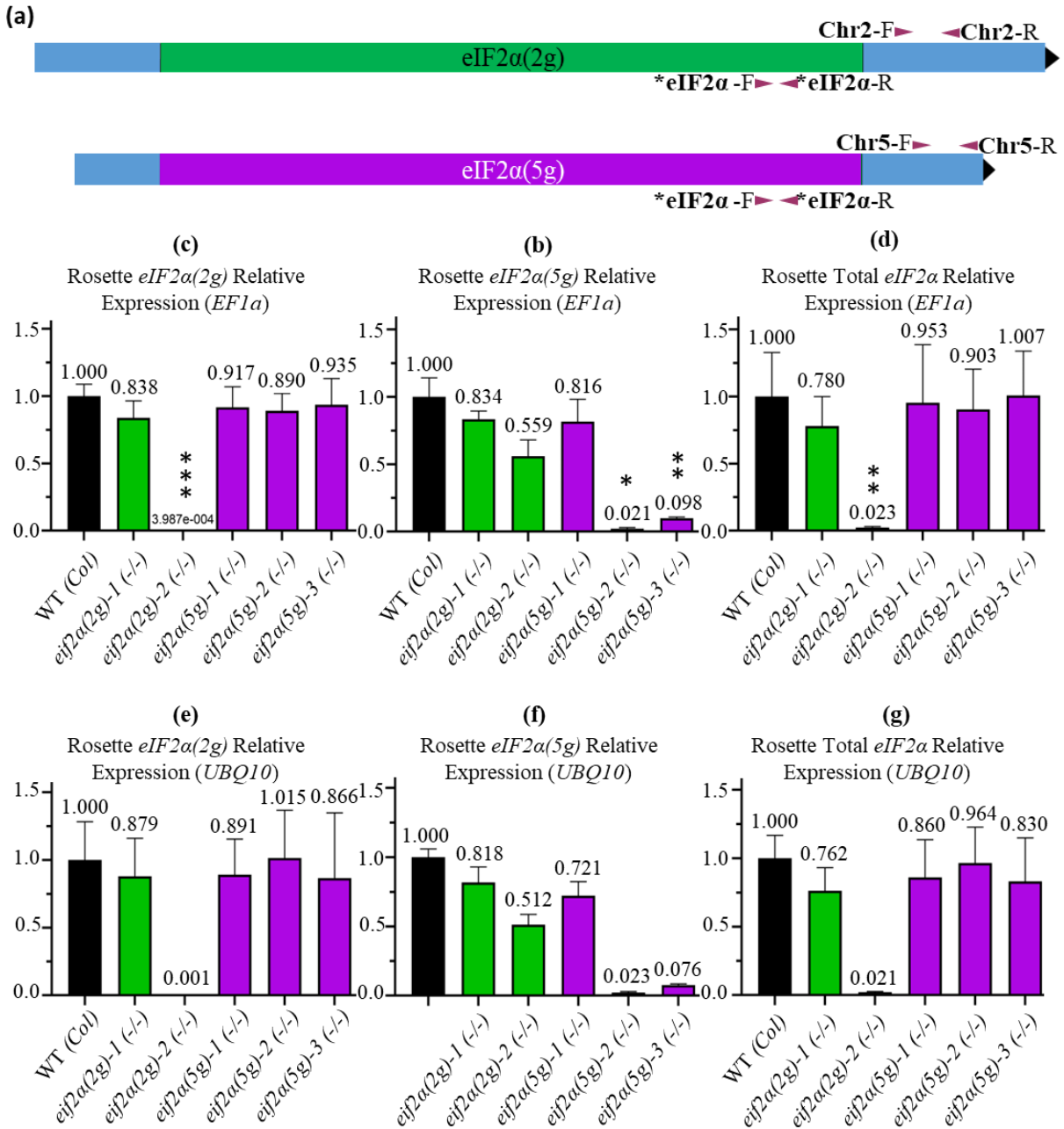
Figure 8 – MRNA Expression Profiles for *eIF2α* Paralogs in Arabidopsis *eIF2α* Mutant Rosette Plants

Figure 8. (a) Shows the relative expression of *eif2α(5g)* mRNA in Arabidopsis rosette leaves for each mutant allele. (b) Demonstrates the relative expression of *eif2α(2g)* mRNA in rosette leaves. (c) Shows the relative expression of total *eIF2α* mRNA in rosette leaves. In all cases, either *EF1α* or *UBQ-10* expression was used as a reference to compare relative *eIF2α* expression and an unequal variance t-test was used. *P < 0.01; **P<0.001; ***P<0.0001.

3.2.5 Rosette protein expression

Total eIF2α protein levels in rosettes were measured by western blot utilizing a rabbit anti-eIF2α antibody in the same manner as the seedling protein expression. In this experiment, we saw that *eif2α(2g)-2* and *eif2α(5g)-3* have greatly increased levels of eIF2α protein expression compared to WT and the *eif2α(2g)-1, eif2α(5g)-1,* and *eif2α(5g)-2* alleles displayed WT expression of eIF2α protein [Figure 7c,d]. This result is surprising since one would not expect such a disparity in the expression of eIF2α protein considering the similarity between the seedling and rosette leaf mRNA expression data [Figure 6; Figure 8; Figure 7]. It is difficult to reconcile the protein expression data of the rosette stage plants with the mRNA expression, however, the result was reproducible and was replicated 4 times. The decreased synthesis of rubisco which is seen in the *eif2α(2g)-2* whole seedling protein extract is not present in rosette leaves protein extract [Figure 7a,b].

3.3 Generation and Genotyping of Plants Harboring *eIF2α* Transgenes in Arabidopsis

The eIF2α protein plays a critical role in the human ISR and yeast GAAC. The core mechanism in the integrated stress response, leading to a change in translation state of eukaryotes is the phospho-regulation of eIF2α. Evidence is emerging that the plant eIF2 mediates a similar stress response [13]. To further address this hypothesis, we introduced an *eIF2α* transgene with mutated phosphoserine 56, the target for the protein kinase GCN2, into the Arabidopsis genome to explore the mechanistic action of this phospho-regulation and its physiological consequences. Thus, generating transgenic lines that are overexpressing wild-type, phospho-deficient (S56A) and phospho-mimetic (S56D) alleles of the *eIF2α* S56 phosphoserine may have consequences on the regulation of endogenous *eIF2α*, the GCN2-mediated ISR, and plant growth and development. One may expect that plants overexpressing S56A could show a phenotype that stems from the lack of a GCN2-mediated stress response while S56D eIF2α may show a constitutive stress response.

Several transgenic lines were generated from agrobacterium strains harboring WT *eIF2α*, phospho-deficient *eIF2α* (S56A) and phospho-mimetic (S56D) *eIF2α* alleles under native and constitutive promoters, respectively. These lines are in a wild-type or an *eIF2α* paralog mutant background with an N-terminal FLAG affinity tag, HA affinity tag, or fluorescence tag [Table 3]. The available transgenic line attributes and the constructs are displayed in Table 3 .

The screening strategy for the *eIF2α* transgenic lines relied on resistance genes and PCR genotyping by taking advantage of the BASTA or hygromycin resistance genes and the unique DNA sequence of the vector DNA. The T_1 seeds were sown onto BASTA or hygromycin agar media and resistant, putative transgenic seedings were selected for growth on soil. In total, 23 sets of 3 independently transformed lines were generated and propagated [Table 3].

Analytical PCR was performed on the putative transgenic mutants to determine the presence or absence of the transgene and endogenous *eIF2α* [Figure 9; Table 4]. Figure 9 shows a representative approach to the PCR genotyping for the sets of transgenic lines. Several of the PCR products that are characteristic of endogenous Arabidopsis paralog genes, endogenous *GCN2*, *eif2α(5g)-2* TDNA, and the *gcn2-1* transposon element are shown in Figure 9a. Figure 9b shows results for a line from construct TG15 and confirms the presence of the endo *eIF2α(2g)* gene and the *eIF2α(2g)* transgene in the wild-type background. Likewise, figure 9c-f are lines with constructs TG21, TG1, TG9 and TG12. Various PCR reactions were performed to confirm the genetic background and the transgene in a specific line. All transgenics in Table 3 were genotyped and the constructs from which they were generated were analyzed by sequencing to confirm promoter structure and phospho S56 alleles before propagation.

Table 3 – List of Transgenic Mutant Constructs Confirmed by Genotyping

Table 3. This table outlines the structure of *eIF2α* transgenes by displaying the *eIF2α* allele, promoter, tag and abbreviated designation used in text. All transgenic lines shown here are stored as T2 seeds.

Transgene	Promoter	Tag	Background	S56	Vector	Selection	Abbr
eif2α(5g) Transgenic Mutants							
eIF2α(5g)	Native	None	*eif2α(5g)-2*	WT	pEG301	BASTA	TG1
eIF2α(5g)	Native	None	*eif2α(5g)-2*	S56A	pEG301	BASTA	TG2
eIF2α(5g)	CaMV 35S	YFP	WT (*Col*)	WT	pEG100	BASTA	TG3
eIF2α(5g)	CaMV 35S	YFP	WT (*Col*)	S56A	pEG100	BASTA	TG4
eIF2α(5g)	CaMV 35S	YFP	WT (*Col*)	S56D	pEG100	BASTA	TG5
eIF2α(5g)	CaMV 35S	YFP	*eif2α(5g)-2*	WT	pEG100	BASTA	TG6
eIF2α(5g)	CaMV 35S	YFP	*eif2α(5g)-2*	S56A	pEG100	BASTA	TG7
eIF2α(5g)	CaMV 35S	YFP	*eif2α(5g)-2*	S56D	pEG100	BASTA	TG8
eIF2α(5g)	CaMV 35S	Flag	WT (*Col*)	WT	pEG100	BASTA	TG9
eIF2α(5g)	CaMV 35S	Flag	WT (*Col*)	S56A	pEG100	BASTA	TG10
eIF2α(5g)	CaMV 35S	Flag	WT (*Col*)	S56D	pEG100	BASTA	TG11
eIF2α(5g)	CaMV 35S	Flag	*eif2α(5g)-2*	WT	pEG100	BASTA	TG12
eIF2α(5g)	CaMV 35S	Flag	*eif2α(5g)-2*	S56A	pEG100	BASTA	TG13
eIF2α(5g)	CaMV 35S	Flag	*eif2α(5g)-2*	S56D	pEG100	BASTA	TG14
eif2α(2g) Transgenic Mutants							
eIF2α(2g)	CaMV 35S	Flag	WT (*Ler*)	WT	pEG100	BASTA	TG15
eIF2α(2g)	CaMV 35S	Flag	WT (*Ler*)	S56A	pEG100	BASTA	TG16
eIF2α(2g)	CaMV 35S	Flag	WT (*Ler*)	S56D	pEG100	BASTA	TG17
eIF2α(2g)	CaMV 35S	HA	WT (*Col*)	WT	pEG100	BASTA	TG18
eIF2α(2g)	CaMV 35S	HA	WT (*Col*)	S56A	pEG100	BASTA	TG19
eIF2α(2g)	CaMV 35S	HA	WT (*Col*)	S56D	pEG100	BASTA	TG20
eIF2α(2g)	CaMV 35S	HA	*gcn2-1 (-/-)*	WT	pEG100	BASTA	TG21
eIF2α(2g)	CaMV 35S	HA	*gcn2-1 (-/-)*	S56A	pEG100	BASTA	TG22
eIF2α(2g)	CaMV 35S	HA	*gcn2-1 (-/-)*	S56D	pEG100	BASTA	TG23

3.3.1 Transgenic mutant seedling growth and development

First, I examined the effect of overexpressing the minor paralog, *eIF2α(5g)*, in the wild-type or *eif2α(5g)-2* mutant background. A root length assay was performed, comparing the growth of the transgenic seedlings expressing three Flag-tagged *eIF2α(5g)* alleles under the CaMV 35S promoter [Figure 10b]. In the *eif2α(5g)-2* background (lines TG12, TG13, TG14), roots grew slightly longer in the presence of the transgene. This may suggest that eIF2α levels are mildly rate-limiting for root growth and that expression of additional eIF2α protein from the 35S promoter reduces this limitation. A similar effect was seen in the wild-type background (lines TG9, TG10, TG11), except for the S56D isoform. This exception may be due to insufficient expression of this transgene in the chosen line and was not investigated further.

The wild type and S56A isoforms of *eIF2α(5g)* were also introduced into the *eif2α(5g)-2* mutant background under their own native promoter. These lines (TG1 and TG2) recapitulated the slight increase in root length seen in the CaMV 35S driven lines [Figure 10a].

Figure 9 – Representative Genotyping of Plants Harboring *eIF2α* Transgenes

Figure 9. 1% Agarose gels showing results of analytical PCR to determine the genotype of transgenic plants. In the interest of concision, five genotyping results of transgenic plants representative of all the lines in Table 3 are shown here. Primer names and primer targets referenced here are also referenced in the methods section 2.8.2. **(a)** This panel shows a series of positive controls for the PCR reactions used to genotype transgenic plants. Note that in the endogenous *GCN2* reaction, there is a large primer dimer present on the gel. **(b)** This panel displays the representative genotyping result from the pEG100-eIF2α(2g) lines. **(c)** This panel is the representative genotyping result from the pEG100 eIF2α(2g) *gcn2-1(-/-)* lines. Shown here, from left to right, is the presence of the endogenous *eIF2α(2g)* gene amplicon, the presence of the pEG100-35S-eIF2α(2g) transgene (TG) amplicon, the absence of the endogenous *GCN2* gene amplicon, and the presence of the *gcn2-1* mutant allele transposon element amplicon. **(d)** This panel shows the genotyping results for the pEG301-Native-eIF2α(5g) *eif2α(5g)-2 (-/-)*. **(e)** This panel displays the representative genotyping data for the pEG100-eIF2α(5g) lines. **(f)** This panel displays the representative genotyping data for the pEG100-eIF2α(5g) *eif2α(5g)-2 (-/-)* lines.

(a) WT (Col) / eif2α(5g)-2 (-/-) gcn2-1 (-/-)

Targets of PCR (a): Endo eIF2α(2g), Endo eIF2α(5g), Endo GCN2, eif2α(5g)-2 TDNA, gcn2-1 transposon

(b) eIF2α(2g) Wild-type

Targets of PCR (b): Endo eIF2α(2g), gcn2-1 transposon, eIF2α(2g) TG

(c) eIF2α(2g) gcn2-1 (-/-) ← Transgene / ← Background

Targets of PCR (c): Endo eIF2α(2g), eIF2α(2g) TG, Endo GCN2, gcn2-1 transposon — Target of PCR

(d) Nat eIF2α(5g) eif2α(5g)-2 (-/-)

Targets of PCR (d): Endo eIF2α(5g), eif2α(5g)-2 TDNA, Nat eIF2α(5g) TG

(e) eIF2α(5g) Wild-type

Targets of PCR (e): Endo eIF2α(5g), eif2α(5g)-2 TDNA, eIF2α(5g) TG

(f) eIF2α(5g) eif2α(5g)-2 (-/-) ← Transgene / ← Background

Targets of PCR (f): Endo eIF2α(5g), eif2α(5g)-2 TDNA, eIF2α(5g) TG — Target of PCR
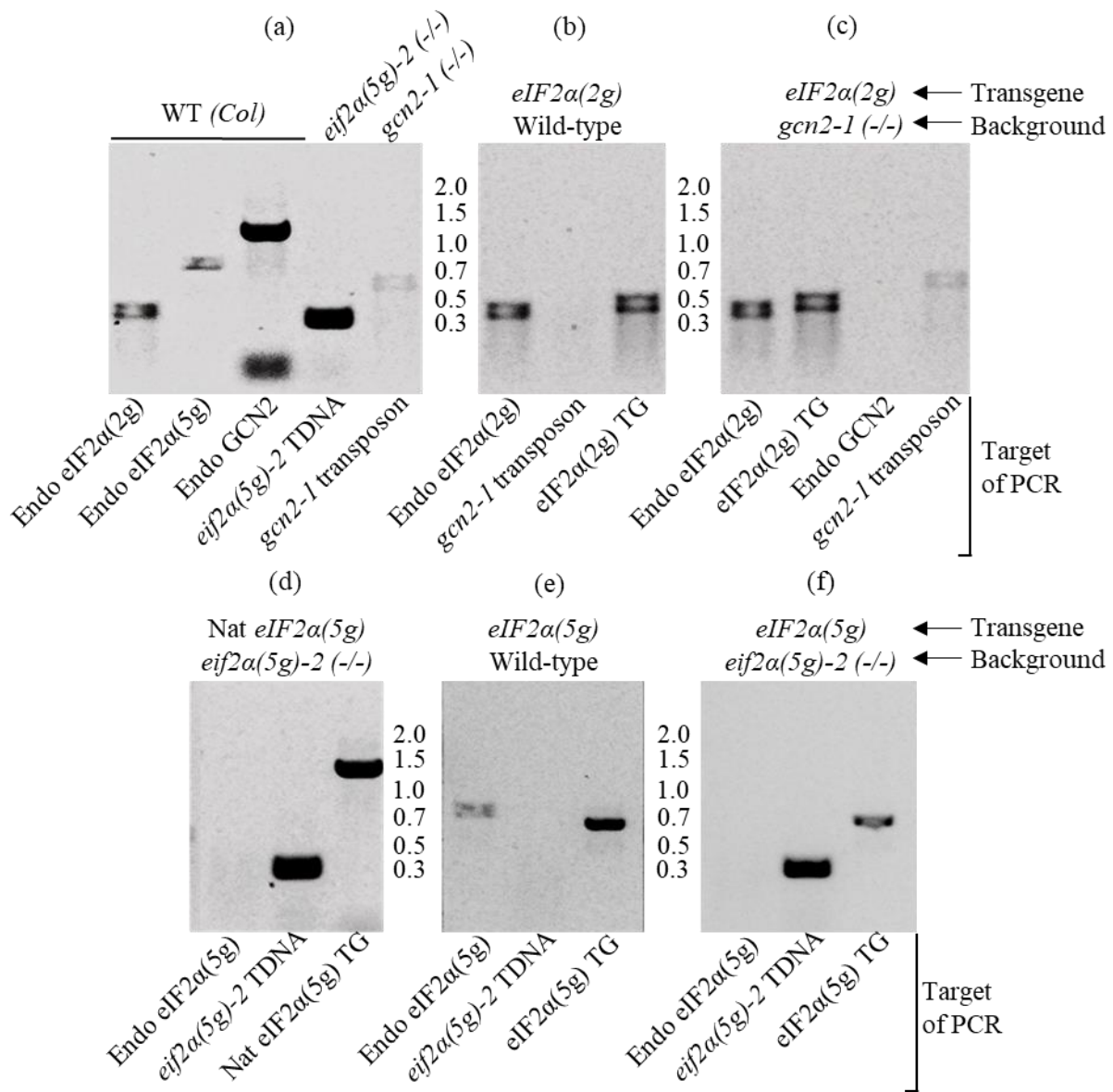
44

Table 4 – Transgenic Line Genotyping Primer Pair Targets and Amplicon Lengths

Table 4. This table displays PCR primer pair targets referenced in Figure 9 and the names of the corresponding primer pairs as their name appears in the methods section and in materials labeled in the von Arnim lab. Additionally, the length in base pairs of the resulting amplicon is displayed.

Primer Target	Amplicon length (bp)	Primer pair
Endogenous *eIF2α(2g)* gene	348	gAT2-eIF2α-intron5-F + qPCR AT2 5'UTR-R
Endogenous *eIF2α(5g)* gene	718	AT5-5'INTeIF2α-F + AT5-eIF2α-mid-R2
Endogenous *GCN2* gene	1200	SKC-049 + SKC-048
eif2α(5g)-2 TDNA	320	AT5-eIF2α-mid-F2 + LBP1.3ext
gcn2-1 transposon element	600	SKC-049 + DS3-2
pEG100-eif2α(2g) transgene	466	qPCR-AT2-CDSeIF2α-F + JucMarkSEQ-R
pEG301-Native-eif2α(5g) transgene	1252	AT5-eIF2α-mid-F2 + JucMarkSEQ-R
pEG100-eif2α(5g) transgene	630	AT5-eIF2α-mid-F2 + JucMarkSEQ-R

In all, the effect of transgenic *eIF2α* expression in seedlings was mild. The root length assay demonstrated that the transgenic lines generally grow slightly faster than wild-type. When the growth between transgenic lines was compared, they generally grew at the same rate regardless of their promoter, S56 allele, or N-terminal tag. Additionally, all the transgenic seedlings appeared identical to wild-type in morphology and germinated at the same rate as wild-type. There was essentially only a nominal effect of the *eIF2α* transgene expression on seedling growth and development. The prediction that S56D would reveal a severe growth defect because of a constitutive ISR-effect did not materialize. Likewise, no difference was detected between WT and S56A alleles. This may suggest that the S56A is functional for eIF2 activity and that our growth conditions do not trigger an ISR, in keeping with low levels of eIF2-P under normal growth conditions. Transgenic lines expressing YFP-tagged *eIF2α(5g)* were also generated but were not examined at the seedling stage. It would have been interesting to examine the transgenes for their ability to rescue the mutant phenotypes for *eIF2α(2g)*, which are much more severe than those for *eIF2α(5g)*. However, no such lines were generated.

3.3.2 Adult plant morphology

The expression of eIF2α protein must be tightly regulated and balanced, as *eIF2α* is expressed ubiquitously in the plant and initiating translation is of core biological importance. The proper growth of shoots, leaves and reproductive structures is inherently related to the diverse, differentiating meristems in plants. Typically, as the shoot apical meristem (SAM) drives shoot elongation, there are other meristems that differentiate into critical plant structures such as axillary node branches and inflorescences. It has already been observed that lower expression of *eIF2α* in the *eif2α(2g)-2* mutants results in infertile, stunted plants [Figure 5]. Conversely, the overexpression of *eIF2α* transgenes may result in developmental defects in vegetative and reproductive structures due to mis- or dysregulation of SAMs and inflorescence meristems (IMs) which give rise to floral meristems (FMs). To address this hypothesis, WT and *eif2α(5g)-2* mutant background seedlings that harbor natively or constitutively expressed *eIF2α(5g)* transgenes with varying

S56 alleles were germinated on agar and transferred to soil after 7 days for screening of growth phenotypes [Table 5].

The data in Table 5 catalogues the observed phenotypes and the frequency at which they occur in each T_3 transgenic strain at 12 and 33 days of growth. The effect of the transgene varied from plant to plant but there was consistency in the type of phenotypes shown, including bunching of developing siliques due to failure of stem elongation, bent or kinked stems, jagged or rough leaf edges, twisted leaf morphology, silique dysmorphia, infertility, and plants with only a single shoot affected while the rosette and other shoots appeared as wild-type [Table 5]. In all of the transgenic lines assayed in Table 5, the seedlings appeared identical to WT in morphology. The observed morphological phenotypes in Table 5 were all accompanied by expression of anthocyanins and in some cases, plants succumbed to the mis-expression of *eIF2α* while displaying extreme levels of anthocyanin [Figure 11]. Any plant that displayed a morphological phenotype had first expressed anthocyanin and either halted further development [Figure 11a] or continued to develop [Figure 11b,c]. All plants that expired on soil did not mature past the early rosette stage and were expressing very high levels of anthocyanin [Figure 11a; Table 5]. Additionally, transgenic mutants such as TG11 in figure 11c showed normal growth and development of the primary shoot, but the rosette leaves began expressing high levels of anthocyanins concomitant with halting of secondary shoot elongation and infertile, flowering structures. The stunted siliques fail to separate in the canonical *pole-ladder* fashion seen on the primary shoot.

Figure 11d, e and f show a plant from set TG10 and represent another typical pleiotropic phenotype. Many rosette leaves and cauline leaves are notched at the edges and asymmetric, the shoots fail to elongate, causing the cauline leaves and flowers to remain bunched together. In the flowers, the petals remain underdeveloped, thus exposing the stigmas. Flowers remain infertile and no seeds develop. Though the phenotypes of the plants are similar between lines, the TG10 and TG11 phenotypes in Figure 11 represent the upper and lower extremes of overexpressing transgenic plant phenotypes.
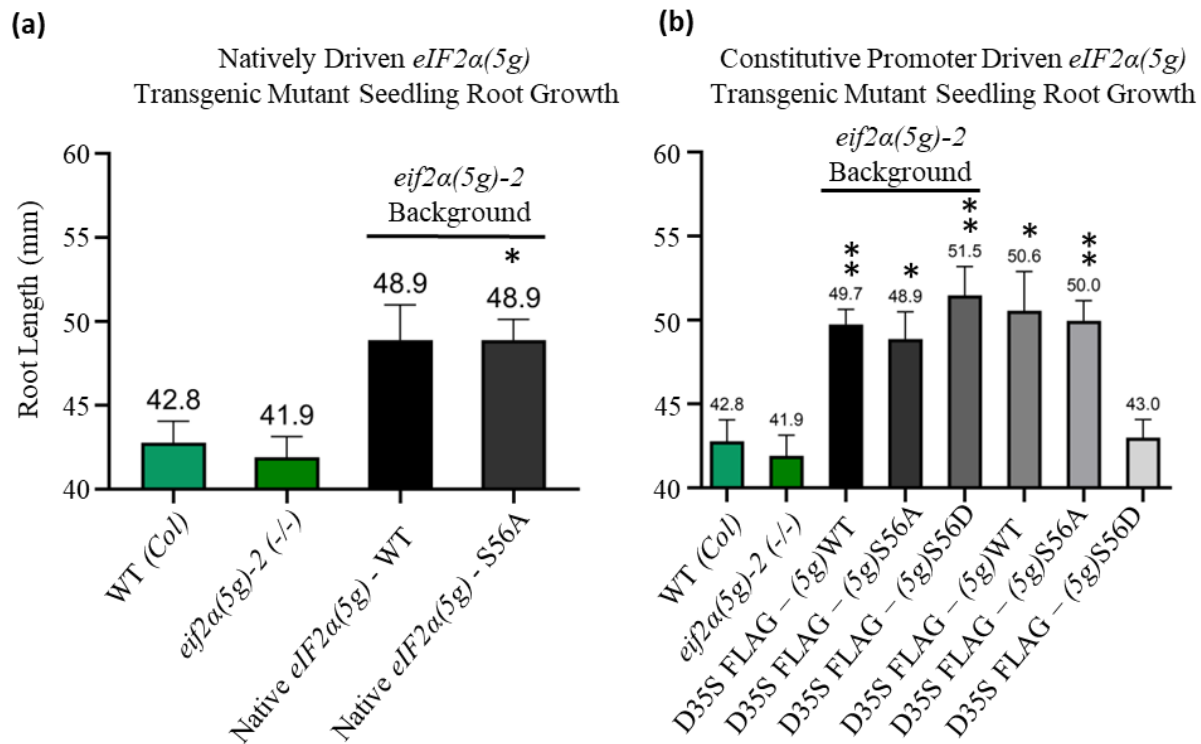
Figure 10 – Root Growth in Transgenic Mutants Expressing *eIF2α(5g)* under Native and Constitutive Promoters

Figure 10. Quantification of transgenic seedling root length assay results after 10 days of growth on ½ MS + 0.5% sucrose media for (a) WT(Col), *eif2α(5g)-2*, TG1 and TG2 *eIF2α(5g)* complementation lines. In panel (b), growth of WT(Col), eif2α(5g)-2, TG12, TG13, TG14, TG9, TG10 and TG11. *P < 0.01; **P<0.001; ***P<0.0001.

Altogether, plants expressing *eIF2α* transgenes under the native promoter, regardless of the S56 allele, developed like wild-type plants with no abnormal morphological phenotypes besides longer root lengths at the seedlings stage [Figure 10; Table 5]. However, all constitutively promoted *eIF2α* transgenic lines showed a subset of siblings (about 8%) with abundant anthocyanins and morphological phenotypes that are caused by an apparent developmental dysregulation of the SAM, axillary meristems (AM), IMs, FMs and leaf growing edges [Table 5; Figure 11]. The plants that presented phenotypes had a 16% rate of death following transgene introduction, 8.3% in the FLAG-tagged lines and 23.1% in the YFP-tagged lines [Table 5]. However, the disparity in the rate of abnormal phenotypes, including death between the lines is likely due to the small sample size.

In summary, the overexpression constructs had a tendency to generate pleiotropic developmental phenotypes at the adult stage of development. This type of phenotype is not typical for Arabidopsis lines generated with 35S promoters or FLAG or YFP tags and must therefore be due to the *eIF2α* coding region. It is possible that some transgenic lines expressed a partial eIF2α fragment that may interfere with eIF2 protein function, but the frequency of aberrant phenotypes makes this seem unlikely. The variability

Table 5 – Tabulated and Qualitative Analysis of Plants Harboring *eIF2α(5g)* Transgenes

Table 5. This table displays the phenotypes of the eIF2α transgenic mutants grown on soil and the rate of occurrence of the phenotypes at 12 days and 33 days on soil. The mutant column shows the construct and background being assayed. The N column shows the sample size for the given mutant data. The % Phenotype columns display how many plants out of the total sample size that show a phenotype differing from WT plants. Terminal phenotype column shows phenotypes present at the point which plants displayed a lethal phenotype or at 33 days of growth on soil when most plants began to senesce— (A+), extreme levels of anthocyanin being expressed and severely stunted growth; A, anthocyanin is being expressed; BS, siliques are bunched due to failure of SAP to elongate; KS, stalk is jagged and kinked; LE, jagged or rough rosette leaf edges; SA, a single shoot is affected; TL, rosette or axillary leaves appear twisted; SD, silique dysmorphia. The % lethal column shows how many plants did not survive on soil by day 33.

Mutant	N	%Phenotype 12 days	%Phenotype 33 days	Terminal Phenotypes	% Lethal
Native-eIF2α(5g)-WT *eif2α(5g)-2*	18	0 % (0/18)	0 % (0/18)	None	0
Native-eIF2α(5g)-S56A *eif2α(5g)-2*	18	0 % (0/18)	0 % (0/18)	None	0
35S-YFP-eIF2α(5g)-WT Wild-type	18	5.6 % (1/18)	5.6 % (1/18)	A, BS, KS, LE, SD, SA	0 (0/1)
35S-YFP-eIF2α(5g)-S56A Wild-type	20	5 % (1/20)	5 % (1/20)	(A+)	100 (1/1)
35S-YFP-eIF2α(5g)-S56D Wild-type	36	16.7 % (3/36)	22.2 % (4/36)	(A+), A, KS, LE, SD, TL	25 (1/4)
35S-YFP-eIF2α(5g)-WT *eif2α(5g)-2*	36	2.8 % (1/36)	16.7 % (3/36)	A, BS, KS, LE, SS, SD	0 (0/3)
35S-YFP-eIF2α(5g)-S56A *eif2α(5g)-2*	17	0 % (0/17)	5.9 % (1/17)	A, BS, KS, SS, TL, SA	0 (0/1)
35S-YFP-eIF2α(5g)-S56D *eif2α(5g)-2*	36	0 % (1/18)	8.3 % (3/18)	(A+), A, BS, LE, KS, SS	33.3 (1/3)
35S-FLAG-eIF2α(5g)-WT Wild-type	18	5.6 % (1/18)	11.1 % (2/18)	A, BS, SS, SD	0 (0/2)
35S-FLAG-eIF2α(5g)-S56A Wild-type	17	0 % (0/17)	17.6 % (3/17)	(A+), A, BS, LE, SD, TL	33.3 (1/3)
35S-FLAG-eIF2α(5g)-S56D Wild-type	16	6.3 % (1/16)	6.3 % (1/16)	A, KS, LE, SD, TL, SA	0 (0/1)
35S-FLAG-eIF2α(5g)-WT *eif2α(5g)-2*	14	0 % (0/14)	7.1 % (1/14)	A, SS, SD, TL	0 (0/1)
35S-FLAG-eIF2α(5g)-S56A *eif2α(5g)-2*	18	5.6 % (1/18)	11.1 % (2/18)	A, BS, KS, LE, SS, TL, SD	0 (0/2)
35S-FLAG-eIF2α(5g)-S56D *eif2α(5g)-2*	33	3 % (1/33)	9.1 % (3/33)	A, BS, KS, TL, SD, SA	0 (0/3)
YFP-tagged mutants	**163**	**4.3 % (7/163)**	**8.0 % (13/163)**	–	**23.1 (3/13)**
Flag-tagged mutants	**116**	**3.4 % (4/116)**	**10.3 % (12/116)**	–	**8.3 (1/12)**
Total	**315**	**3.5 % (11/315)**	**7.9 % (25/315)**	–	**16 % (4/25)**

between siblings from the same parent plant is most reminiscent of epigenetic variation in gene expression, which is common with Arabidopsis transgenes. It is most interesting to speculate whether this phenotype is due to ectopic expression of *eIF2α* in the wrong cell type, due to overexpression, or due to under-expression from RNA interference. Under-expression of the *eIF2α(5g)* gene alone would most likely have no consequence, while under-expression of *eIF2α(2g)* would be expected to cause the small size and infertility seen in the *eif2α(2g)-2* mutants. However, none of the transgenic plants showed the characteristic *eif2α(2g)-2* phenotype of delayed growth, small size, pale green pigmentation, and complete infertility. This tends to argue against an RNA interference effect and leaves overexpression or ectopic expression as more likely explanations. Because eIF2 is thought to be expressed ubiquitously, ectopic expression is unlikely to occur. This leaves overexpression or perhaps a combination of uniform overexpression and stochastic RNA interference as a possible basis for the abnormal phenotypes. The expression level of the transgenes was not examined.
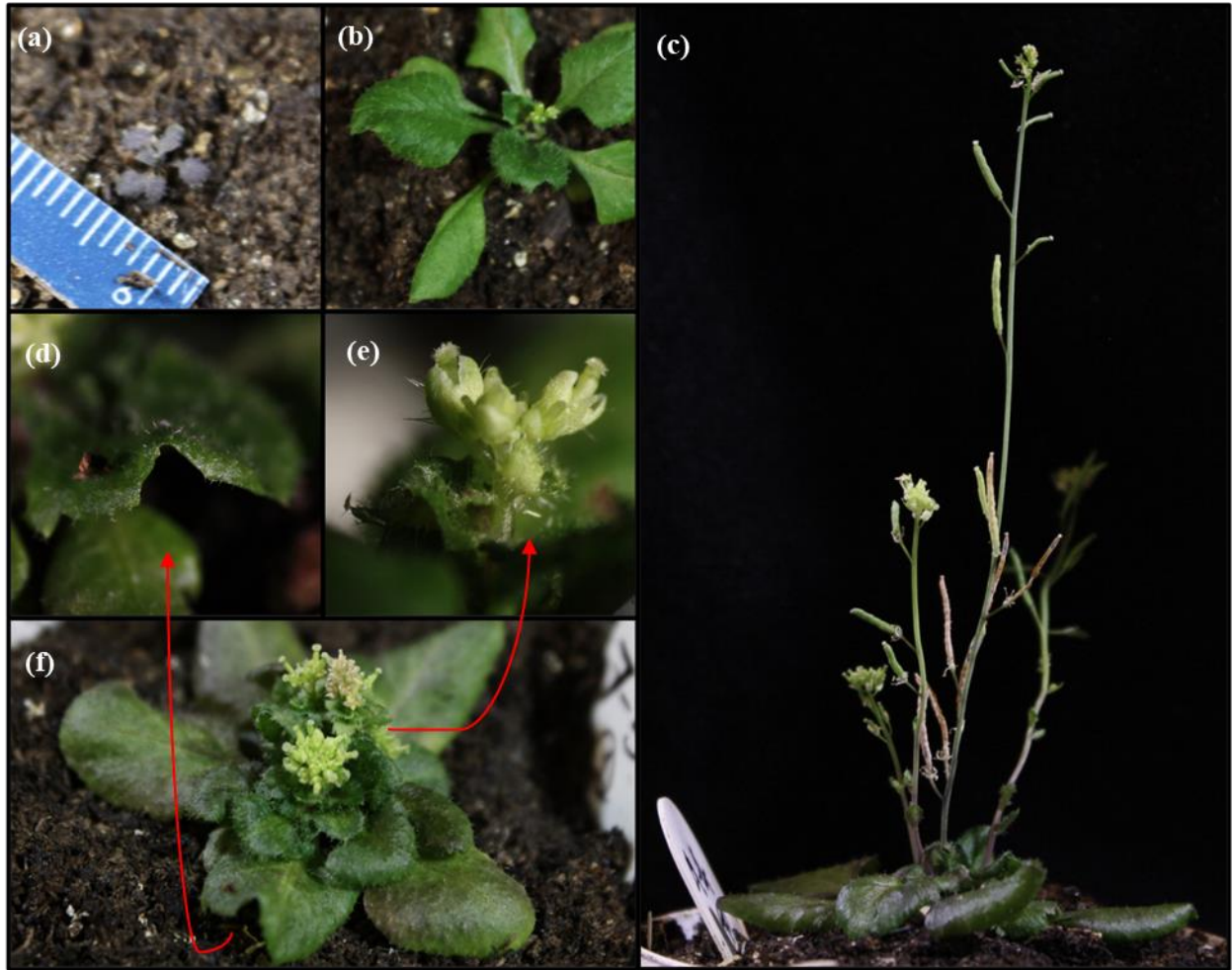
Figure 11 – Representative Phenotypes of Transgenic Plants

Figure 11. (a) Shows a plant expressing the 35S *eIF2α(5g)*-S56D transgene in a wild-type background at 24 total days of growth, 14 days on soil. (b) Shows a plant expressing the 35S *eIF2α(5g)*-S56D in a wild-type background at 24 days growth, 14 days on soil. This panel shows that the 3 newest rosette leaves are greener than the others, indicating increased levels of anthocyanins. (c) Shows the plant from panel B at 50 days growth, 40 days on soil. (d-f) Show a plant expressing the 35S *eif2α(5g)*-(S56A) transgene in a wild-type background. Panels D and E are magnified structures from panel F.

3.4 Expression of recombinant ISR-relevant proteins: eIF2α, ABCF1, GCN2-KD

In this project, three versions of *eIF2α*, wild-type, phospho-deficient (S56A), and phospho-mimetic (S56D) alleles, in addition to the kinase domain of *GCN2*, two regulators of translation in the ISR pathway were cloned and expressed in *E. coli* and purified by 6xHIS affinity tag. Additionally, recombinant *ABCF1*, which is one potential interactor of GCN1 was expressed and purified by use of a Glutathione S-Transferase (GST) affinity tag.

3.4.1 Expression profile

An expression profile experiment was performed to determine optimal expression conditions for recombinant *ABCF1* and *GCN2-KD*. Expression of *eIF2α* was already optimized by a previous post-doc, Ansul Lokdarshi, when this project began. Antibodies against the 6xHisTag and GST tags of recombinant *GCN2-KD* and *ABCF1*, respectively, were used to measure recombinant protein expression.

For recombinant *GCN2-KD* expression, recovery in the soluble fraction was best when expression was induced with 1mM IPTG for 24 hours at 12 ℃. Other conditions: (i) 1mM IPTG for 3 hours at 37 ℃, (ii) 1mM IPTG induction for 18 hours at 22 ℃, and a control condition (iii) with no IPTG induction for 24 hours at 12 ℃. Interestingly, there is a significant amount of protein in the insoluble fraction in all other conditions except the uninduced control [Figure 12a].

ABCF1 expression was assayed in the same manner using an anti-GST antibody in growth conditions: (i) 0.5 mM IPTG induction for 18 hours at 22 ℃, (ii) 1mM IPTG induction for 3 hours at 37 ℃, (iii) 1mM IPTG induction for 18 hours at 22 ℃, and (iv) 0.5mM IPTG induction for 24 hours at 12 ℃. *GST-ABCF1* was most highly expressed in the soluble fraction at 1 mM IPTG induction for 18 hours at 22 ℃ condition [Figure 12b].
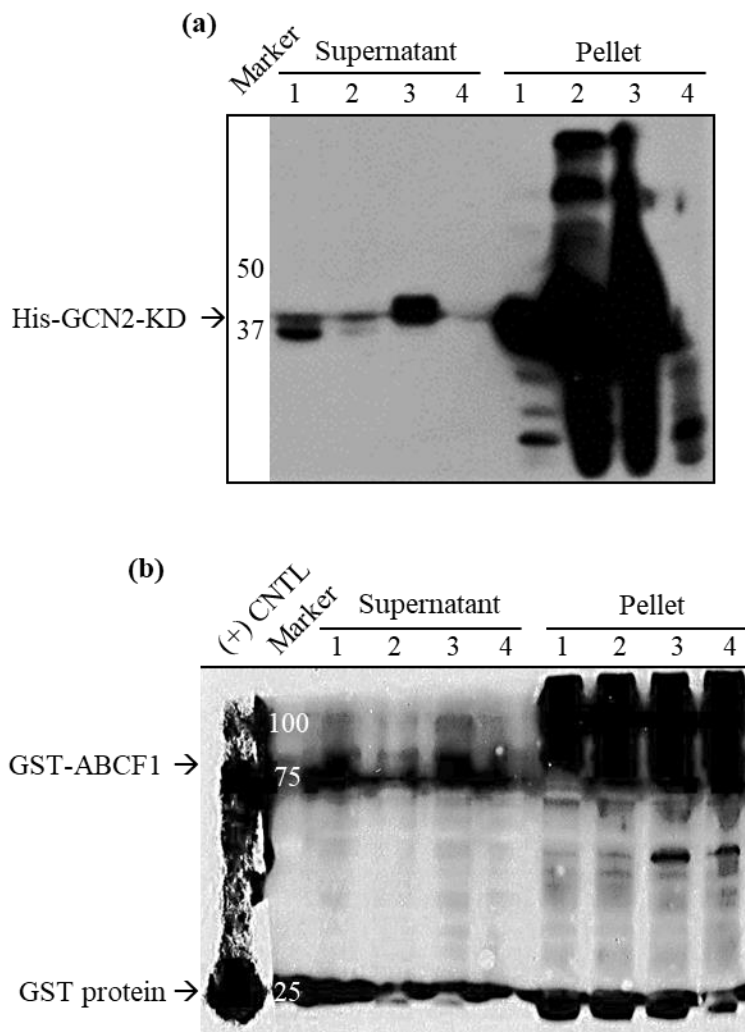
Figure 12 – *E. coli* Expression of Recombinant GCN2-KD and ABCF1

Figure 12. (a) This panel displays the expression profile for recombinant GCN2-KD from the soluble protein fraction and the pellet fraction in condition 1 (1 mM IPTG induction for 3 hours at 37 °C), condition 2 (1 mM IPTG induction for 18 hours at 22 °C), condition 3 (1 mM IPTG induction for 24 hours at 12 °C), condition 4 (un-induced for 24 hours at 12 °C). (b) This panel displays the expression profile for recombinant ABCF1 from the soluble protein fraction and the pellet fraction in condition 1 (0.5mM IPTG induction for 18 hours at 22 °C), condition 2 (1 mM IPTG induction for 3 hours at 37 °C), condition 3 (1 mM IPTG induction for 18 hours at 22 °C), condition 4 (1 mM IPTG induction for 24 hours at 12 °C). The positive control lane contains purified GST-ABCF1 mixed with GST-protein. The Bio-Rad precision plus pre-stained ladder (Bio Rad cat#1610374) is used for the panel a and panel b blots.

54

3.4.2 Purification

The recombinant proteins were purified utilizing column chromatography according to their respective affinity tags. The recombinant His-tagged eIF2α WT, S56A, and S56D allele proteins were purified in a Ni-NTA resin column using imidazole as an eluant. In Figure 13a, it is demonstrated the eIF2α protein was expressed and purified out of *E. coli* BL21-STAR cells. The protein is purified out with a 110kDa contaminant protein. In attempts to remove the contaminant, expression, and purification with varying inductions at 0.25mM and 0.5mM IPTG, extended length of induction at lower temperatures, induction at OD_{600} 0.4, and incubating cell lysate with Mg·ATP before and during resin binding were all performed but to little effect. To address the possibility that the contaminant could be an oligomer or tightly bound to eIF2α, samples were incubated with 4M Urea, only to reveal the same result as Figure 13a. Additionally, anion exchange HPLC purification was performed on Ni-NTA-purified recombinant eIF2α. Comparing panel 13a and 13b shows that the anion exchange method further purified the recombinant eIF2α protein, yet the primary contaminant band remained in each fraction [Figure 13b].

The purified fractions of recombinant ABCF1 are shown in Figure 13c. In each fraction, the desired 92kDA GST-ABCF1 protein is visible in the eluate, along with shorter fragments that I interpret as cleaved 67kDa ABCF1, 26 kDa GST and their degraded fragments [Figure 13c].

The recombinant GCN2-KD protein was purified in an Ni-NTA resin column using imidazole as an eluent. The purified GCN2-KD is shown in Figure 13d. The recombinant GCN2-KD protein was successfully expressed and purified; however, it is still unknown whether this truncated version of GCN2 is capable of kinase activity.
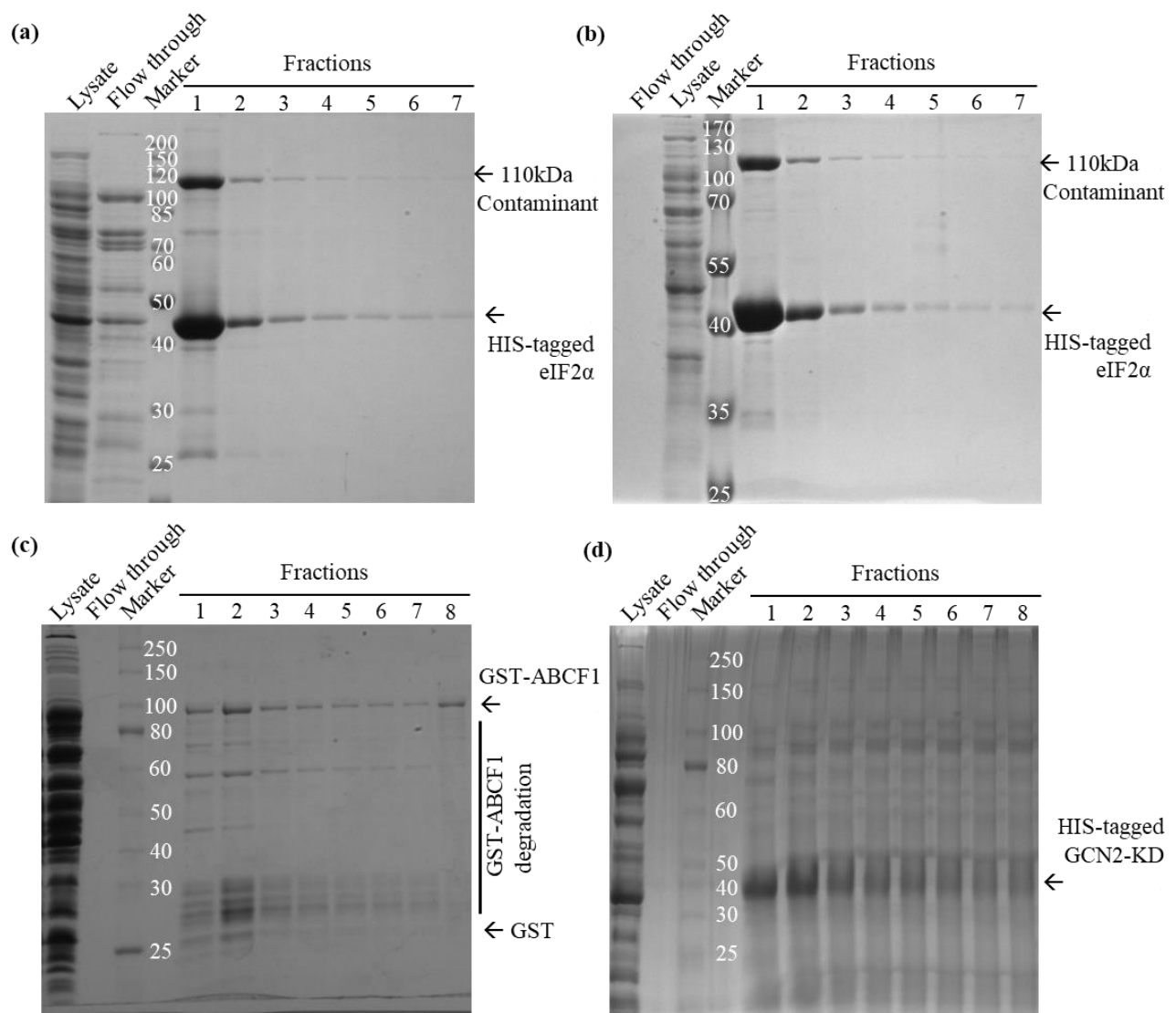
Figure 13 – Purification of Recombinant Proteins

Figure 13. This figure shows the purified fractions of recombinant eIF2α after (a) His-tag column purification [marker was Thermo Pageruler 26614] and (b) additional HPLC-anion exchange [marker was Thermo Pageruler 26617] (c) This panel shows the GST-ABCF1 purification using glutathione resin as a GST binding agent and reduced glutathione as an eluant [NEB P7703]. The higher concentration and greater purity in the F8 elution was due to a longer eluent incubation of 20 minutes after most contaminants had been eluted in earlier fractions. (d) This panel shows the purification fractions of GCN2-KD [NEB P7703].

Chapter 4.
Discussion and Conclusions

In this work, I characterize the growth and development of *eIF2α* mutant plants and the expression of *eIF2α* at the mRNA and protein level of several *eIF2α* single-paralog mutants. Results from the RT-qPCR experiment demonstrates that of the two *eIF2α(2g)* mutant alleles, *eif2α(2g)-2* is an *eIF2α(2g)*-null allele and of the three *eIF2α(5g)* mutant alleles, *eif2α(5g)-2* and *eif2α(5g)-3* are *eIF2α(5g)*-null alleles [Figure 6; Figure 8]. The *eif2α(2g)-1* and *eif2α(5g)-1* mutants show no defect in expression of *eIF2α* mRNA or protein [Figure 6; Figure 7; Figure 8].

The *eIF2α(2g)* gene provides the majority of *eIF2α* mRNA and protein in wild-type plants. This can be concluded because the *eIF2α(2g)* is responsible for expression of 96% of *eIF2α* transcripts in Arabidopsis seedlings as evidenced by the *eif2α(2g)-2* loss of function mutant in the RT-qPCR experiment [Figure 6d]. The RT-qPCR observation is also consistent with the stunted plant development and low protein expression of the *eif2α(2g)-2* mutant, alone [Figure 5; Figure 6]. As the plant develops into the rosette stage, the RT-qPCR data reveals similar results as during the seedling stage. The *eif2α(5g)-2 and eif2α(5g)-3* mutants are both loss of function mutations, however, the *eif2α(5g)-2 and eif2α(5g)-3* mutants are similar to wild-type in total protein expression and indistinguishable as developing seedlings [Figure 5; Figure 7].

Although *eif2α(2g)-2* is an mRNA-null allele, the seedlings do express some eIF2α protein which must come from *eIF2α(5g)* expression. Additionally, though the *eIF2α(5g)* gene does not express high levels of *eIF2α* mRNA, protein from the *eIF2α(5g)* gene can accumulate to levels detectable by western blot. More strikingly, the rosette-stage *eif2α(2g)-2* plants are expressing excess eIF2α protein. Interestingly, however, the *eif2α(2g)-2* and *eif2α(5g)-3* mutants are expressing 2.9 and 5.3-fold more eIF2α protein than wild-type, respectively. It is more puzzling that the *eif2α(2g)-2* mutant, which is expressing 2.3% of the wild-type plant mRNA levels, is expressing eIF2α protein at roughly three times the level of wild-type. Perhaps the reduced rate of eIF2α protein synthesis in the *eif2α(2g)-2* mutant results in an altered rate of

degradation for the *eif2α(5g)* paralog or the translational product of the 6th intron TDNA insertion mutants somehow evades degradation. Another possibility is that a truncated 6th intron transcript results in a truncated eIF2α protein which lacks its native C-terminal peptide sequence, resulting in evasion of C-terminal recognition degron pathways. The potential for C-terminal degron evasion in the *eif2α(2g)-2* and *eif2α(5g)-3* line could offer an alternative explanation to the staggering degree of compensation coming from the WT paralog in each mutant.

Unfortunately, there is not a unique epitope target for an eIF2α antibody to differentiate *eIF2α(2g)* and *eIF2α(5g)* protein. However, if both *eif2α(2g)-2 and eif2α(5g)-3* are loss of function alleles as evidenced by the RT-qPCR experiment, then we can infer that the unaffected paralog in each mutant must be responsible for the expression of total eIF2α protein. It is difficult to determine why there is a growth phenotype in the *eif2α(2g)-2* when there is apparent compensatory expression from the *eIF2α(5g)* allele. One could speculate that the *eIF2α(5g)* protein might not be fully functional or it may not be expressed highly enough in cell types where protein levels of *eIF2α* are rate limiting. Another explanation could be that the lack of eIF2α protein in the early stages of seedling development results in lasting developmental consequences that prohibit normal development by the time the *eIF2α(5g)* paralog can compensate for *eIF2α(2g)* loss of function.

The question remains: Is the elevated level of eIF2α protein due to compensatory translation from the poorly expressed *eIF2α(5g)* mRNA, or due to a change in the rate of degradation of translated eIF2α? We can infer from these results that the expression of *eIF2α(5g)* protein is carefully regulated depending on the status of the *eIF2α(2g)* gene. Moreover, it is interesting that a similar compensatory effect was seen in the *eif2α(5g)-3* mutant background, where excess *eIF2α* protein is present, possibly expressed from the *eIF2α(2g)* gene. This is striking because no such effect was seen in the other null allele for *eIF2α(5g)*, *eif2α(5g)-2*, and there are no striking phenotypes in any of the *eIF2α(5g)* mutants. These data suggest that the expression of eIF2α paralogs is carefully calibrated by a form of allele-specific cross-regulation, which may occur at the level of mRNA translation or protein degradation. The mechanism for this is unclear. It is

noteworthy that both mutants for which there is excess eIF2α protein have a TDNA insertion that interrupts the 6th intron.

Wild-type, *eif2α(5g)-2*, and *gcn2-1* background plants that harbor an *eIF2α* transgene under various promoters and biochemical tags were cloned, screened, and propagated. The pEG100 constructs harboring *eIF2α(2g)* transgenes were archived for future work and not included in the growth or morphological analyses shown in this thesis. The transgenic mutants harboring an *eIF2α(5g)* transgene were characterized according to their root length, growth, and morphology. All seedlings with an *eIF2α(5g)* transgene under a native promoter grew as wild-type or outgrew the wild-type in a root length assay, and they appeared to be normally developing seedlings. However, it was observed that the pEG100 constructs that constitutively express an *eIF2α(5g)* transgene from the 35S promoter could over-accumulate anthocyanin and present many growth defects in stochastic fashion during growth on soil— defects in SAM derived organs, affecting phyllotaxy of floral initiation, stem elongation, leaf expansion and fertility. These phenotypes ranged widely in severity and frequency and appeared to occur regardless of the S56 allele, affinity tag, or fused fluorescent reporter in each construct. It is not likely that the phenotypes exhibited by the constitutively expressed *eIF2α(5g)* is due to silencing of *eIF2α* since the transgenic plants did not resemble the *eif2α(2g)-2* mutant. Additionally, the *eIF2α(5g)* transgenes that were driven by a native promoter were identical to wild-type in morphology. Therefore, the phenotypes of the plant are most likely due to mis-regulation of *eIF2α* when it is over-expressed. The apparent mis-regulation of *eIF2α* results in growing meristem defects as evidenced by abrogated shoot elongation and failure of inflorescences to produce mature reproductive structures.

In this work, two proteins implicated in the mammalian ISR, eIF2α and the GCN2-kinase domain in addition to ABCF1, a protein which shares homology with the GCN1 cofactor, GCN20, were purified after being cloned and expressed in *E. coli*. Though much work has been done to characterize the ISR pathway in yeast and mammals, the overarching goal of the recombinant protein expression was to develop an *in-vitro* assay to tease apart the plant ISR pathway.

The coding sequence of *eIF2α(2g)* wild-type, S56A and S56D alleles were digested out of another template and ligated into pET28a under an IPTG inducible promoter and His-tag. After transformation to the BL21STAR *E. coli* cell line, the recombinant *eIF2α* alleles were expressed according to an already-optimized protocol. However, it remains to be seen if the S56D and S56A eIF2α alleles function as phospho-mimic and phospho-deficient variants of S56.

The purification of eIF2α protein was frustrated by the presence of an unknown contaminant at approximately 110kDa in size. The contaminant was suspected to be an aggregate of eIF2α and other *E. coli* proteins, a heat shock 70 protein bound to the 42 kDa recombinant eIF2α, a protein which has an epitope that binds Ni·NTA resin, or a recombinant eIF2α trimer. It is unlikely that an eIF2α trimer or protein aggregate could remain intact after SDS-PAGE. To address this possibility, the aliquots of protein used for SDS-PAGE were treated with either 10X SDS or 4M Urea at 37 °C for 5 minutes, but no change in ratio between the 42kDa eIF2α and the 110kDa contaminant band intensity was observed. To address the possibility of a Ni·NTA-binding *E. coli* protein or an HSP70-bound eIF2α, the lysate was incubated with Mg·ATP during equilibration with the Ni·NTA resin and, following column purification, HPLC-anion exchange purification was performed. Though the HPLC-anion exchange experiment resulted in purer fractions, the effort ultimately failed to purify eIF2α from the 110kDa contaminant.

The recombinant *ABCF1* protein was cloned, expressed, and purified successfully. The recombinant gene differs slightly from the endogenous gene CDS since a single codon underwent a silent mutation for the purpose of removing a restriction site. Although the original direction was to clone Arabidopsis *GCN1*, the direction was frustrated by a database mis-annotation of *ABCF1* as *GCN1*. Regardless, it is fortuitous that *ABCF3*, which shares high binding domain sequence homology with *ABCF1*, has been identified as the plant *GCN20* homologue. This means that *ABCF1* is still a relevant protein to be used in any kinase activity assay for which these proteins were intended to be used. In Arabidopsis, all the *ABCF*-family proteins share some level of sequence homology in their functional

nucleotide-binding domains [26]. It has not been determined whether *ABCF3* function can be complemented by other *ABCF*-family proteins in Arabidopsis.

A recombinant *GCN2-KD* was cloned, expressed, and purified successfully. A full-length recombinant *GCN2* was not achieved in this project. Each attempt to clone full-length *GCN2* in *E. coli* was undermined by random, disparate mutations across the gene, including the kinase domain. It is worth noting that other groups that express recombinant *GCN2* do not use *E. coli* [5, 31]. Expressing the truncated version of *GCN2* in *E. coli* was successful, however, domain activity of the recombinant GCN2-KD protein has not been confirmed.

The overarching goal of the cloning work was to design an in-vitro system by which the GCN2 kinase pathway can be explored. Some kinase domain activity assays were attempted which resulted in no eIF2α phosphorylation (not shown), but all avenues of experimental conditions and kinase buffers have not been explored. The groundwork for such an assay exists, but the conditions for kinase activity of GCN2 must be explored and confirmed.

List of References

1. Browning, K.S. and J. Bailey-Serres, *Mechanism of cytoplasmic mRNA translation.* Arabidopsis Book, 2015. **13**: p. e0176.
2. Urquidi Camacho, R.A., A. Lokdarshi, and A.G. von Arnim, *Translational gene regulation in plants: A green new deal.* Wiley Interdiscip Rev RNA, 2020. **11**(6): p. e1597.
3. Beilsten-Edmands, V., et al., *eIF2 interactions with initiator tRNA and eIF2B are regulated by post-translational modifications and conformational dynamics.* Cell Discovery, 2015. **1**(1): p. 15020.
4. Yamamoto, Y., et al., *The eukaryotic initiation factor (eIF) 5 HEAT domain mediates multifactor assembly and scanning with distinct interfaces to eIF1, eIF2, eIF3, and eIF4G.* Proc Natl Acad Sci U S A, 2005. **102**(45): p. 16164-9.
5. Hinnebusch, A.G., *Translational regulation of yeast GCN4. A window on factors that control initiator-trna binding to the ribosome.* J Biol Chem, 1997. **272**(35): p. 21661-4.
6. Golovko, A., et al., *The eIF2A knockout mouse.* Cell Cycle, 2016. **15**(22): p. 3115-3120.
7. Kashiwagi, K., et al., *Structural basis for eIF2B inhibition in integrated stress response.* Science, 2019. **364**(6439): p. 495-499.
8. Kenner, L.R., et al., *eIF2B-catalyzed nucleotide exchange and phosphoregulation by the integrated stress response.* Science, 2019. **364**(6439): p. 491-495.
9. Marintchev, A. and T. Ito, *eIF2B and the Integrated Stress Response: A Structural and Mechanistic View.* Biochemistry, 2020. **59**(13): p. 1299-1308.
10. Lokdarshi, A., et al., *Light Activates the Translational Regulatory Kinase GCN2 via Reactive Oxygen Species Emanating from the Chloroplast.* Plant Cell, 2020. **32**(4): p. 1161-1178.
11. Wek, R.C., H.Y. Jiang, and T.G. Anthony, *Coping with stress: eIF2 kinases and translational control.* Biochem Soc Trans, 2006. **34**(Pt 1): p. 7-11.
12. Wang, L., et al., *The inhibition of protein translation mediated by AtGCN1 is essential for cold tolerance in Arabidopsis thaliana.* Plant Cell Environ, 2017. **40**(1): p. 56-68.
13. Cho, H.Y., et al., *Ethylene modulates translation dynamics in Arabidopsis under submergence via GCN2 and EIN2.* Sci Adv, 2022. **8**(22): p. eabm7863.
14. Liu, X., T. Afrin, and K.M. Pajerowska-Mukhtar, *Arabidopsis GCN2 kinase contributes to ABA homeostasis and stomatal immunity.* Commun Biol, 2019. **2**: p. 302.
15. Lokdarshi, A. and A.G. von Arnim, *Review: Emerging roles of the signaling network of the protein kinase GCN2 in the plant stress response.* Plant Sci, 2022. **320**: p. 111280.
16. Pakos-Zebrucka, K., et al., *The integrated stress response.* EMBO Rep, 2016. **17**(10): p. 1374-1395.
17. Pereira, C.M., et al., *IMPACT, a protein preferentially expressed in the mouse brain, binds GCN1 and inhibits GCN2 activation.* J Biol Chem, 2005. **280**(31): p. 28316-23.
18. Sattlegger, E. and A.G. Hinnebusch, *Separate domains in GCN1 for binding protein kinase GCN2 and ribosomes are required for GCN2 activation in amino acid-starved cells.* Embo j, 2000. **19**(23): p. 6622-33.
19. Castilho, B.A., et al., *Keeping the eIF2 alpha kinase Gcn2 in check.* Biochim Biophys Acta, 2014. **1843**(9): p. 1948-68.
20. Pochopien, A.A., et al., *Structure of Gcn1 bound to stalled and colliding 80S ribosomes.* Proc Natl Acad Sci U S A, 2021. **118**(14).
21. Han, T.T., W.C. Liu, and Y.T. Lu, *General control non-repressible 20 (GCN20) functions in root growth by modulating DNA damage repair in Arabidopsis.* BMC Plant Biol, 2018. **18**(1): p. 274.
22. Faus, I., et al., *Arabidopsis ILITHYIA protein is necessary for proper chloroplast biogenesis and root development independent of eIF2α phosphorylation.* J Plant Physiol, 2018. **224-225**: p. 173-182.

23.   Izquierdo, Y., et al., *Arabidopsis nonresponding to oxylipins locus NOXY7 encodes a yeast GCN1 homolog that mediates noncanonical translation regulation and stress adaptation.* Plant Cell Environ, 2018. **41**(6): p. 1438-1452.

24.   Johnston, A.J., et al., *Genetic subtraction profiling identifies genes essential for Arabidopsis reproduction and reveals interaction between the female gametophyte and the maternal sporophyte.* Genome Biol, 2007. **8**(10): p. R204.

25.   Zhu, S., A.Y. Sobolev, and R.C. Wek, *Histidyl-tRNA synthetase-related sequences in GCN2 protein kinase regulate in vitro phosphorylation of eIF-2.* J Biol Chem, 1996. **271**(40): p. 24989-94.

26.   Faus, I., et al., *The ABCF3 Gene of Arabidopsis Is Functionally Linked with GCN1 but Not with GCN2 During Stress and Development.* Plant Molecular Biology Reporter, 2021. **39**(4): p. 663-672.

27.   He, H., et al., *Crystal structures of GCN2 protein kinase C-terminal domains suggest regulatory differences in yeast and mammals.* J Biol Chem, 2014. **289**(21): p. 15023-34.

28.   Lageix, S., et al., *Enhanced interaction between pseudokinase and kinase domains in Gcn2 stimulates eIF2α phosphorylation in starved cells.* PLoS Genet, 2014. **10**(5): p. e1004326.

29.   Masson, G.R., *Towards a model of GCN2 activation.* Biochem Soc Trans, 2019. **47**(5): p. 1481-1488.

30.   Dong, J., et al., *Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain.* Mol Cell, 2000. **6**(2): p. 269-79.

31.   Inglis, A.J., et al., *Activation of GCN2 by the ribosomal P-stalk.* Proc Natl Acad Sci U S A, 2019. **116**(11): p. 4946-4954.

32.   Lageix, S., et al., *Arabidopsis eIF2alpha kinase GCN2 is essential for growth in stress conditions and is activated by wounding.* BMC Plant Biol, 2008. **8**: p. 134.

33.   De Vylder, J., et al., *Rosette tracker: an open source image analysis tool for automatic quantification of genotype effects.* Plant Physiol, 2012. **160**(3): p. 1149-59.

34.   Browning, K.S., et al., *Determination of the amounts of the protein synthesis initiation and elongation factors in wheat germ.* J Biol Chem, 1990. **265**(29): p. 17967-73.

35.   Pfaffl, M.W., *A new mathematical model for relative quantification in real-time RT-PCR.* Nucleic Acids Res, 2001. **29**(9): p. e45.

Vita

Mark Edens was born in the month of January 1992 in East Tennessee. His family moved to the area from Pennington Gap, Virginia. He received his BS in Biochemistry, Cellular and Molecular biology from the University of Tennessee Knoxville in the Spring of 2017 and started his graduate program at the University of Tennessee in the Fall of 2018. After finishing his master's degree, he plans on beginning work in the agricultural plant science industry.