Developing antibodies to *Arabidopsis thaliana* translation initiation factor eIF4G

Joceline Liu

Dean's Scholars Thesis
Spring 2007

Abstract

In eukaryotes, a key way to regulate biochemical pathways and protein synthesis is through translation, the reading of an mRNA sequence to make proteins and enzymes. In translation, one of the main steps that can be regulated is the formation of the initiation complex, which is necessary to position the machinery at the right spot on the mRNA before protein can be generated. A major player in translation initiation is the eIF4F complex (***e***ukaryotic ***i***nitiation ***f***actor ***4F***), made up of three different proteins: eIF4A, eIF4E and eIF4G. The eIF4F complex specifically functions to recruit the small ribosomal unit, which serves as an interface between mRNA and the new protein being made, to the correct place on the existing mRNA strand. In this way, eIF4F may control the rate of protein synthesis and, ultimately, the abundance, or conversely the absence, of specific proteins in organisms. Therefore, the eIF4F complex may be useful in controlling expression systems, which can range in application from hormone production in animals to synthesizing useful medicines in plants.

One of the major goals of the Browning lab is to understand better the effects of eukaryotic initiation factor mutants on phenotype in the *Arabidopsis thaliana* plant. One part of the lab research thus focuses on the generation of "knock out" *A. thaliana* mutants, eliminating one or a combination of initiation factors. Antibodies specific to the different initiation factors are useful for assaying protein expression and for immunolocalization, which allows you to visualize the presence or location of the protein which bind to that antibody. While antibodies have been generated for several initiation factors already, some still remain elusive. In a previous project, I produced antibodies to the *Arabidopsis* eIF4F complex and the lab already possesses *Arabidopsis* (At) eIF4E antibody. However, antibodies specific to AteIF4G have not yet been made, due to difficulties

producing ample AteIF4G protein for antibody generation. It is important to differentiate between *Arabidopsis* and wheat initiation factors, as the possessed wheat eIF4G antibody does not cross-react with AteIF4G. A large part of the problem may be attributed to *A. thaliana* eIF4G protein instability compared to wheat eIF4G; the protein is prone to rapid degradation and low expression levels in *E. coli*.

In order to optimize the expression of AteIF4G, different vectors, expression systems, growth media, bacterial growth conditions and protein size were tried. In testing different vectors and expression systems, I looked for more stable folding of the protein. However, no large differences in expression were seen. Different growth media were compared in protein expression levels, using media that slowed expression over a long amount of time and allowed for better folding and stabilizing of the protein, or media that provided excess amounts of nutrients to increase the output of all proteins, or media that simply followed basic protocols. The slow-expression medum did not result in more final protein. Conversely, the excess-nutrient medium resulted in an increase in eIF4G protein, but also an even larger amount of contaminant proteins that were not the desired eIF4G. Compared to the basic media, there was not a significant increase in yield of pure eIF4G. Lastly, because the full-length eIF4G protein is approximately 150 kD, the excessive length of an already unstable protein could explain the surprisingly low protein expression level. A cDNA for an N-terminal truncation of eIF tr4G (~ 24 kDa) was prepared. This construct has a His-tag on the C-terminus to simplify purification. Antibodies raised to the N-terminal portion of eIF4G will presumably react with the full-length protein. From this shortening of eIF4G, a noticeable increase in protein expression was observed. From purified eIFtr4G protein, antibodies were produced in rabbits. This eIF4G-specific antibody can be used in the future to determine the presence of eIF4G protein or eIF4F complex, and can serve to clarify their roles in certain phenotypes and biochemical pathways.

Contents

1 Introduction

One of the key processes in cell survival and functioning is the generation of proteins and enzymes, which are necessary for the cell in such important processes as DNA replication, cell division and recognition of substances and other cells. This essential production of protein results from two processes that convert information encoded in the sequence of DNA to the amino acid sequence of proteins via an RNA intermediate. The transfer of information from DNA to RNA occurs through transcription, while the reading of RNA to synthesize protein occurs via translation as shown in *Figure 1*. In both transcription and translation, the process occurs in three phases: initiation, elongation and termination.

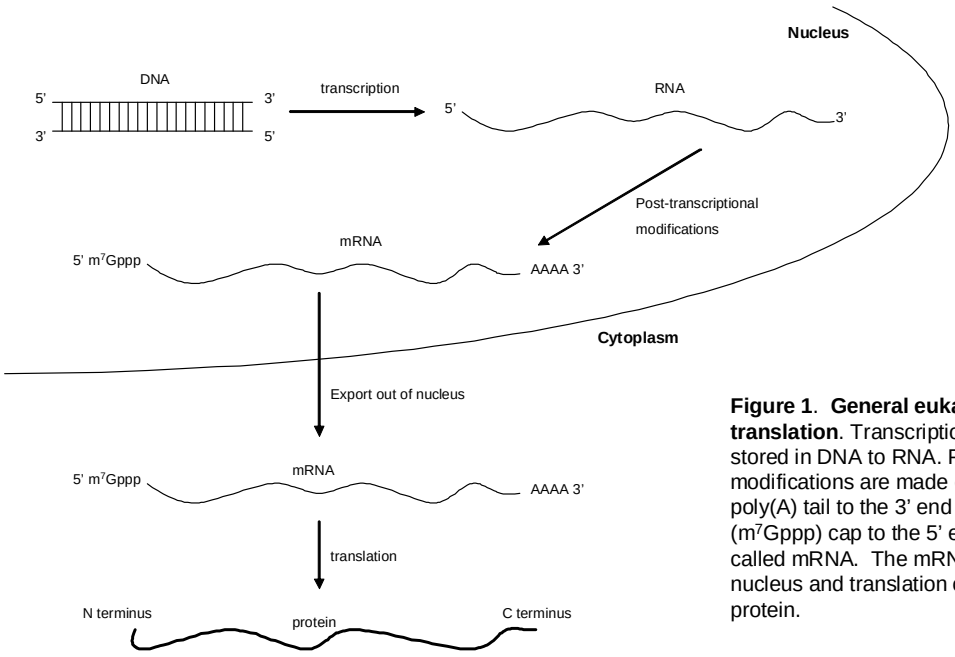

Figure 1. **General eukaryotic transcription and translation**. Transcription converts information stored in DNA to RNA. Post-transcriptional modifications are made on the RNA, adding a poly(A) tail to the 3' end and a methylguanosine (m^7Gppp) cap to the 5' end. The resulting strand is called mRNA. The mRNA is exported out of the nucleus and translation converts mRNA into protein.

1.1 The process of eukaryotic translation initiation

Throughout this project, focus was placed specifically on the initiation phase of translation in eukaryotes. In most eukaryotes, before the RNA transcript is exported from the nucleus to the cytoplasm of the cell, post-transcriptional modifications are made on the RNA, adding a cap to the 5' end, the "front", and a poly(A) tail on the 3' end, the "back" (*Figure 1*). These two components are vital for the efficient initiation of translation. During cap-dependent translation initiation, the most common type in eukaryotic cells, a multiprotein complex is formed on the messenger RNA (mRNA) near the 5' end. This multiprotein complex, called the 43S pre-initiation complex, is formed through sequential recruitment of different eukaryotic initiation factors (eIFs) to the mRNA. This pre-initiation complex serves to orient the translation reading machinery to the right place on the mRNA.

One of the first steps in forming this pre-initiation complex is the recognition of the 5' cap of the mRNA transcript by eukaryotic initiation factor 4E (eIF4E), which binds eukaryotic initiation factor 4G (eIF4G) as seen in *Figure 2a* (Imataka 1997). Second, eIF4G recruits eukaryotic initiation factor 4A (eIF4A) demonstrated in *Figure 2b*, which functions as an ATP-dependent RNA helicase (Von der Haar 2000). Third, eIF4G binds to eIF4B, which binds poly(A) and stabilizes the binding of ATP to eIF4A, promoting the ATP-dependent RNA helicase activity of eIF4A. Also, poly(A)-binding protein (PABP), which also binds to the 3' poly(A) tail as in *Figure 2c* (Kawaguchi 2002). This assembly of proteins presumably circularizes the mRNA, through the simultaneous interaction of eIF4E, binding the 5' end and both eIF4G and PABP bound to the 3' end of the mRNA. eIF4G also

associates with eukaryotic initiation factor 3 (eIF3), shown in *Figure 2d*, the ribosome-associated initiation factor, which allows the small subunit (40S) of the ribosome to bind the

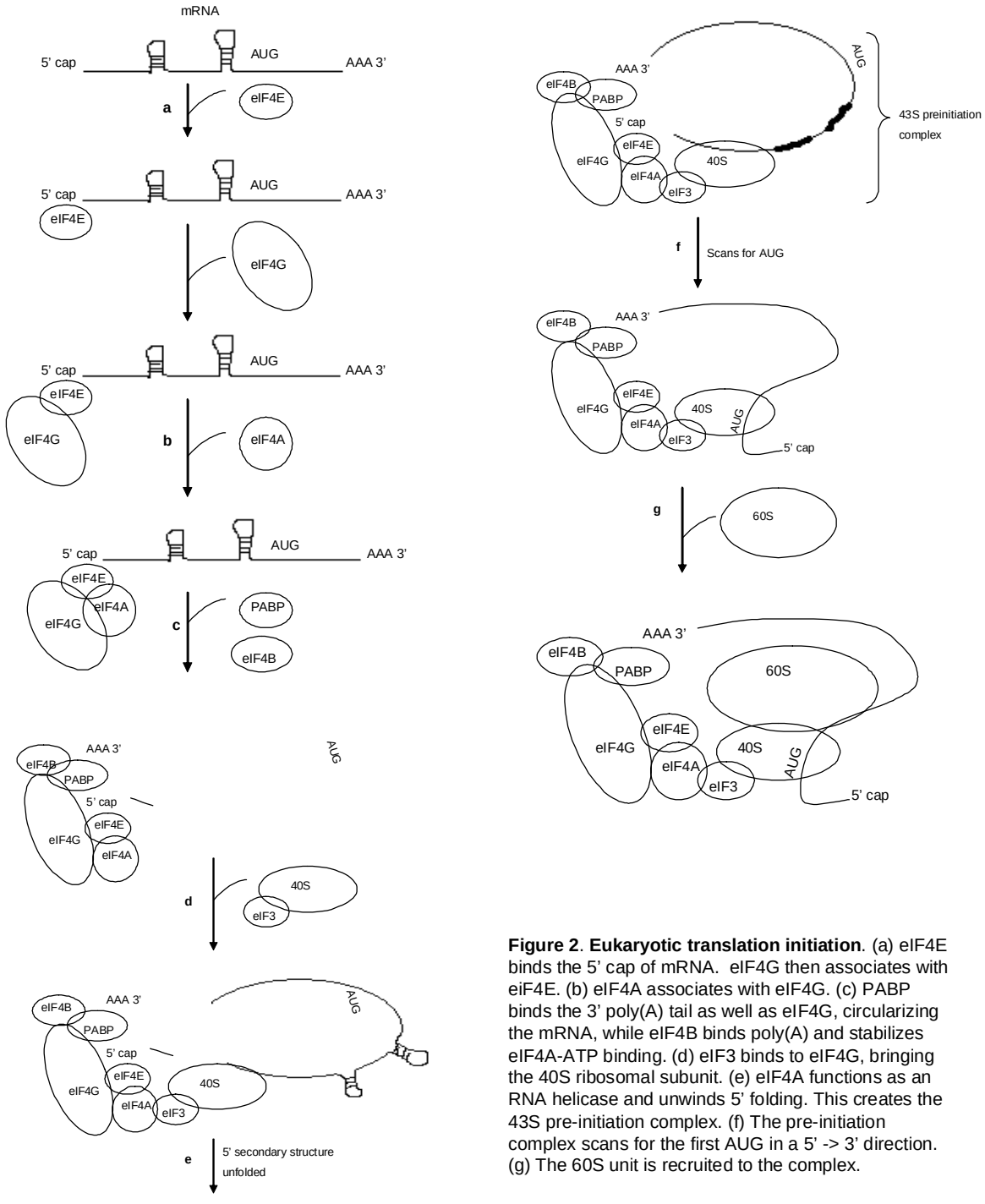

Figure 2. **Eukaryotic translation initiation**. (a) eIF4E binds the 5' cap of mRNA. eIF4G then associates with eiF4E. (b) eIF4A associates with eIF4G. (c) PABP binds the 3' poly(A) tail as well as eIF4G, circularizing the mRNA, while eIF4B binds poly(A) and stabilizes eIF4A-ATP binding. (d) eIF3 binds to eIF4G, bringing the 40S ribosomal subunit. (e) eIF4A functions as an RNA helicase and unwinds 5' folding. This creates the 43S pre-initiation complex. (f) The pre-initiation complex scans for the first AUG in a 5' -> 3' direction. (g) The 60S unit is recruited to the complex.

assembled pre-initiation complex so that translation can begin (Von der Haar 2000). After the association of the mRNA and the 40S subunit, the eIF4A RNA helicase unwinds folding on the 5' end of the mRNA that would otherwise impede the progression of the ribosome (*Figure 2e*). Then, as shown in *Figure 2f*, the 43S pre-initiation complex scans towards the 3' end of the transcript until it finds the correct translation start point (Kawaguchi 2002). The final step in translation initiation is the association of the large subunit (60S) of the ribosome (*Figure 2g*), completing the ribosome complex and allowing the elongation phase of translation to begin.

1.2 Translation initiation as a point of regulation

Because of the necessity of proteins and enzymes for the basic functioning of cells and organisms, translation is a process that is heavily regulated. The recruitment of the 40S ribosomal subunit to the mRNA is usually the rate-limiting step in protein synthesis, which makes translation initiation, more specifically the formation of the 43S pre-initiation complex, a logical point of translational regulation. Translation initiation can control the abundance, or absence, of proteins and enzymes necessary to perform a certain function.

Translational control has also been implicated in regulating gene expression (Han 2002). In other words, the amount of a gene transcript is not always proportional to the amount of protein that is synthesized because of the regulation of translation (Kawaguchi 2002). Because of the central role that the formation of the pre-initiation complex plays as the admission point of mRNA into translation, modulation of its formation can selectively promote the translation of certain mRNAs. This can explain how a cell can be re-

programmed to express genes in different levels in response to environmental or temporal signals (Morley 1997). Studies have shown that in changing plants from light to dark environments, or exposing them to heat stress or water shortage, protein expression levels are changed for selected mRNAs (Kawaguchi 2002).

1.3 The eIF4F complex

The eIF4F complex is composed of eIF4E and eIF4G. eIF4A and eIF4B, interact with this complex and are responsible for the binding of mRNA to the ribosome. By convention, mammalian systems usually consider eIF4A as a part of the eIF4F complex, while in plant systems, the eIF4F complex is considered only eIF4E and eIF4G based on the ease of dissociation of eIF4A during extensive purification.

1.3.1 eIF4E: the cap-binding component of eIF4F

The cap-binding protein component, eIF4E specifically recognizes and binds the cap structure at the 5' end of mRNAs, thought to be the first step in assembling the pre-initiation complex (Ruud 1998). The eIF4E directly interacts with eIF4G, and when it is within the eIF4F complex, eIF4E's binding affinity for the 5' cap increases about 10-fold. It is hypothesized that the binding of eIF4G to eIF4E induces a conformational change, leading to a change in binding affinity (Prévôt 2003). From these observations, it seems that although eIF4E alone can bind the cap structure, an eIF4E-eIF4G interaction, such as in the eIF4F complex, is needed for optimal binding of the cap (Prévôt 2003).

1.3.2 eIF4A: the RNA helicase of eIF4F

eIF4A is an ATP-dependent RNA helicase, meaning it harnesses energy from the hydrolysis of ATP to melt secondary structure and unwind RNA, facilitating ribosome binding and scanning for an AUG codon, or the starting point for translation. In a similar change as in eIF4E, the rate of ATP hydrolysis and of RNA unwinding is greatly increased when eIF4A is part of the eIF4F complex (Ruud 1998).

1.3.3 eIF4B: an RNA-binding protein

eIF4B is an RNA-binding protein that bind the poly(A) tail and eIF4G, promoting the association of ATP with eIF4E (Kawaguchi 2002). In this way, it promotes the ATP-dependent RNA helicase activity of eIF4E to unwind the RNA strand. eIF4B only promotes this activity when a part of the eIF4G/eIF4E/eIF4A protein complex, while not in the case of eIF(iso)4G/eIF(iso)4E/eIF4A complex (Kawaguchi 2002).

1.3.4 eIF4G: the scaffold of eIF4F

The common thread between all of the eIFs in the eIF4F complex, and most of the other eIFs involved in forming the 43S pre-initiation complex, is eIF4G. eIF4G doesn't act as a fixed docking station for the other eIFs, but is rather a changing scaffold that is modified by the proteins it associates with (Prévôt 2003). Through all of these interactions, eIF4G acts as a molecular bridge between the 5' cap structure and the 3' poly(A) tail, and at the same time, between the mRNA and the 40S ribosomal subunit.

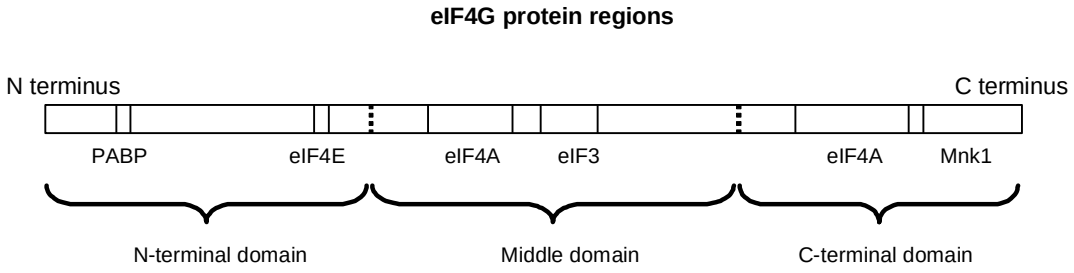

Figure 3. The domains of eIF4G. eIG4G is divided into three domains. The N-terminal domain contains the PABP and eIF4E binding sites. The middle domain contains binding sites for eIF4A and eIF3. This middle domain has been determined as necessary for ribosomal recruitment. The C-terminal region contains a second eIF4A binding site as well as a Mnk1 binding site. This domain has been shown to play a modulatory role.

Since most of the research I conducted was related to this important eIF4G protein, the majority of the remaining information presented will pertain to eIF4G: its characteristics, interactions, and applications. The eIF4G protein can be divided into three domains: the N-terminal region at one end, the middle region, and the C-terminal region at the other end (*Figure 3*). The N-terminal region contains the eIF4E binding site and the PABP binding site. The middle region of eIF4G contains the eIF4A and eIF3 binding sites. The C-terminal region contains a second eIF4A binding site and a docking site for Mnk1, a protein kinase that modulates the activity of mammalian eIF4E (Morino 2000).

Not all of eIF4G is necessary to perform translation initiation, so to determine the part of eIF4G that was necessary, researchers replaced eIF4G's N-terminal domain with an RNA binding protein, thus deleting the eIF4E and PABP binding sites (De Gregorio 1999). This allowed researchers to attach the still intact middle region to certain RNAs through the RNA binding protein. By doing this, they could look at the ability or inability ability of different parts of eIF4G to initiate translation of a reporter mRNA, producing an easily

detectable protein. From this study, it became evident that the middle core region of eIF4G acts as the active "ribosome recruitment core" and is absolutely necessary for eIF4G to be translationally active (De Gregorio 1999). In a different study, the C-terminal domain of mammalian eIF4G was shown to play a modulatory role.

While in the De Gregorio study, the eIF4A binding site in the middle region was shown to be necessary for any translation initiation to occur, a second study shows that the second eIF4A binding site in the C-terminal region is necessary for vigorous translation. How the C-terminal region modulates eIF4G activity is still unknown, but two possible models have been proposed in which the C-terminal region changes conformation to allow eIF4A binding in two different ways (Morino 2000). Both of these models exemplify how eIF4G is a dynamic molecule that continually changes shape to modulate initiation.

Another essential function performed by eIF4G is the circularization of the mRNA, which occurs through the 5' cap – eIF4E – eIF4G – PABP – 3' poly(A) tail interactions (*Figure 4*). Similar to the increases in efficiency of eIF4A and eIF4E when bound in complex with eIF4G, PABP – eIF4F interaction increases both the affinity of the eIF4F complex for

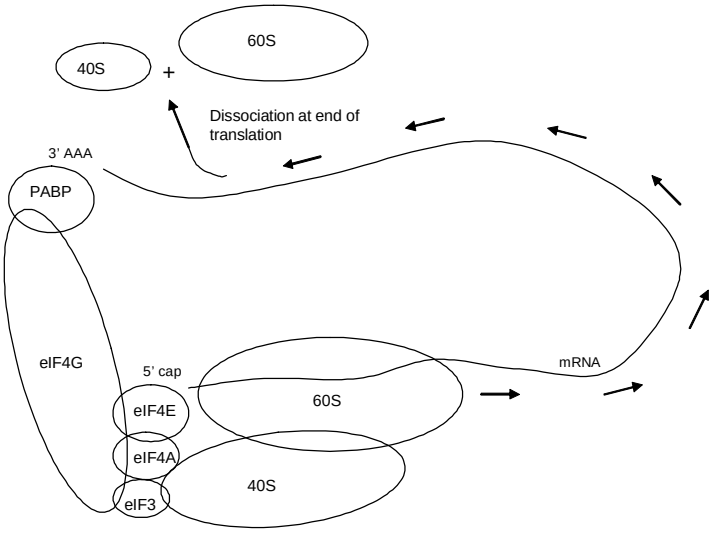

Figure 4. 5'-3' circularization of mRNA. This circularization serves two purposes: 1) it makes sure only intact mRNA's will be translated and 2) it promotes efficient re-initiation of translation because of the proximity of the ribosome subunits and the eIFs after dissociation.

the 5' cap about 40-fold, as well as the affinity of PABP for poly(A) tail (Borman 2002). The significance of 5' to 3' bridging is that it may provide a proofreading mechanism so that only intact mRNAs, containing both the 5' cap and the 3' poly(A) tail, are translated (Craig 1998). Through this mechanism, translation of nonfunctional or damaged mRNAs would be prevented. Secondly, the circularization of mRNA would promote more efficient re-initiation of translation. Because of the proximity of the ends, when ribosomes dissociate at the 3' end of the mRNA after termination and eIF4F dissociates from the 5' end after initiation, the ribosomes will be more readily recaptured to the eIF4F complex because they are close together, increasing the overall rate of translation (Borman 2002).

1.4 Circularization of mRNA during initiation

This circularization of mRNA has been shown in yeast and plants, with PABP binding the poly(A) tail, and in humans, with PAIP-1 playing this role (Craig 1998). The importance of the 5' – 3' interaction can also be seen in the way that viruses have evolved to exploit this tactic to increase translation of viral proteins. For example, the Barley yellow dwarf virus (BYDV) which does not have the 5' or 3' end modifications, and yet is still translated at a moderate level in plant hosts. BYDV accomplishes this through binding host eIF4F to its own RNA. After the eIF4F is recruited, the viral RNA interacts with itself to form a stem loop that essentially creates a "kissing-loop" that resembles the 5' – 3' circularization of the host (Kawaguchi 2002).

Since the 5' – 3' interaction facilitates translation initiation, it follows that disruption of this interaction would inhibit translation on a global level. Some mRNAs in plants have evolved a mechanism to deal with this. For instance, heat shock in plants causes a general

decrease in translation because of the disruption of the 5' to 3' interaction. However, the group of Heat shock proteins (HSP), which aid the plant in dealing with excessive heat, are still able to initiate translation efficiently because the HSP mRNAs establish a thermostable 5' – 3' complex (Kawaguchi 2002). Thus plants have evolved a way to use the 5' – 3' circularization to their advantage in making sure heat shock coping proteins are able to continue translation even under adverse conditions.

1.5 Plant vs. general eukaryotic translation initiation

My research focuses on plant-specific translation initiation. In almost all aspects, plant translation is identical to general eukaryotic translation. However, one difference is the presence of an alternate form, or isoform, of eIF4F, called eIF(iso)4F, in addition to eIF4F. Just as eIF4F contains eIF4E and eIF4G as its major components, eIF(iso)4F also contains their isozyme forms, eIF(iso)4E and eIF(iso)4G which differ in size and amino acid sequence. The eIF(iso)4F complex has been found in all higher plants tested thus far including wheat, maize, cauliflower, *Arabidopsis* and barley(Browning 1996). The eIF4G and eIF(iso)4G proteins show significant sequence homology, with the best conserved motif at the eIF4E binding site (Morley 1997). Both the regular and iso- forms are present in higher plants. The reason for this is still unclear, as they appear to have the same functional traits (Browning 1996). eIF(iso)4F complex can substitute for the eIF4F complex in the absence of eIF4F, to recruit the 40S subunit to mRNA, and to unwind RNA when eIF4A is present.

The few differences that have been noted include an eIF(iso)4E preference for hypermethylated caps and that eIF4F is more sensitive to 5' secondary structure than eIF(iso)4F (Browning 1996). Also, in wheat germ, eIF(iso)4F is present in 3- to 5-fold

higher amounts than eIF4F (Browning 1996). Despite these distinguishing factors, plants in which eIF(iso)4E has been eliminated still show normal development under normal growth conditions, raising questions as to its purpose (Duprat 2002).

1.6 Regulation of translation initiation through phosphorylation

One of the ways to regulate translation initiation in mammalians is through the rate of formation of the eIF4F complex. This can be modulated either through the phosphorylation state of eIF4E or that of 4E binding protein (4E-BP) shown in *Figure 5a.*

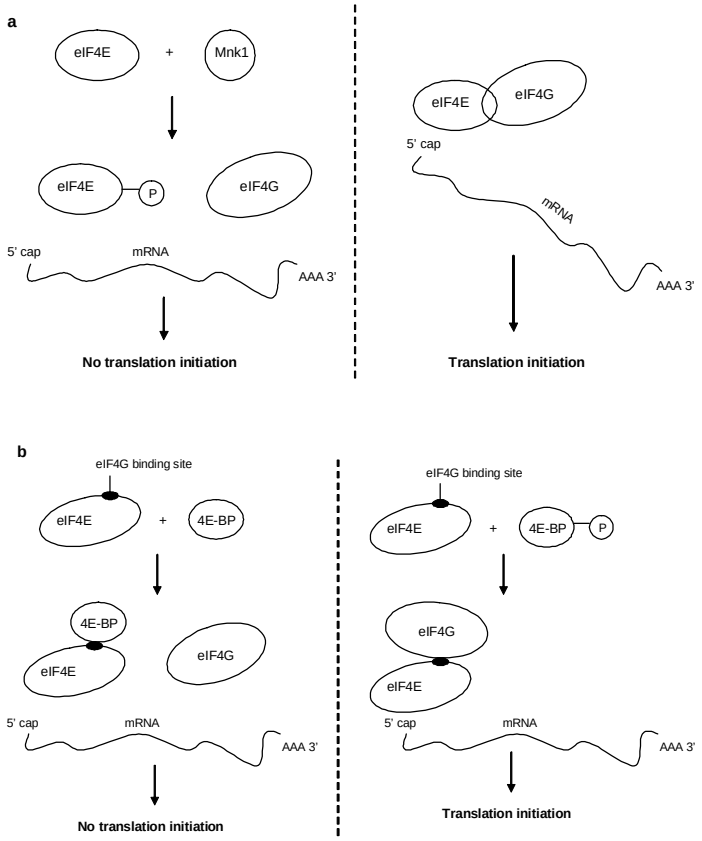

Figure 5. Mammalian translation initiation regulation by phosphorylation. (a) When eIF4E is phosphorylated by Mnk1, it is inhibited from binding eIF4G, thus inhibiting translation initiation. (b) 4E-BP binds eiF4E at the same site that it usually binds eIF4G, thus inhibiting translation initiation. However, when 4E-BP is phosphorylated, it is unable to bind eIF4E, thus allowing translation initiation to continue.

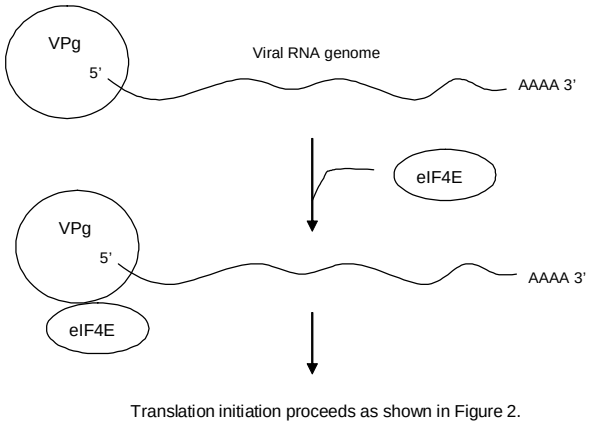

Figure 6. *Potyvirus* mechanism of taking over host translational machinery. The viral genome is a single-stranded RNA with a 3' poly(A) tail but a capless 5' end. They produce VPg, a protein which binds the 5' end of the viral genome and acts as a cap mimic. VPg then recruits eIF4E and proceeds in translation initiation as shown in Figure 2.

The mammalian eIF4E protein can be phosphorylated by Mnk1, a protein kinase. In its phosphorylated state, eIF4E does not bind well with eIF4G and also has a lower affinity for the 5' cap thus decreasing translation initiation (Kawaguchi 2002).

Another way to modulate eIF4E in mammals is through 4E-BP, which binds to eIF4E at the same site that it binds with eIF4G in its unphosphorylated state, thus inhibiting eIF4F formation shown in *Figure 5b* (Imataka 1997). Conditions that promote growth, such as treatment of cells with insulin or growth factors, results in the phosphorylation of 4E-BP which leads to an increase in eIF4F complex formation, and a resulting increase in translation. Another protein which regulates translation levels through interaction with the initiation complex in *S. cerevisiae* is Dcp1, a protein involved in the decapping process that marks mRNA for degradation. Dcp1 is only able to act when bound to eIF4G when not bound to eIF4E. In this way, eIF4G, depending on association with translation factors or decay proteins, can act as a sensor molecule to modulate the translation and degradation of mRNA (Prévôt 2003).

New research has identified a novel disulfide bond in wheat eIF4E, which is also conserved in other plants. The state of the Cys involved in the disulfide bond can be altered based on either a reducing or oxidizing environment (Mozingo 2007). The redox state of a plant can be influenced by light or environmental stress and may regulate photosynthesis, seed development and germination (Mozingo 2007). It is hypothesized that the redox state of the plant influences the formation of the disulfide bond, which then affects the binding affinity of eIF4E for the 5' cap, eIF4G and other proteins and machinery that interact with eIF4E (Mozingo 2007).

1.7 Viral and apoptotic strategies of taking over plant translational apparatus

Studying plant virus infections and how they commandeer host translational machinery to translate their own genome provides a new take on translation initiation. This area of research is one example where more knowledge about translation initiation could help in treating and preventing viral infections. The largest genus of plant viruses, *Potyvirus*, have a single stranded RNA genome, with a viral protein, VPg, bound to the 5' end and a 3' tail. It is thought that VPg functions as a cap mimic and binds to eIF4E or eIF(iso)4E and is then able to recruit the rest of the initiation complex and translate its genome (*Figure 6*).

Interestingly, once either eIF4E or eIF(iso)4E was mutated, plants developed resistance against several potyviruses after just a few amino acid changes in eIF(iso)4E. This shows the necessity for eIF4E in plant viral infections and also implies that different potyviruses use different isoforms of eIF4E. Similar results have been observed in other

studies involving viruses using eIF4G and/or eIF4E to hijack host translational machinery (Robaglia 2006). Because viruses rely completely on host translational machinery to produce viral proteins, their ability to recruit translation initiation factors such as eIF4G is an important target for the control of viral diseases.

Viruses have evolved many different, very effective mechanisms to take over host translation initiation machinery for their own viral genome translation. One way in which some mammalian picornaviruses are able to do this is through internal initiation, rather than cap-dependent initiation, as there is no 5' cap on the viral genome. Internal initiation occurs in mRNAs that have a long 5' untranslated region (UTR) with a special secondary structure that serves as an internal ribosome entry site (IRES) as shown in *Figure 7a*. Infection with picornaviruses, such as poliovirus, is accompanied with an abrupt end of host protein synthesis, while virus translation continues. This effect is largely mediated through 2A and L viral proteases, which cleave eIF4G, separating the binding sites for eIF4E and PABP (*Figure 7b*). By cleaving off the cap-binding component of eIF4G, the virus has inhibited cap-dependent translation (*Figure 7c*), the usual method for host cell mRNAs, and made cap-independent translation more likely, the method for viral RNA genome (Prévôt 2003). Cap-independent translation proceeds as the remaining C-terminal end of eIF4G binds to the IRES and continues to assemble the pre-initiation complex (*Figure 7d*). Understanding all the varied ways that viruses are able to replicate so efficiently and steal resources from normal cell processes can contribute to developing new ways of controlling expression.

The importance of eIF4G is further highlighted by its essential function in apoptosis, or programmed cell death in mammalian cells. Apoptosis is a normal part of development, providing a way to remove damaged or old cells from tissues. Apoptosis is split up into two

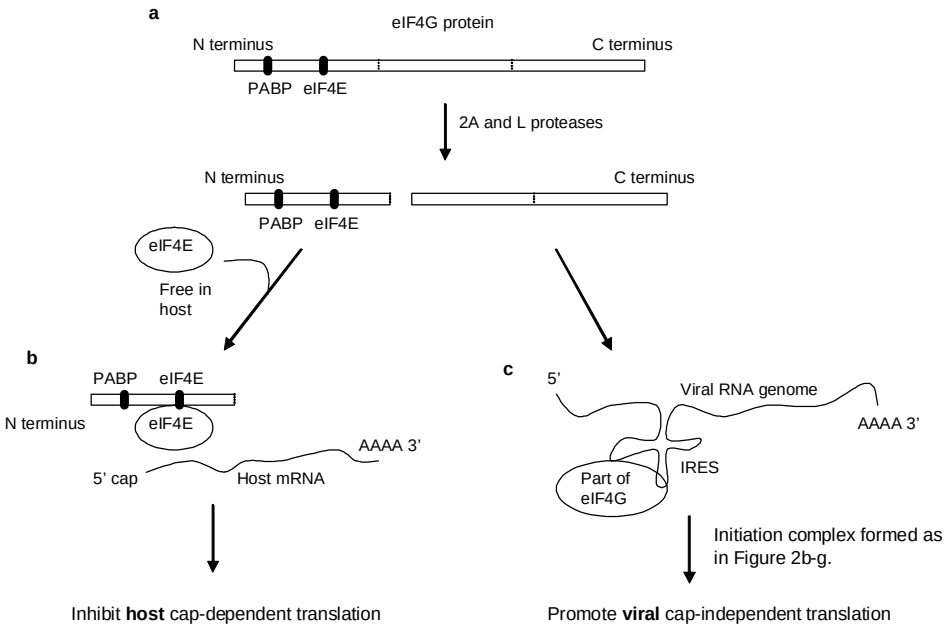

Figure 7. Viruses use host machinery for cap-independent translation initiation via an IRES. (a) eIF4G is cleaved by 2A and L proteases, separating the binding sites for eIF4E and PABP from the rest of the protein. A more detailed version of eIF4G can be seen in Figure 3. (b) The N-terminal domain of eIF4G binds with eIF4E in the host, but is unable to initiate translation without the rest of eIF4G. Thus, host cap-dependent translation is inhibited. (c) Viral genomic RNA does not contain a 5' cap, but contains 5' UTR folding, an IRES, that can be used to initiate translation with the middle and C-terminal regions of eIF4G. Thus, viral cap-independent translation is promoted.

phases, with the first being the induction phase, where signaling leads to the activation of death signaling proteins, and the second being the execution phase, in which there is rapid degradation and cell death. One class of the death signaling proteins are the caspases: proteases that cleave the amino acid cysteine. Because of the precedence of eIF4G as a target for the termination of host translation in viruses, it was estimated to be a likely target for caspases during apoptosis. As theorized, caspase 3 is responsible for eIF4G cleavage and eIF4G cleavage corresponded to the extinction of translation (Marissen 1998). By cleaving eIF4G, a crucial scaffolding protein in translation initiation, cell translation is shut off and disruption of cell processes ensues, explaining the quick execution phase of apoptosis.

While this has been shown in mammals, it is still unclear whether plants use a similar method for apoptosis.

1.8 Current Work

The Browning laboratory studies eukaryotic initiation of translation in plants, using *Arabidopsis thaliana* and wheat as model systems and is seeking a molecular description of the interaction of mRNA with initiation factors eIF4F, eIF4A and eIF4B. The focus of my research is the expression of AteIF4G protein to use to raise rabbit antibodies. Antibodies are a useful tool for the visualization and analysis. Recombinant protein may be used to raise rabbit antibodies specific to that protein, which is used both in *in vitro* methods, such as Western analysis or immunoprecipitation, and *in situ* methods, such as immunolocalization. One of the research projects of the laboratory is to investigate the effects of knockouts of different combinations of eIFs in *A. thaliana*, looking at the resultant phenotypic changes, differences in development, and environmental responses. When conducting these studies, antibodies specific to the knockout initiation factor of interest are very useful in confirming the absence of gene products and in monitoring eIF levels in response to different conditions. Because there have been difficulties in the past with expressing sufficient full length AteIF4G for antibody production, I created a shorter version of the protein, AteIFtr4G. The vector that houses the gene is a possible reason for low expression, because when eIF4G was originally cloned into its vector, pET22b, the vector had to be modified, creating vector XpET22b. This modification could have damaged the vector to operate at a less than optimal level, resulting in lower protein expression. Therefore, we will be inserting

the eIFtr4G gene back into the XpET22b vector as well as into an untouched vector, pET23a, and comparing expression levels to see whether the vector was one of the reasons expression has been chronically low.

To increase protein expression levels even further, I varied the environmental growth conditions, including the temperature, the level of bacterial growth at which protein expression is induced, the strain of *E. coli* used, and the concentration of isopropyl β-D-1-thiogalactopyranoside (IPTG) used to induce protein expression. Each of these variables must be optimized for different proteins. Induction tests were performed to find the most favorable conditions for expression of *Arabidopsis* eIFtr4G and subsequent purification to obtain sufficient amounts of protein for antibody production.

One of the uses for the eIF4G specific antibody will be to help study gene knockout plants. T-DNA insertion mutants have been obtained for all the Arabidopsis genes for eIF4G, eIFiso4G, eIF4E and eIFiso4E. In a mutant plant study similar to what the eIF4G antibody is needed for, the effects of knocking out the eIF(iso)4E gene in *A. thaliana* were studied using eIF(iso)4E-specific antibody. The resulting *A. thaliana* mutant line did not show any morphological or developmental changes in normal growth conditions. However, it was found that when the mutant plant was exposed to potyviruses TuMV and LMV, it was resistant to infection, implying that these two potyviruses required eIF(iso)4E to translate the viral genome(Duprat 2002). In contrast, the mutant plants were not resistant to CMV or *Tomato black ring virus strain ED* (TBRV-ED), implying that these two viruses use a different initiation factor to translate their genome (Duprat 2002). In this study, it was important to verify the total lack of eIF(iso)4E RNA and protein in order to consider the plant an eIF(iso)4E knockout plant line. Western analysis probes protein with an antibody, which detects the protein that the antibody was developed against. If the eIF(iso)4E gene has been

successfully knocked out, the plant extract, when treated with antibody against the protein product of the gene, should not react since there is no gene encoding for eIF(iso)4E. Also by Western analysis, comparison of protein levels in different plants can be made when plant extract is treated with antibody against the protein of interest. In this way, one can look at the effects of knocking out one initiation factor on the amount of others. For example, in the eIF(iso)4E knockout study, it was found that in knockout plants, the amount of eIF4E was elevated compared to wild-type plants (Duprat 2002).

2 Materials and Methods

2.1 Cloning

The *Arabidopsis* full-length eIF4G in XpET22b was first cloned by Lauren Harkinson and Elizabeth Burks. eIF4G was cloned in two pieces into the pET22b vector (Novagen), but was modified to no longer contain a XbaI site. This modified vector is now called XpET22b. From this original plasmid DNA, a truncated version of eIF4G was made, consisting of the first 609 bp on the 5' end of the eIF4G gene, called eIFtr4G. This tr4G was generated by cutting at two restriction sites to remove most of the eIF4G gene and only retain the N-terminal region. The XpET22b was relegated closed. This construct should place the his-tag in frame with the eIFtr4G. A second construct was made by removing the N-terminal region with two restriction enzymes and ligating into pET23a plasmid (Novagen). This construct should also be in frame with the his-tag in pET23a. Moving the N-terminal region into another vector was to be sure that the poor expression in XpET22b was not due to some unintended alteration during the modification XpET22b. The truncated contructs were transformed into competent cells for expression to compare for vector effects on expression.

2.1.1 Creating the eIFtr4G fragment and insertion into a vector (*Figure 8*)

NdeI and HindIII were selected as sites for removal of the N-terminal region of AteIF4G from XpET22b for recloning into pET23a. HindIII was selected to remove the C-terminal portion of AteIF4G and to religate XpET22b closed so that it only contained the N-terminal region. However, it was found that the eIF4G gene possessed by the Browning Laboratory had several point mutations that did not affect the peptide sequence of the protein. The diagrams in this study correspond to those seen in the Browning Lab eIF4G gene (refer to Appendix B for AteIF4G cDNA sequence).

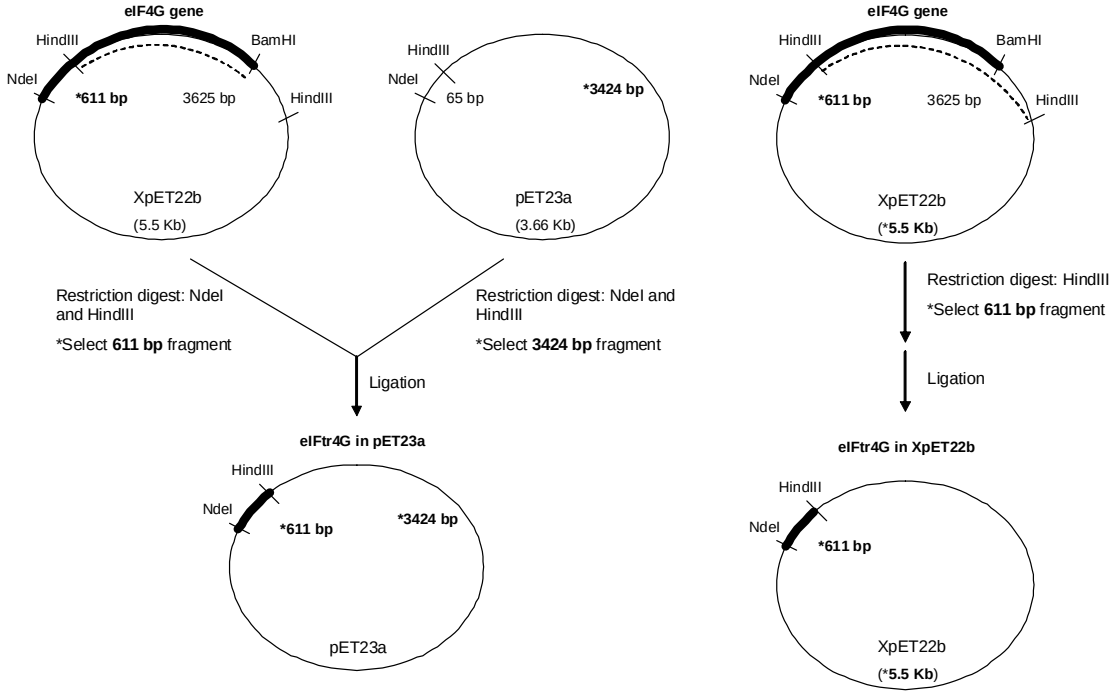

Figure 8. Creating the eIFtr4G gene and insertion into vectors.

eIF4G in XpET22b plasmid DNA was obtained from previous stocks, or prepared fresh by Mini-prep (QIAgen). Restriction digest of eIF4G in XpET22b and pET23a vector was performed, using NdeI and HindIII restriction enzymes (New England Biolabs) and appropriate buffers according to the detailed procedure in **Appendix C**. eIF4G in XpET22b plasmid DNA was also digested with only HindIII and appropriate buffers following the same restriction digest procedure.

The restriction digest mixtures were run on a DNA agarose gel. From the digest of pET23a (NdeI/HindIII), the major plasmid vector band (3427 bp) was recovered. From the AteIF4G/XpET22b digest (NdeI/HindIII), the N-terminal region (609 bp) was recovered. From the AteIF4G/XpET22b digest (HindIII), the plasmid plus the N-terminal region (6100 bp) was recovered. The DNA was recovered from the agarose using Ambion spin columns and centrifugation (3000 rpm for x min). The eluted DNA was ethanol precipitated (see **Appendix C)** and dissolved in 5.5 µl of water.

The insertion of the N-terminal AteIF4G fragment (609 bp) into the pET23a vector (3427 bp) was carried out by standard ligation procedures as described in **Appendix C**. This ligation should produce tr4G in pET23a. The N-terminal AteIF4G fragment retained in XpET22b digest (HindIII) is self-ligated and will produce tr4G in XpET22b.

The ligation reactions were transformed into *E. coli* DH5α competent cells as described in **Appendix C**. Single colonies were selected and plasmids prepared to confirm the ligation products were correct. Plasmid DNA was used for a restriction digest to test whether the vector contains the eIFtr4G gene (refer to **Appendix C**). NdeI and HindIII restriction enzymes were used to screen the plasmids and the restriction digests were separated by electrophoresis on a 1% agarose gel. The correct products for ligation into pET23a or retention in XpET22b will show a band at about 609 bp corresponding to the N-

termial region and no other DNA except the vector. Positive colonies were confirmed using DNA sequencing using T7 promoter and T7 terminator primers, performed by the ICMB Core Facility at the University of Texas at Austin.

2.1.2 Transformation of plasmid into competent cells

The tr4G in XpET22b and pET23a were transformed into *E. coli* competent cells (BL21 or Artic Express, Invitrogen??) using the procedure as described in **Appendix C**.

Begin by adding 2 µl of β-mercaptoethanol to 75 µl of Arctic Express competent cells on ice. Swirl each tube every 2 minutes for 10 minutes. Then, add 150 ng of plasmid DNA and incubate the mixture on ice for 30 minutes. During the incubation step, preheat SOC medium in a 42º water bath. Heat shock the tubes in the 42º water bath for 20-25 seconds and incubate the mixture on ice again for 2 minutes. Then, add 0.9 mL SOC and incubate the mixture at 37º for 1 hour, shaking. Finally, plate about 200 µl of the cell mixture on an LB ampicillin plate and incubate at 37º overnight.

2.2 Expression

The basic expression procedures are identical for both the tr4G/XpET22b and tr4G/pET23a are described in detail in **Appendix C**.

Several conditions of expression may be optimized such as the strain of *E. coli*, the culture media used, the concentration of IPTG, the temperature pre- and post-induction or the time the cells grow post-induction

Optimization of IPTG used for induction. To determine the optimal IPTG concentration for induction, a small-scale expression was performed on tr4G/pET23a in BL21 cells. Single colonies were used to innoculate six 250 mL LB cultures containing ampicillin (x mg/ml). The cultures were grown with shaking at 37ºC. The OD_{600} was checked periodically and when the OD_{600} reached 0.4-0.6, the cultures were induced with IPTG. Two cultures were each induced with 0.1 mM IPTG, 0.4 mM IPTG, and 0.5 mM IPTG. After induction, the cultures were incubated with shaking at 30ºC for 3 hours. The cells were harvested by centrifugation and stored at -80ºC. A small portion of the harvested cell pellet was analyzed on a Western blot using anti-His C-terminus-HRP antibody to determine if expression occurred.

Comparison of tr4G/XpET22b and tr4G/pET23a expression. The comparison of tr4G in pET23a and tr4G in XpET22b in BL21 cells was performed through side-by-side expression under the same conditions using LB media. After inoculation of six 1 L flasks at 37ºC shaking, at an OD_{600} between 0.4 and 0.6, the cultures were both induced to a final concentration of 0.5 mM IPTG. Post-induction, the flasks were incubated at 30ºC for 3 more hours before harvesting. A small portion of the harvested cell pellet was analyzed on a Western blot using anti-His C-terminus-HRP antibody to determine if expression occurred.

Expression in Artic Express cells. The tr4G /XpET22b was transformed into Arctic Express cells to look for improvement in tr4G expression levels. 5 mL LB cultures containing 20 μg/mL gentamycin were innoculated using single colonies. The cultures were incubated with shaking at 37ºC overnight. The 5 mL overnight cultures were used to

inoculate 1 L LB flasks (without antibiotics) shaken at 37ºC. The induction (x mM IPTG??) and harvesting of cells were the same as described above.

Media effects on expression. In order to compare the effects of using different types of media on tr4G expression levels 2XYT and Terrific Broth (TB) media were compared to LB media (see **Appendix A**) were also used. For the 2XYT media, the same procedure as with LB was used. For TB, the conditions were altered after induction. After induction, the flasks were switched to 28ºC and shaken for 18 hours.

2.3 Purification

2.3.1 Purification by His-tag on column

The bacterial cell pellet was thawed on ice and for each liter of culture used for expression, about 10 mL of binding buffer was used. The cells were completely resuspended in ice-cold binding buffer and kept on ice. This cell mixture was sonicated, making sure the cells do not become warm by keeping them on ice, at 3 x 30 seconds of 70%, followed by 2 x 30 seconds of 90% as described in **Appendix A**. The sonicated solution was centrifuged at 49,000 rpm (Ti50.2 rotor), for 35 minutes at 4ºC to remove debris and insoluble proteins. If inclusion bodies (insoluble protein) are present, a large amount of eIFtr4G protein will be present in the pellet. A portion of the pellet is retained for SDS gel analysis.

The supernatant from the centrifugation of 6 L of culture was applied to a 1 mL Ni-NTA Superflow Resin (QIAgen). The column was prepared by rinsing with 5-10X column

volume of binding buffer to establish a baseline UV reading. The supernatant was applied to the column very slowly. The column was washed with binding buffer until the UV reading returns to the baseline, or levels off to show that the proteins bound loosely to the column matrix were removed. The tr4G was eluted using elution buffer (see **Appendix A**) and 0.5 mL fractions were collected until 15-20X the volume of the column was applied. Each fraction was analyzed by the Bradford protein assay and peak fractions of protein were pooled. The pooled fractions are dialyzed against x ml of buffer N' + 0.1 M KCl buffer (see **Appendix A**) in 20 mm tubing at 4ºC overnight. The protein sample was aliquoted (sizes?) and stored at -80ºC. If the same column is to be used again it should be regenerated as described in **Appendix C**.

The pooled protein sample was analyzed for purity by 12.5 % SDS PAGE and western blotting as described in **Appendix C**.

The pooled protein was concentrated using Ambion spin columns in the JA-20 rotor, at 4ºC at a speed of 9,000 rpm at 15 minutes intervals until the desired level of concentration was achieved.

2.3.2 Removal of contaminants by ion exchange chromatography

Further purification of the tr4G was carried out using phosphocellculose (PC) and diethylaminoethanol (DEAE) ion exchange columns.

Initial small scale purification was performed (0.5 mL columns) to determine if contaminants and eIFtr4G may be separated based on ion exchange columnes. The columns were prepared by washing with N' +0.1 M KCl buffer (5-10x column volumes). The protein sample was applied and the column was washed with N' +0.1 M KCl buffer.

The protein was eluted using N' + 0.2 M KCl buffer (see **Appendix A**) and 0.25 mL fractions were collected. The resulting fractions were analyzed for purity by12.5% SDS-PAGE.

Further purification was tested using a 1 mL HisTrap HP FPLC column (Amersham) and elution with an imidazole concentration gradient. All buffers were filtered through a 0.2 μm membrane. The column was washed with 5X the column volume of de-ionized water at a rate of 1 mL/min and the flow rate held constant throughout the entire run. The column was equilibrated with binding buffer (~5 column volumes). The protein sample was injected and the column was then washed with 10-15X volumes of binding buffer. The protein was eluted using a gradient from 20 mM to 250 mM imidazole over 10 minutes, collecting 0.5 mL fractions. Any remaining proteins were removed by elution with 500 mM imidazole. The column was washed with 5-10X column volumes of de-ionized water, 5-10X 1.5 M NaCl, 5-10X de-ionized water, and then stored in 20% ethanol.

The purity was assessed by 12.5% SDS-PAGE and the fractions containing the highest purity tr4G were pooled and dialyzed overnight in N' + 0.1 M KCl at 4ºC. The protein was analyzed by Bradford for protein concentration, aliquoted and stored at -80ºC.

2.3.3 Analysis of protein by mass spectroscopy

The gel apparatus was cleaned with 20% SDS and de-ionized water and stored in clean bags under a chemical hood. Disposable materials, such as pipette tips, razor blades and gloves were brand new and changed each time it was exposed to the environment outside of the hood. All solutions must be filter sterilized. The protein sample was separated on a NuPAGE Novex Bis-Tris Gel (Invitrogen) using the XCell SureLock Mini-

Cell apparatus. The sample was prepared with NuPAGE sample buffer (4X), 1 µl 2M DTT, and de-ionized water to a total volume of 20 µl . This sample was heated at 75ºC for 10 minutes. The XCell gel apparatus was assembled and filled with NuPAGE running buffer. Prior to sample application, 0.5 mL NuPAGE Antioxidant Solution added to the front chamber. A protein molecular weight marker was also applied to the gel The gel was run at 200 V until the dye marker reach the bottom of the gel.

Precautions must be taken to prevent contamination of the sample, gel, buffer or apparausThe gel was taken out from the XCell apparatus and placed fixer solution provided with the NuPAGE kit, and shaken, for 20 minutes. The gel is transferred to 25 mL of NuPAGE Stain A, shakenfor 10 minutes. 2 mL of NuPAGE Stain B is added and shaken for an additional 30 minutes. The gel was destained and then shaken in de-ionized water for 20 minutes, or until the protein bands were clearly visible. Bands of interest are cut out, using a sterile razor blade, and placed in 1.5 mL eppendorf tubes. The bands were futher destained with 1 mL of 50% methanol for 20 minutes. This step of destain may be repeated if necessary to remove residual dye. The gel band was then dehydrated by removing the 50% methanol and addition of 1 mL acetonitrile for 5 minutes. This step was repeated with fresh acetonitrile for 10 minutes. The acetonitrile was removed and 0.5 mL of 100 mM DTT added for 30 minutes. The DTT was removed and 0.5 mL of 100 mM iodoacetamide added for 30 minutes. The iodoacetaamide was removed and 0.5 mL 100 mM ammonium bicarbonate added for 30 minutes. The aceonitrile step was repeated for 30 min and then all liquid removed. The gel band was air dried for 10-15 minutes. To the dried tube 20 µl trypsin (conc?) was added and placed on ice for 5 minutes. The trypsin was removed and 10 µl of 100 mM ammonium bicarbonate was added. The reaction was incubated overnight at 37ºC.

The reactions were centrifuged briefly to get the liquid and gel bands to the bottom of the tube and 10 µl of formic acid was added and incubation continued at room temperature for 15 minutes to inactivate the trypsin. A 15 µl aliquot of the solution was removed and placed into a clean 1.5 mL eppendorf tube. Add an additional 10 µl of formic acid to the sample to increase the sample volume. The protein sample prepared by this method was analyzed by mass spectroscopy for analysis (performed at the ICMB Core Facility at the University of Texas at Austin).

3 Results

Introduction of Results

The main problem in reaching the goal of developing eIF4G-specific antibodies was obtaining enough eIF4G protein. Because AteIF4G is a relatively large protein (180 kDa) and very prone to degradation (Browning 1996), it has been difficult to attain sufficient amounts of this protein. Previous work I have done has included protein expression of *Arabidopsis* eIF4F complex (both eIF4G and eIF4E together) and purification for antibody production. In achieving that goal, I faced many difficulties with the eIF4G component because of its low expression levels and propensity for degradation. We were able to increase the yield of full-length eIF4G by combining cells producing eIF4E and cells producing eIF4G before breaking open the cells. By mixing eIF4E and eIF4G expressing cells before the cells were lysed, when the proteins were released, they stabilized each other by presumably encouraging correct folding and overall stability. We were able to take advantage of the stabilizing effect of eIF4E binding to eIF4G, as well as the complex's increased affinity for the 5'-cap structure in the purification of the eIF4F complex from other proteins naturally produced in *E. coli* cells. The use of the affinity matrix, m^7GTP Sepharose, which mimics the cap and binds eIF4E, also retains any proteins bound to eIF4E, such as eIF4G. It is important to attain protein without a large amount of

contaminants, because when antibodies are made, they are made to all proteins in the sample provided. There is no way to predict the reactivity of the host animal in which the antibodies are raised to certain proteins over others. Therefore, with contaminants, one runs the risk that the contaminant may have much better antigenicity, making the majority of antibodies produced against the contaminant and not the protein of interest, in this case eIF4G. Thus, using an antibody for identification of a protein that contains a contaminant may produce false results or at the very least a high background of non-specific binding to the antibody preparation.

The biggest hurdle in this previous project was expressing eIF4F; even with the eIF4E-eIF4G stabilization, it was still difficult to obtain enough eIF4F complex. The main reason for this is thought to be the large size of the eIF4G, making it less likely that the translational machinery stays on the mRNA long enough to finish translating eIF4G into protein or that the protein itself is toxic to *E. coli*. Therefore, I tried to improve the situation using a co-expression approach. The eIFs4G gene and eIF4E gene were made into a dicistronic construct, where the two genes are adjacent to each other. Whereas in the previous procedure the eIF4G and eIF4E proteins were expressed in separate cells and then combined after cells were broken open, with the dicistronic construct, eIF4E and eIFs4G will be expressed in the same cell, one after the other. Because the two proteins will be expressed in such close proximity to each other, they should form the stabilized eIF4F complex more quickly, and degradation should be inhibited. In comparing the two different methods of eIF4F expression, there was not a significant difference in protein yield. However, in the end, this *Arabidopsis* eIF4F complex was successfully expressed at a very low level and enough protein obtained to produce eIF4F complex antibodies.

While the eIF4F antibody has been very useful, there is also a need for *Arabidopsis* eIF4G-specific antibody. The *Arabidopsis* eIF4F-specific complex, while it recognizes eIF4G protein, also recognizes eIF4E, making it inadequate for pull downs and identifying eIF4G. In developing eIF4G antibody, the same difficulties that were faced in eIF4F antibody production will apply; however, because the antibody will be specific to eIF4G, we cannot use eIF4E to stabilize and increase the protein yield of eIF4G. To address this need, instead of using a full-length eIF4G gene, I constructed a short version of eIF4G (eIFs4G) that consisted of the N-terminal half of eIF4G, producing a protein that is only 542 amino acids long. Theoretically, if the protein is smaller, there is a greater chance that the entire protein will be produced before the translational machinery falls off the mRNA, increasing protein yield and it is possible that a shorter version of eIF4G maybe less toxic to *E. coli*. Because we are using the eIF4F protein for antibody production, the eIF4G does not need to be full-length, as antibodies directed against a portion of the protein will still be able to recognize the full length eIF4G effectively.

3.1 Cloning eIFtr4G

Possible factors that could cause the consistently low expression levels of eIF4G include: 1. the plasmid vector, XpET22b; 2. the strain of *E. coli* is not optimal for expression; 3. the growth media and/or conditions for the bacteria is suboptimal; and 4. the toxicity of eIF4G to *E. coli*.

A truncated version of eIF4G, named truncated eIF4G (eIFtr4G), that consisted of the first 223 amino acids of the N-terminus of eIF4G was engineered by two methods. This

N-terminal region was selected because it is relatively un-conserved among different species, even plants, making the antibody specific to *Arabidopsis* eIF4G. This gene was specifically engineered to produce the eIFtr4G protein with a protein tag of six histidines on the C-terminal end to simplify purification on a Ni-NTA column.

Full-length eIF4G, originally in XpET22b, was digested in order to yield eIFtr4G. For its insertion into the pET23a vector, both the pET23a vector and eIF4G in XpET22b were digested with NdeI and HindIII restriction enzymes. As shown in *Figure 9a*, the restriction digests were separated on a DNA agarose gel. Lane 1 shows the DNA size marker while Lane 2 contains pET23a digest (NdeI/HindIII), with the linearized plasmid of 3427 bp. Lane 4 contains the eIF4G in XpET22b digest (NdeI/HindIII) and should contain DNA fragments of 609 bp, 3625 bp, and 5380 bp.

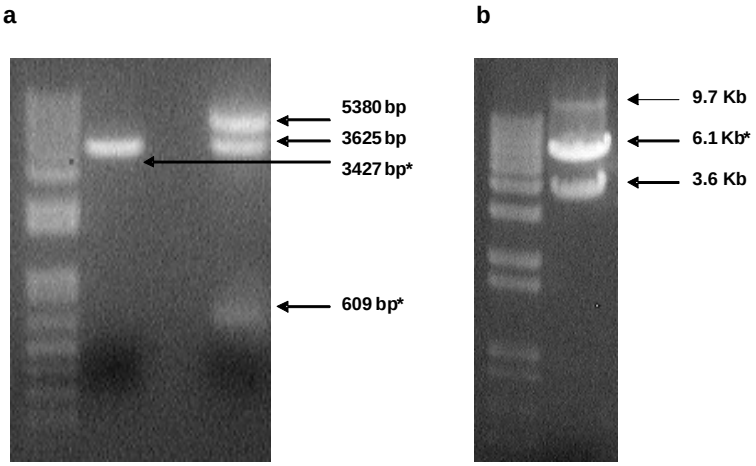

Figure 9. Restriction digest of 4G in XpET22b and pET23a vector(a) Restriction digest using NdeI and HindIII, for the insertion of eIFtr4G into pET23a. Lane 1: DNA marker. Lane 2: pET23a digest. Lane 3: empty. Lane 4: eIF4G in XpET22b digest, with the piece of interest starred. (b) Restriction digest using HindIII, for eIFtr4G in XpET22b. Lane 1: DNA marker. Lane 2: eIFtr4G in XpET22b digest, with the piece of interest starred.

The desired fragments pET23a (3427 bp) and eIFtr4G (609 bp) were recovered from the gel and ethanol precipitated. The DNA sequence was confirmed.

For the creation of eIFtr4G in XpET22b, full-length eIF4G in XpET22b was digested using HindIII restriction enzyme. In *Figure 9b*, the results of the restriction digest are shown, for eIF4G in XpET22b digested with Hind III. The digest shows 3 bands: undigested full-length eIF4G in XpET22b of about 9.7 Kb, the front third of eIF4G in XpET22b at about 6.1 Kb, and the last two-thirds of eIF4G at about 3.6 Kb. The desired band is the one containing the first third of eIF4G (6.1 Kb). This DNA fragment was then self-ligated and sequenced in order to confirm the sequence of eIFtr4G. The eIFtr4G final protein will have the amino acid sequence and properties shown in **Appendix B**.

3.2 Finding the optimal IPTG concentration for induction

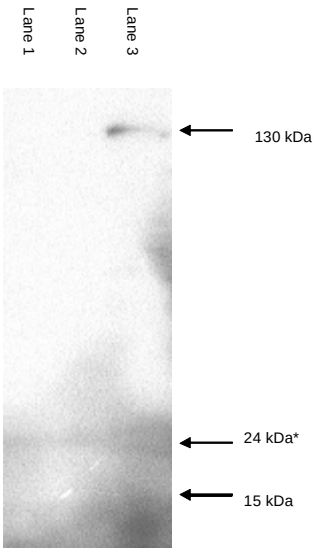

Figure 10. IPTG concentration induction test. This Western blot, performed on crude culture samples, used an Arabidopsis recombinant eIF4F primary antibody for eIFtr4G protein identification (eIFtr4G in XpET23a in BL21 cells, raised in LB Broth media). Lane 1: [IPTG] = 0.1mM. Lane 2: [IPTG] = 0.2mM. Lane 3: [IPTG] = 0.3mM.

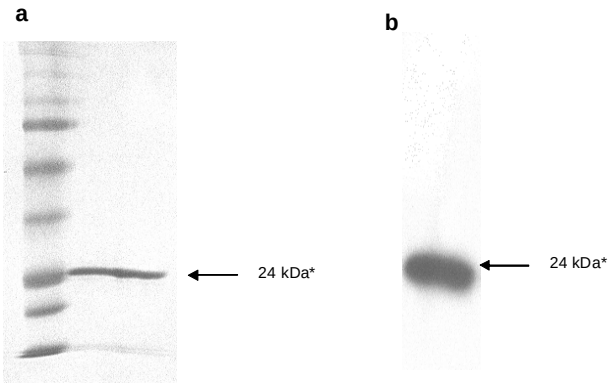

Figure 11. Pooled IPTG induction test. eIFtr4G in pET23a in BL21 cells, expressed in LB Broth media. (a) sample on a 12.5% SDS page gel, showing the major band at about 24 kDa, where eIFtr4G should be. (b) sample analyzed by Western blot, showing a single band at about 24kDa as well.

A preliminary protein expression was performed, using the eIFtr4G/pET23 construct to determine the optimum concentration of IPTG for induction. Cultures were inoculated with eIFtr4G/pET23a in BL21 competent cells according to the protocol described in the methods section. Samples were taken after 3 hours of induction and used to determine the best IPTG concentration through Western blot (see *Figure 10*). Western blot was performed using the rabbit anti-Arabidopsis recombinant eIF4F antibody. From this Western blot, one may see that with increasing IPTG concentration an increasing amount of protein is produced. Since the 0.5 mM IPTG concentration yielded the most eIFtr4G protein, it can be concluded that 0.5 mM IPTG is the best condition for induction. Thus, in all future preparations, induction was performed with 0.5 mM IPTG.

To confirm the expression of AteIFtr4G, a protein extract was prepared and purified on a Ni-NTA column as described in the methods section. Peak fractions were collected and combined. The pooled fractions were analyzed on 12.5% SDS-page gel (*Figure 11a*) as well as by Western blot, using an anti-His C-terminal-HRP antibody, as shown in *Figure 11b.* In the protein gel, the single clear band at 24 kDa indicates that a protein of the correct size of eIFtr4G was purified on the Ni-NTA column. In order to confirm the band

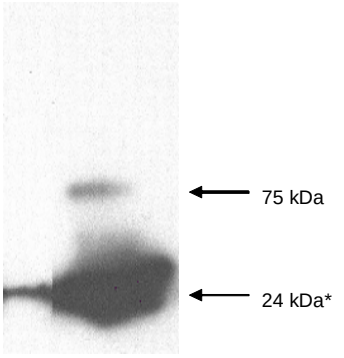

Figure 12. Test for inclusion bodies during expression. eIFtr4G in XpET22b was expressed in BL21 cells. Pellet and pooled fractions were compared by Western blot analysis, probed with anti-His C-terminus-HRP antibody. Lane 1: crude cell pellet. Lane 2: pooled fractions.

as eIFtr4G, the Western blot was performed, probing for proteins that contain a His-tag. The clear, single band on the Western blot confirms that the band seen on the protein gel was in fact eIFtr4G. Using Bradford analysis, the total protein content of the pooled fractions was approximately 0.1186 mg from 750 mL of culture. This yield is low and it is possible that the majority of the expression protein may be present in insoluble material.

3.3 Test for inclusion bodies during protein expression

In order to assess whether there is formation of insoluble protein in inclusion bodies, a 3 L LB Broth culture was inoculated (using eIFtr4G in XpET22b in BL21 cells) and induced with 0.5 mM IPTG. According to the methods described above, the bacteria was harvested and the protein was purified. A crude sample of bacterial pellet was taken after harvest. This pellet sample was analyzed alongside the pooled purified fractions in a Western blot, using the anti-His C-terminal-HRP antibody, as shown in *Figure 12*. In the crude pellet sample, a small amount of eIFtr4G protein was detected, while, in the pooled peak fractions, a very large band indicating presence of eIFtr4G protein and a band showing

a protein of about 75 kDa were detected. Therefore, it is possible that there are a small amount of inclusion bodies, but it is more likely that some cells were not completely broken open during sonication, so that some protein was trapped inside cells. Bradford analysis showed that the pooled fractions contained 1.894 mg of protein from 3 L of cells.

3.4 Comparing AteIFtr4G expression levels in different vectors

One of the possible reasons eIF4G protein expression levels are so low could be due to the alteration of the pET22b vector, to XpET22b. Some undesired mutation may have occurred during the alternation which could interfere with the ability to promote robust translation. To test this, comparisons were made between expression levels of eIFtr4G in XpET22b and in pET23a, an unaltered vector. Identical conditions were used: LB Broth media in six 1 L cultures each, induction with a total IPTG concentration of 0.5 mM for 3 hours at 30ºC, and harvesting and purification according to the basic procedure described in the methods section. Peak fractions from both preparations were pooled separately and analyzed using Bradford analysis. The preparation using eIFtr4G in XpET22b contained 2.284 mg, while the one using eIFtr4G in pET23a yielded 3.893 mg.

From these results, there is an obvious improvement when eIFtr4G gene was inserted into an unaltered vector, pET23a. However, while this direct comparison shows eIFtr4G in pET23a produced more protein than when in XpET22b, one could also look at the previous induction test, which used eIFtr4G in pET23a, and the inclusion body test, using eIFtr4G in XpET22b. In the induction test preparation, 0.1186 mg was purified from 1.5 L of culture (0.08 mg/L) using pET23a, while in the inclusion body test, 1.894 mg was

obtained from 3 L of culture (0.65 mg/L) using XpET22b. From this comparison, eIFtr4G in XpET22b was much more productive. This discrepancy from the direct comparison could be due to several factors: 1) inexperience because the induction test was the first preparation done, 2) the induction test was induced at a higher OD_{600}, when less cells were in a productive state, and 3) the difference due to the use of 250 mL cultures for the induction test versus 1 L cultures for the inclusion body test. Based on these inconsistencies between the two tests, the direct comparison, done side-by-side, gives a much more representative picture of which vector results in better protein expression levels. Therefore, future preparations all used the eIFtr4G in pET23a construct. However, it should be noted that the difference is quite small and there is likely little or no difference in the constructs.

3.5 Comparing AteIFtr4G expression using different *E. coli* strains

E. coli BL21 cells are traditionally used for protein expression in the Browning laboratory and for most proteins good expression is obtained. *E. coli* Arctic Express cells (Novagen) are engineered to express at cold temperatures (4ºC). At this temperature, proteins will be expressed slower, but the cell also produces cold-loving chaperonins, which help to fold proteins into the correct, stable configuration, and rare tRNAs, which reduces pausing of the translational apparatus, reducing its dissociation from the mRNA. Thus, although the Arctic Express cells produce protein at a lower rate, they should ultimately produce more correctly folded and stable protein. This strain will be used to determine if the production of eIFtr4G may be improved.

Six 1 L cultures of LB Broth were inoculated with eIFtr4G in XpET22b in Arctic Express cells. Incubation and expression conditions were as described by the manufacturer for Arctic Express cells (Novagen). The cells were harvested and the protein purified by Ni-NTA column. The pooled peak fractions were run on a 12.5% SDS-page gel to check for eIFtr4G protein expression. However, there was no detectable eIFtr4G present. This effect could have been due in part to either the cold temperature, which may have lowered already low protein expression levels or the problem with eIFtr4G protein expression is not due to misfolding. Therefore, Arctic Express *E. coli* cells did not improve eIFtr4G protein expression in comparison to BL21 cells.

3.6 Comparing AteIFtr4G expression in different growth media and temperature

The growth media was varied to find the conditions that encourage production of the most eIFtr4G protein. The most basic growth medium is LB, which contains all necessary nutrients for bacterial cells to grow. Another medium that will be tried will be 2xYT medium, which contains the minimal nutrients for cell growth, thus slowing down expression, but increasing folding time for proteins following expression to increase stability. Also, Terrific Broth (TB) will be tested, which provides a very robust environment for bacteria with excessive nutrients, to encourage an increase in global protein expression levels.

To compare Terrific Broth media expression versus that observed in LB Broth media, the same volume to surface area ratio was maintained by using six 1 L media in 6 L flasks were used in all preparations. The eIFtr4G in pET23a construct in BL21 cells were

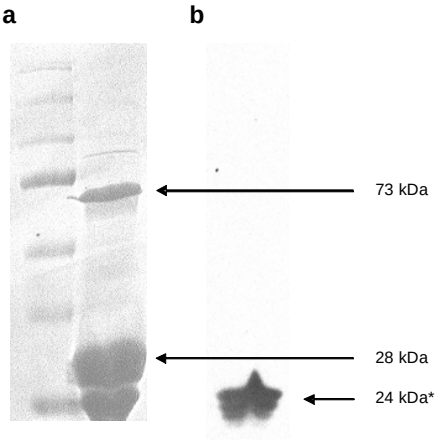

Figure 13. eIFtr4G in pET23a was expressed in BL21 cells, using Terrific Broth media and modified induction conditions . (a) Peak fractions were pooled and run on a 12.5% SDS - page gel. Lane 1: Protein molecular weight marker. Lane 2: peak fractions show the presence of eIFtr4G protein in addition to 2 contaminants. (b) Peak fractions were analyzed by Western blot, probed with anti-His C-terminus-HRP antibody. Lane 1: peak fractions, identify the 24 kDa protein as the eIFtr4G protein and the 73 kDa and 28 kDa proteins as contaminants.

used for expression. Induction of expression in TB broth was a final IPTG concentration of 0.5 mM, at 28ºC overnight, for about 18 hours. Purification of protein was performed as described for a basic Ni-NTA column. The pooled fractions were analyzed on a 12.5% SDS-page gel (*Figure 13a*). The gel revealed significant levels of contaminants in addition to eIFtr4G, one of which is similar in size to eIFtr4G. To distinguish the protein of interest from contaminants, the sample was analyzed with Western blot, probed with anti-His C-terminal-HRP antibody (*Figure 13b*). This Western blot revealed that eIFtr4G protein was the smaller protein in the doublet, running as a 24 kDa protein, while the two other protein bands, ran at about 28 kDa and 73 kDa.

These contaminant bands were not seen in preparations using LB Broth media and an induction temperature of 30ºC. The appearance of these bands could be due to the overall, large increase in global protein expression. Using Bradford analysis, the peak fractions of this preparation yielded 18.970 mg of protein. However, one must also consider that a large portion of this protein consists of contaminants. While these proteins could have also been present in LB broth preparations, they may not have been visible with that level of protein expression. In order to be able to assess by how much Terrific Broth and the new overnight

induction condition improve eIFtr4G protein expression, purification of the protein of interest from contaminants was necessary.

3.7 Purification of AteIFtr4G

To separate the contaminants from eIFtr4G protein, we assumed that they had different binding affinity for the Ni-NTA resin. Because the eIFtr4G protein is engineered to have six His residues, it should bind the resin with the greatest affinity. The peak fractions were applied over a Ni-NTA column, which was then washed with 20X column volumes of binding buffer, making sure that all proteins not bound to the matrix would be washed out. Then, instead of proceeding to the normal elution step, a second wash was applied, binding buffer containing 50 mM imidazole for 15X column volumes, intended to wash off proteins bound loosely to the matrix, leaving those that bind very tightly to the matrix. Then, elution using the elution buffer, with 250 mM imidazole, bumps off all remaining proteins on the column. From each step, peak fractions were pooled, and then run on a 12.5% SDS-page gel (*Figure 14*). In this gel, the 20 mM imidazole wash eliminated some of the 28 kDa and 73 kDa contaminants. Therefore, one of the reasons that there was so much contamination in the initial purification was due to insufficient washing at low imidazole after the extract was applied to the column. The 50 mM imidazole wash eliminated the majority of the remaining 28 kDa and 73 kDa proteins, as well as some of the eIFtr4G protein (24 kDa). The eluted protein from the 250 mM imidazole wash contained mostly eIFtr4G and small amounts of 28 kDa and 73 kDa proteins. While it is true that

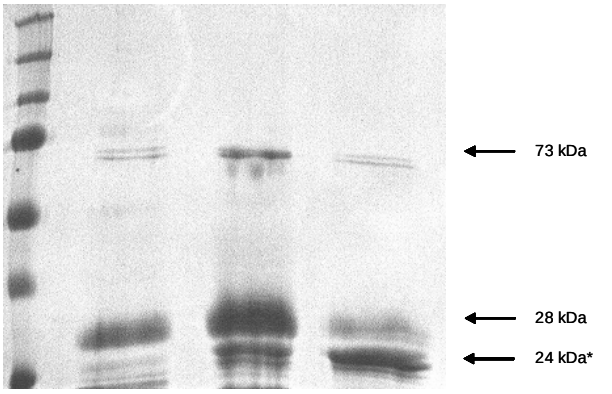

Figure 14. Separation of contaminants from eIFtr4G protein over a Ni -NTA column. After three successive washes, peak fractions were pooled and run on a 12.5% SDS -page gel. Lane 1: protein molecular weight marker. Lane 2: peak fractions following a wash with binding buffer (20 mM imidazole). Lane 3: peak fractions following a wash with binding buffer (50 mM imidazole). Lane 4: peak fractions following a wash with elution buffer (250 mM imidazole).

eIFtr4G had the greatest affinity for the Ni-NTA matrix, there was not a clear separation between eIFtr4G and the contaminanting proteins; about one-third of eIFtr4G protein is also lost at the concentration of imidazole that removes most of the contaminants.

The relatively pure protein sample obtained from the elution step contained 7 mg of protein. Thus, there is a marked increase in eIFtr4G protein expression when expressed in Terrific Broth and using overnight induction when compared to LB Broth. However, in this 7 mg, there is still ~20% contaminanting proteins.

A second protein preparation identical to the one described above using Terrific Broth was performed, and after the two purification steps on the Ni-NTA column, Bradford assay showed ~7 mg of protein. This sample was run on 12.5% SDS-page gel (*Figure 15*) and results were consistent with those seen previously, the same contaminants were found at about the same levels as previously observed.

While the additional purification of peak fractions on a Ni-NTA column reduced contaminant levels, further purification was still needed. One way that they could be further separated could be based on the binding affinity of the desired protein (eIFtr4G) or the contaminants to either PC or DEAE based on their charge. PC columns purify proteins by.

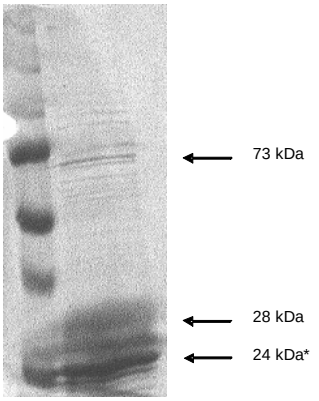

Figure 15. eIFtr4G in pET23a was expressed in BL21 cells, using Terrific Broth media and modified induction conditions . Peak fractions were pooled and run on a 12.5% SDS -page gel. Lane 1: Protein molecular weight marker. Lane 2: peak fractions show the presence of eIFtr4G prot ein in addition to 2 contaminants seen previously.

binding positively charged proteins. AteIFtr4G could potentially, selectively bind PC as the protein's predicted isoelectric point (pI) is basic, 9.82. On the other hand, AteIFtr4G could also selectively bind DEAE as it contains a string of acids on the N-terminal end of the protein. If the contaminants do not bind either PC or DEAE and the AteIFtr4G does, then the contaminants will be removed in the wash. Conversely, if the contaminants bind and AteIFtr4G does not, then the desired protein will be recovered in the run through of the column.

Small scale purifications were performed on both types of columns and comparisons were made between the run through, the re-equilibrating wash, and the elution fractions on a 12.5% SDS-page gel (*Figure 16*). The PC and DEAE column run-throughs contained most of the eIFtr4G and contaminant proteins, while the remaining protein was found in the re-equilibration wash. Therefore, neither eIFtr4G nor the contaminant proteins showed differential binding affinity for PC or DEAE.

Since there is no way to separate eIFtr4G and contaminants by ion exchange chromatograpy, the Ni binding affinity column will be used with a gradient to try and separate the proteins based on their different affinities. To get a better separation of contaminants and eIFtr4G protein, an imidazole gradient, from 20 mM to 250 mM was

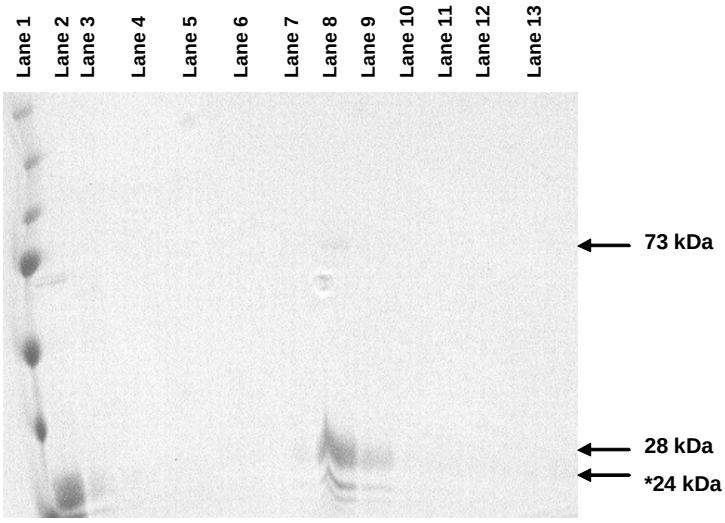

Figure 16. Separation of contaminants from eIFtr4G protein over PC and DEAE columns. Fractions collected during loading, wash and elution were collected and run on a 12.5% SDS-page gel. Lane 1: protein molecular weight marker. Lane 2: PC column run through. Lane 3: PC column wash. Lanes 4-7: PC column elution. Lane 8: DEAE run through. Lane 9: DEAE wash. Lanes 10-13: DEAE elution fractions.

applied over a HisTrap HP column using FPLC, as described in the methods. The ~14 mg of protein from the combination of two preparations grown in Terrific Broth were applied over the column. In *Figure 17a*, the UV analysis of fractions during elution is graphed along with the imidazole gradient. The peaks in the UV reading correlate to the specific protein coming off by matching the corresponding fractions to the protein seen on a protein gel. The peak fractions were immediately run on a 12.5% SDS-page gel (*Figure 17b*) to determine which fractions contained the purest eIFtr4G. The peak fractions were pooled and dialyzed. The first, small peak on the UV graph, corresponds to the first protein that came off of the FPLC column, the 73 kDa protein with little overlap with eIFtr4G. The large, second peak corresponds to both the 28 kDa contaminant as well as eIFtr4G. Over the course of the gradient, the gel shows that while the contaminant binds less well to the column than eIFtr4G, there is a large amount of overlap. Therefore, one has to sacrifice a large portion

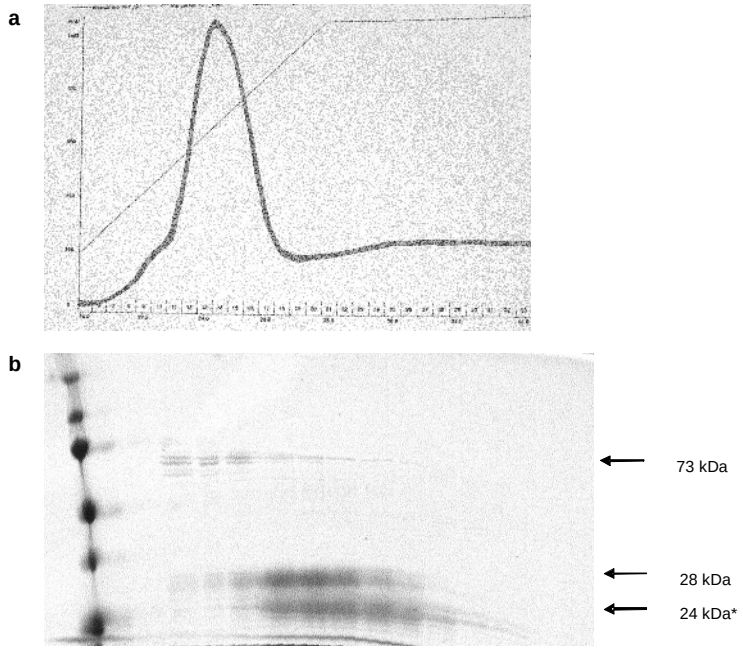

Figure 17. HisTrap HP FPLC column using an imidazole gradient. eIFtr4G in pET23a in BL21 cells were raised using Terrific Broth media. (a) UV reading was graphed against fraction number during elution. The black line indicates the imidazole gradient, from 20-250 mM, while the blue line indicates the UV reading. Each bump represents a different protein's elution. These fractions were run on a 12.5% SDS-page gel to determine the identity of the protein being eluted. (b) Elution fractions were run on a 12.5% SDS-page gel. Lane 1: protein molecular weight. Lane 3: run through. Lane 4-14: Peak fractions.

of the total eIFtr4G in the sample for purity. The most pure fractions were pooled and Bradford assay indicated a total of ~5 mg of protein.

3.8 Mass Spectroscopy

This new purification step resulted in a much cleaner protein sample, containing predominantly eIFtr4G protein; however, before raising antibodies it was desirable to confirm that the major band was eIFtr4G and to determine the identity of the remaining

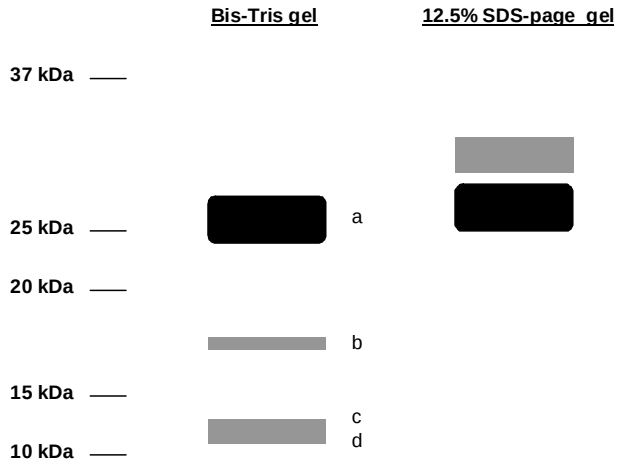

Figure 18. Protein banding patterns on Bis-Tris gels versus 12.5% SDS-page gels. A Bis-Tris gel was required for mass spectroscopy and showed a different banding pattern than usually seen on 12.5% SDS-page gels. Each band was analyzed using mass spectroscopy and identities were determined. (a) eIFtr4G. (b) eIFtr4G degradation products. (c) E. coli secB and eIFtr4G degradation products. (d) eIFtr4G degradation products.

contaminants. To prepare the protein for mass spectroscopy, it was run on a NuPAGE Novex Bis-Tris gel and prepared as per the methods described above. When run on the Bis-Tris gel, a different banding pattern was seen when compared to previous 12.5% SDS-page gels. A cartoon of the banding patterns is shown in *Figure 18*. This is not entirely surprising as proteins may run differently on different gels due to different interactions with the gel and chemical environment. Results indicated that the major band (a) was in fact eIFtr4G protein, the band just below (b) was a degradation product of eIFtr4G, the next minor band was *E.coli* secB (c)protein with trace amounts of eIFtr4G, and the last band (d) was another degradation product of eIFtr4G. Therefore, while there is a confirmed contaminant, the overall amount of secB in comparison to the amount of eIFtr4G protein is insignificant and should not affect the overall quality of the antibody produced.

The mass spectroscopy results showed that only a very small portion of the final protein sample contained contaminant protein. The very faint band on the Bis-Tris gel that represented this contaminant protein did not run in the same pattern that was normally seen on a 12.5% SDS-page gel (refer to *Figure 18*). There are four bands seen on the Bis-Tris gel, while there are only two bands on the 12.5% SDS-page gel, indicating that several bands that usually run together on the 12.5% SDS-page gel are separated out on the Bis-Tris gel. In trying to determine which band on the 12.5% SDS-page gel corresponded to *E. coli* secB protein, I considered several factors. Looking at *Figure 13*, by performing the Western blot alongside the 12.5% SDS-page gel, one may determine that the 28 kDa protein is a contaminant, as it does not bind with anti-His C-terminus HRP antibody. Also, it cannot be a degradation product of eIFtr4G as it is larger in size than the band that does associate with the antibody. From mass spectroscopy results, one knows that only eIFtr4G and *E. coli* secB proteins are present in the sample. Thus, the only possible identity of the 28 kDa band is secB protein, while the eIFtr4G band contained full-length eIFtr4G and its degradation products. However, it was somewhat unexpected as the secB band seen on the Bis-Tris gel was considerably more faint than the 28 kDa band on the 12.5% SDS-page gel.

4 Discussion

From the results obtained, out of the conditions tested, the expression levels maximizing eIFtr4G protein expression were found to be 1) a high IPTG concentration, 0.5 mM; 2) a gene construct which used an unaltered vector, in this case pET23a; 3) using BL21 *E. coli* cells; 4) raising cultures in Terrific Broth; and 5) inducing cells overnight at a temperature of 28°C. It was confirmed that AteIFtr4G protein does not form inclusion bodies in solution, thus eliminating this as a possible factor in low protein yield.

While there is basal expression of AteIFtr4G protein at low concentrations of IPTG during induction, results clearly showed a marked increase in expression levels as IPTG concentration was increased. Although this increased IPTG concentration increased eIFtr4G protein expression, it also increased global *E. coli* protein expression, increasing the amount of contaminant proteins. While this is a drawback, it is outweighed by the benefit of increasing tr4G protein expression, reducing the number of protein preparations that would have to be done to accumulate the same amount of protein. Then, almost all of the contaminants can be removed through an extra purification step. If a similar obstacle arises in the future in which protein expression is very low, increasing IPTG concentration could help. Conversely, if protein expression levels are high, with a high level of contaminant

protein expression as well, one may consider attempting to reduce IPTG concentration during induction for more pure protein expression.

Because there was a modest improvement of protein expression levels when the AteIFtr4G gene was in pET23a, as compared to XpET22b, one may conclude that deleting the XbaI restriction site in pET22b may have compromised the vector's overall functioning. Knowing this, better expression of full-length AteIF4G protein may be possible if the gene was excised and inserted into a new vector. This general principle could also be applied to other proteins that are troublesome to express; placing the gene in a fresh vector that has not been altered may eliminate the interference of a reworked vector on protein expression.

Another hypothesized cause of low AteIFtr4G protein expression was the misfolding and degradation of protein. Therefore, to test this, I compared the differences in protein yield using two different *E.coli* strains: BL21 and Arctic Express. BL21 is a standard E. coli strain used in protein expression, while the Arctic Express strain is special in that it encodes cold-loving chaperonins that help the folding of the protein being expressed during induction at 4°C. Theoretically, using the Arctic Express cells should decrease the amount of contaminanting proteins, while helping the stability of the protein of interest. Results indicated that using Arctic Express cells did not show any improvement in AteIFtr4G protein expression and actually decreased the amount of protein produced. This outcome may imply that AteIFtr4G does not have much difficulty with protein folding. This would not be surprising as eIFtr4G is not an especially large protein. However, one may not conclude that this is case as the temperature change to 4°C could have played a role in the initial difficulty of protein expression, which would also not be surprising as eIF4G and eIFtr4G proteins seem to be fragile and particular about the condition and environment during protein expression.

When media and induction temperature and duration were varied, a large difference in protein expression level was seen. Growing cultures in Terrific Broth at an induction temperature of 28°C overnight, versus using LB Broth at an induction temperature of 30°C for 3 hours, resulted in a very noticeable increase in global protein expression. Using 6 L of LB Broth to raise eIFtr4G in pET23a at standard induction procedures yielded 3.893 mg of relatively pure protein. Using 6 L of Terrific Broth to raise the same construct at modified induction procedures yielded 7.032 mg in one preparation and 6.910 mg in another. Therefore, using Terrific Broth and the modified induction environment produced about a 180% increase in eIFtr4G protein expression. Since this study's main goal was to increase eIFtr4G protein expression, temperature, duration and media were all altered from standard procedure at one time. As a group of changes, an obvious effect was seen. However, future studies could further elucidate the effects of each modification individually, testing whether a single modification caused most of the optimization, or whether the factors worked cooperatively to increase protein expression levels.

Increasing global protein expression using modified media and induction conditions causes the same problem seen with increasing IPTG concentration for induction earlier, in that while an increase in the protein of interest is seen, an increase in contaminant proteins are also observed. For AteIFtr4G, however, this problem was solved by the addition of a FPLC purification step using an imidazole gradient on a Ni column.

Throughout the purification of AteIFtr4G protein from contaminants, when run on a 12.5% SDS-page gel, 2 contaminant bands were seen at 28 kDa and 73 kDa. All three bands showed similar binding affinity to Ni-NTA matrix, as shown through their overlapping elution when using 20 mM, 50 mM, and 250 mM imidazole washes (see *Figure 14*). These three elution steps also showed that the 73 kDa protein had the lowest binding affinity, then

the 28 kDa protein, and eIFtr4G had the greatest affinity. This is expected as the eIFtr4G protein was engineered to contain a His-tag at the C-terminal end of the protein, while the contaminant proteins should not.

One way to purify eIFtr4G was to separate it from the contaminant proteins based on charge. This was attempted through the use of PC, which is negatively charged, and DEAE, which is positively charged. However, results showed that neither eIFtr4G nor the contaminant proteins bound either PC or DEAE.

The contaminants were successfully removed by the use of a HisTrap-HP FPLC column, applying an imidazole gradient to create clearer separation of eIFtr4G and contaminants. By applying the imizadole gradient, the 73 kDa contaminant was removed, while the 28 kDa and eIFr4G proteins eluted together, suggesting very similar binding affinity. Fractions were selected that contained the most eIFtr4G, even though a large portion of the total eIFtr4G protein was lost in fractions that also contained considerable amount of the 28 kDa contaminant. The pooled fractions were then analyzed by mass spectroscopy and were determined to be mostly eIFtr4G, with trace amounts of *E. coli* secB protein. This sample, containing ~5 mg AteIFtr4G protein was used to raise antibodies in rabbits.

Developing eIF4G antibody will provide the tools to probe for levels of eIF4G in knockout eIF4G and eIF(iso)4G *Arabadopsis* plants, or to perform immunoprecipitation, to pull down proteins or RNA complexes that are associated with eIF4G. The strategies used in this study to increase protein expression could be applied to other proteins that are also difficult to express. Finally, the results of these projects may contribute to continuing and future projects that cannot yet be foreseen, but will undoubtedly clarify the roles of eIF4F and eIF(iso)4F, in addition to the translational machinery apparatus in general in plants.

5 References

1. Borman, Andrew M. (2002). Free Poly(A) Stimulates Capped mRNA Translation *in Vitro* through the eIF4G-Poly(A)-binding Protein Interaction. *The Journal of Biological Chemistry* **277**, 36818-36824.

2. Browning, Karen S. (1996). The plant translational apparatus. *Plant Molecular Biology* **32**, 107-144.

3. Craig, Andrew W. B. (1998). Interaction of polyadenylate-binding protein with the eIF4G homologue PAIP enhances translation. *Nature* **392**, 520-523.

4. De Gregorio, Ennio et al. (1999). Translation driven by an eIF4G core domain *in vivo*. *The EMBO Journal* **18**, 4865-4874.

5. Duprat, Anne et al. (2002). The *Arabidopsis* eukaryotic initiation factor (iso)4E is dispensable for plant growth but required for susceptibility to potyviruses. *The Plant Journal* **32**, 927-934.

6. Gazo, Brandy M. et al. (2004) A Novel Interaction of Cap-binding Protein Complexes Eukaryotic Initiation Factor (eIF) 4F and eIF(iso)4F with a Region in the 3'-Untranslated Region of Satellite Tobacco Necrosis Virus. *The Journal of Biological Chemistry* **297**, 13584-13592.

7. Han, Baoguang and Jian-Ting Zhang (2002). Regulation of Gene Expression by Internal Ribosome Entry Sites or Cryptic Promoters: the eIF4G Story. *Molecular and Cellular Biology* **22**, 7372-7384.

8. Imataka, Hiroaki et al. (1997). A new translational regulator with homology to eukaryotic translation initiation factor 4G. *The EMBO Journal* **16**, 817-825.

9. Kawaguchi, Riki and Julia Bailey-Serres (2002). Regulation of translational initiation in plants. *Current Opinion in Plant Biology* **5**, 460-465.

10. Marissen, Wilfred E. and Richard E. Lloyd (1998). Eukaryotic Translation Initiation Factor 4G Is Targeted for Proteolytic Cleavage by Caspase 3 during Inhibition of Translation in Apoptotic Cells. *Molecular and Cellular Biology* **18**, 7565-7574.

11. Monzingo, Arthur F. et al. (2007). The Structure of Eukaryotic Translation Initiation Factor-4E from Wheat Reveals a Novel Disulfide Bond. *Plant Physiology* **143**, 1504-1518.

12. Morino, Shigenobu et al. (2000). Eukaryotic Translation Initiation Factor 4E (eIF4E) Binding Site and the Middel One-Third of eIF4GI Constitute the Core Domain for Cap-Dependent Translation, and the C-Terminal One-Third Functions as a Modulatory Region. *Molecular and Cellular Biology* **20**, 468-477.

13. Morley, Simon J. et al. (1997). eIF4G : Translation's mystey factor begins to yield its secrets. *RNA* **3**, 1085-1104.

14. Prévôt, Déborah et al. (2003) Conducting the initiation of protein synthesis: the role of eIF4G. *Biology of the Cell* **95**, 141-156.

15. Robaglia, Christophe and Carole Caranta. (2006). Translation initiation factors: a weak link in plant RNA virus infection. *Trends in Plant Science* **11**, 40-45.

16. Ruud, Kelley A, et al. (1998) Identification and Characterization of a Novel Cap-binding Protein from *Arabidopsis thaliana*. *The Journal of Biological Chemistry* **273**, 10325-10330.

17. Von der Haar, Tobias et al. (2000). Stabilization of Eukaryotic Inititaion Factor 4E Binding to the mRNA 5'-Cap by Domains of eIF4G. *The Journal of Biological Chemistry* **275**, 30551-30555.

Appendix A

Media

LB Broth

Per liter:

20 g of LB Broth base and filled to a total volume of 1 L with deionized H_2O. Autoclave sterilization for 20 minutes at 15 lb/sq. in. on liquid cycle.

2XYT Broth

40 g of LB Broth base and filled to a total volume of 1 L with deionized H_2O. Autoclave sterilization for 20 minutes at 15 lb/sq. in. on liquid cycle.

Terrific Broth (TB)
(Tartof and Hobbs 1987)

Per liter:

To 900 mL of deionized H_2O, add:

Bacto-tryptone	12 g
Bacto-yeast extract	24 g
Glycerol	4 mL

Autoclave sterilization for 20 minutes at 15 lb/sq. in. on liquid cycle. Allow solution to cool and add 100 mL of sterile solution of 0.17 M KH_2PO_4, 0.72 M K_2HPO_4.

Column Buffers

Binding buffer

Per 150 mL:

50 mM HEPES, pH 7.6	7.5 mL of 1 M
0.6 M KCL	6.71 g
20 mM Imidazole	204 mg
Protease inihibitor cocktail **no EDTA*** (Roche)	1 Complete tablet (or 5 mini-Complete tablets)
deionized H_2O	fill to 150 mL

*no EDTA is necessary because EDTA is a chelator of metal ions and would strip Ni from the column.

Elution buffer

Per 50 mL:

20 mM HEPES, pH 7.6	1.0 mL of 1 M
100 mM KCL	0.37 g
250 mM Imidazole	0.85 g
deionized H_2O	fill to 50 mL

Strip buffer

Per 50 mL:

80 mM HEPES, pH 7.6	4.0 mL of 1 M
0.6 M KCL	2.24 g
100 mM EDTA	1.86 g
deionized H_2O	fill to 50 mL

Charging buffer

Per 50 mL:

50 mM $NiSO_4$	0.657 g
deionized H_2O	fill to 50 mL

N' +0.1 M KCl

Per liter:

10% glycerol	100 mL
20 mM HEPES, pH 7.6	20 mL of 1 M
0.1 M KCL	7.46 g

100 mM EDTA	200 µl 0.5 M
2 mM DTT	1 mL of 2M
deionized H_2O	fill to 1 L

N' +0.3 M KCl

Per liter:

10% glycerol	100 mL
20 mM HEPES, pH 7.6	20 mL of 1 M
0.1 M KCL	22.38 g
100 mM EDTA	200 µl 0.5 M
2 mM DTT	1 mL of 2M
deionized H_2O	fill to 1 L

Western Blot Buffers

Anode 1 buffer

Per 250 mL:

300 mM Tris-HCl, pH 10.4	30 mL of 2.5 M
0.05% SDS	1.25 mL of 10%
10% methanol	25 mL
10 mM BME	1 µl
deionized H_2O	fill to 250 mL

Anode 2 buffer

Per 250 mL:

25 mM Tris-HCl, pH 10.4	2.5 mL of 2.5 M
0.05% SDS	1.25 mL of 10%
10% methanol	25 mL
10 mM BME	1 µl
deionized H_2O	fill to 250 mL

Cathode buffer

Per 250 mL:

25 mM Tris-HCl, pH 9.4	6.25 mL of 1.0 M
40 mM α-amino caproic acid	1.2 g
0.05% SDS	1.25 mL of 10%

10% methanol	25 mL
10 mM BME	1 µl
deionized H_2O	fill to 250 mL

Appendix B

Arabidopsis eIFtr4G: amino acid sequence and protein properties

4G N-terminal/pET23a+ fusion protein

Analysis	Entire Protein
Length	215 aa
Molecular Weight	23061.04 m.w.
1 microgram =	43.363 pMoles
Molar Extinction coefficient	8960
1 A[280] corr. to	2.57 mg/ml
A[280] of 1 mg/ml	0.39 AU
Isoelectric Point	9.82
Charge at pH 7	8.16

```
   1  MHQAQGHGFA TPMGAQIHPQ LGHVGVGLSP QYPQQQGGKY GGARKTTPVK
  51    ITHPDTHEEL RLDRRGDPYS EGDSTALKPH SNPPPRSQPV SSFAPRPVNL
 101    VQPSYNSNTM IYPPVSVPLN NGPMSSAQAP RYHYPVIDGS QRVQLINQPA
 151    HTAPQLIRPA APAHLSSDST SSVKARNAQN VMSSALPVNA KVSVKPAGVS
 201    EKLAAALEHH HHHH*
```

Arabidopsis thaliana eIF4G: Browning Lab cDNA sequence

```
   1 tcgactcgat cctttattgt tatttatgac agttttaaga agcccggcaa tgctcaagga
  61 ggagggcagc ctcgggtgaa tctgccacct gtgaatcatc ctaataatca caacaatggt
 121 cccaatgctc actctcgctc tcaaggtatg atcgttattc acggagtgcc gcaaagtctt
 181 taaatttgta taatttcaga ttaatttggt cagtgtgtga gtgatcgttt ggtttagttc
 241 aatctatctt acactgttag ttacaggaga accgggtgtt ggtggaccaa ccaatccaac
 301 tgaatcgttc aacagaaaca ccggacctat tccaaaggct ccaacttctc agtctaccgt
 361 catgagttcc aagatcaatg agacgcccaa cacagctaaa ggtaggtttt tgttgcatcg
 421 tctatttttt tgtatgcata ctgagctttg aagttgaact aaattcctct gtaattccat
```

481 aaaggtataa gttttataat gtttattttg ttttgtttag tggcagcctc tggagacgct
541 tctcaggcat ttcctctcca gtttgggtca cttggtcctg atttgatggt aatgctgttg
601 ttccttcctt tgtttgtttg atttttccat caacacctga ataatcttcc ctgttggttt
661 tctaggttcc tgctcgaact acctcagcac ctccgaatat ggatgaccag aaacgtgccc
721 aggtggaaca ccttttagta ttatgcatct gccgttactg tgattgttgc ctacataaat
781 aattctttct tataggatga tgaatagtct gtgtctcgaa ctttgcttct gacttttaaa
841 tgctctatca tttgtcttgt attgcgattt gctggcctga ttgcttgatc tttaactcaa
901 attggtttaa agaatatata ctgaaactta agatgtcttt tattcacttt gttctgaact
961 tgacctttcg tttatgacac gagcaaacat tgttttgcag atgcagcaat cttctttaag
1021 aacggcgtca aatgtgccag cttctgtacc caaaaaagat tcatcaaata agggtgcaga
1081 taatcaattg atgaggaaag aggggcacaa tccatcgagt gaaaaagctg atatccaagt
1141 cccacatata gcccctccaa gtcaaacgca gaagtctcca attacaaata ttcgcatgcc
1201 ttctgtgcag acaccatatc agcatactca ggtccctcac cctgtacatt ttggtgggcc
1261 gaatatgcat atgcagcctc ccgtgactgc aacctcgttt cagatgccaa tgccaatggc
1321 attatctatg ggaaatactc ctcaaatccc gccgcaggtg ttttatcagg gacatccacc
1381 acatccgatg catcatcagg gtatgatgca tcaggctcag ggacatggtt ttgcaactcc
1441 aatgggtgct cagattcatc ctcagttagg ccatgtgggt gtgggtttga gccctcagta
1501 tccccagcag caaggtggaa aatatggtgg ggcacgcaag accacccctg taaagattac
1561 acatcctgac acacacgaag agctgaggct tgatcgacgt ggtgacccgt attcagaagg
1621 cgattcaacg gctttaaaac cacattctaa cccacctccc agatcacagc cagtctcatc
1681 atttgctcca agaccagtca atttggtgca accctcatat aactccaata ccatgatata
1741 tcccccggtt tcggtaccgt taaataatgg tccaatgtca tccgctcagg caccgagata
1801 tcattaccca gttattgatg ggtctcagag agtacaactt atcaaccaac ctgctcatac
1861 tgctccacag cttatcagac ccgctgctcc tgcacatctt tcctctgatt cgacttcctc
1921 tgtgaaagca cgcaatgccc aaaatgtaat gtcatctgct ctacctgtaa atgcgaaggt
1981 atcagtgaag ccagctgggg tttctgaaaa gcttggatca ccaaaagaca ggtcacatgg
2041 agaagttaac atttctctgt cacaaaagaa cgtggaggca tgttcgttga gctcttccca
2101 gcagccgaaa cctagctttg tctctggagt accaaattcg tctgctccgc cagcaaagtc
2161 gcctgtggag actgttccgc tagcaaagtc gtctgtggag actgttccgc cagtaaagtc
2221 gtctgtggag actgctccag ttacaacgac tgaaatcaga agagcggaaa tggtgagtga
2281 gtcgatctca gttgaagatc agacatgtaa ggtggaaccc cctcataatc tgactgaggt
2341 atgatactgt gttttgtttt tgggtattat tcattaattc tttttcttat ttgattaaat
2401 tattttcttt ttttggctgg ttcactcaga atcgtggaca gactatgcca gactctctgg
2461 tctctgatcc tgaaacagca accgttgctg ccaaggaaaa tttatcactc ccagctacca
2521 acgggtttag gaagcaactc ctgaaggtgt ctactacatc tgatgctcca acttctgact
2581 cagtagatac aagtattgac aaatctacgg aaggttcaag ccatgcctca tcggagattt
2641 ctggttcttc accgcaagag aaagacctaa aatgtgataa ccggactgct tctgacaagc
2701 tcgatgaaag gtctgtaatt tctgatgcaa aacacgaaac actgtcaggt gtgcttgaga
2761 aggcacagaa tgaggtagat ggtgccacag atgtctgtcc tgtctctgaa aaactagctg
2821 ttacagatga tacgagctct gaccttccac attctactca tgttctgtct tctactgttc
2881 ctcttggaca ttcggaaaca cataaatctg ctgttgaaac aaacacgaga agaaatactt
2941 ctacaaaagg aaagaagaag ataaaagaaa tccttcaaaa agcagatgct gcagggacaa
3001 cttctgatct ctatatggct tacaaagggc ctgaggaaaa gaaagagagc tcaaatgttg
3061 ttcatgatgt ttcgaaccag aacctgttac ctgccatacc tcaggctgtt gaagccattg
3121 tggatactga accagtgaaa aatgaaccag aagactggga agatgcagcc gatgtttcta
3181 caccaaagct ggaaactgca gataattctg tgaatgcaaa gagaggttcc tcagatgagg
3241 tcagcgacaa ctgcatcaat acagaaaaga agtactcccg ggatttcctc ctaaagtttg
3301 cagacctgtg tactgctctt cctgagggat ttgacgtttc gcctgatatt gctaatgcct

3361 tgattgttgc atatatgggt gcatcacatc atgaacatga ttcatatcct actcctggaa
3421 aggttatgga tcgccaagca agtggtgctc gtttagatcg ccgtcccagc aacgtggctg
3481 gtgatgatag atggacgaag aatcagggtt ctcttccagc aggatatggg ggtaacgtag
3541 gtttccgacc tggtcaagga ggaaactcgg gagttttaag aaaccctcgt atgcagggac
3601 caattatatc tagaccgatg caacctgtgg gtcctatggg aggaatgggt agaaataccc
3661 ccgacttaga aaggtggcaa cgtggttcaa atttccaaca aaaaggactt tttccttctc
3721 cgcacactcc tatgcaagtg atgcacaaag ccgagagaaa ataccaagtg gggacaattg
3781 cagatgaaga acaagcaaaa caaaggcagt taaagagcat cctgaacaag ttgaccccac
3841 aaaactttga gaaactgttt gagcaagtta aaagtgtcaa cattgacaac gctgttacac
3901 tttctggtgt catttcacag atatttgaca aagccttgat ggagccaaca ttctgtgaga
3961 tgtatgcaga tttctgtttt catctctctg gggcgttacc tgattttaat gagaatggtg
4021 aaaagattac cttcaaaaga ttgcttctca ataaatgtca ggaagaattc gagaggggggg
4081 agaaagaaga ggaggaagcc agtagagttg ccgaagaagg tcaagtagaa caaaccgagg
4141 aggaaaggga agagaaaaga cttcaggtgc gaaggagaat gcttggtaac atcagactta
4201 ttggtgagtt atacaagaaa aggatgttga ctgagaaaat catgcacgca tgcatccaga
4261 agttgctcgg gtataatcaa gatccacatg aagagaatat tgaagctctg tgtaaactaa
4321 tgagtacgat aggagttatg atcgatcaca acaaagctaa gttccagatg gatggatatt
4381 ttgagaaaat gaaaatgcta tcatgcaaac aagaattgtc ttctagggtg aggttcatgt
4441 tgatcaatgc catcgatctg agaaagaaca aatggcagga gagaatgaag gtcgaagggc
4501 cgaaaaaaat tgaggaagtg cacagagatg ctgcacaaga acgccaaact caagcgaata
4561 ggctttcacg tggaccctca atgaattcgt caggaagaag agggcatatg gagtttagta
4621 gtcctagggg aggaggagga atgctatcac ctccagctgc ccaaatgggt agttaccatg
4681 gaccacctca aggtcgtggc tttagtaatc aggacattcg atttgatgac aggccatctt
4741 atgagcctag gatggttcca atgccgcaaa ggtcagtatg tgaggagcct attaccttgg
4801 gtccgcaagg tggtcttggt cagggaatgt ctattagaag gcctgcagta gcatcaaaca
4861 cttatcagtc tgatgctact caggccggtg gtggagattc taggcgaccg gccggtggtt
4921 tgaatggttt tggctcacat agacctgcaa gtcctgttac tcacggacgg tcaagtcctc
4981 aagagcgggg aacagcttat gttcataggg aatttgcaag tctgtcgcgt gcttctgatc
5041 tgtcaccaga agtttcgtcc gctaggcaag tactacaagg gccatcagct acagtaaaca
5101 gtcctcgaga aaatgctttg tctgaagaac agttagagaa tctgtcattg tccgcaatta
5161 aggaatatta caggtactat atctctcctt tcttgctggt catttgtttc atttttcgca
5221 aagtcatcag atactgccac atttaggaaa ttattctgtg ggaacctgtg taggtataat
5281 tggatcaaat tacagttcat tcgtttacag ttagagcgtt tgtttacaac gacagatcat
5341 ctgataaagt gtatatttct tgaagtatgt aaattagaat ttgctccaaa caaaacggct
5401 taaaaataca tatgagagac tttgttcttt gactttaata gaatctctaa tacaggggta
5461 aactctgttt tttttactga ctgtgcgttt tatgtatgat gttagtttct attcgtttac
5521 cctctttatt attgataagc actctggtgc gttccaatgt gttcttgtgc agtgcccgag
5581 atgagaatga gattggtatg tgcatgaaag atatgaattc accagcttac cacccaacaa
5641 tgatttctct ctgggtaact gattcgtttg agagaaaaga caaagaaagg gatctcttag
5701 caaagctcct tgtgaacctc gtgaaatctg ctgacaacgc cttaaacgaa gtccagctag
5761 tgaaagggta agtaaacaaa gcccataaat cttgaagccc tagtctaatc aaacgtctgt
5821 tgcttaaatc ttttactgct ttttttttca caggtttgaa tcggttttga aaaccctgga
5881 ggatgcagta aatgatgctc caaaagcagc agagtttctt ggtagaatat ttgggaaaag
5941 tgtgacagag aaagtagtga cattgacaga gattggtcgg ttaatccagg aaggaggaga
6001 agaaccagga agtctgatag agtttggatt aggcggcgat gttcttggga gtgttttgga
6061 gatgataaaa acagaagctg gagaagaaac gttggttgag attcgccgga gctcaggtct
6121 gaggattgaa aatttcaaac ctcatgcacc taaccggtct aagatattag agaaatttac
6181 ttaggaaaaa aaaatggaac catcttttgg gttcctttct tcttctcttt ttttgtttct

6241 ctcttaaaag tcttttctct tttcaagtgc ttcaaacaaa actaatttgt tataaaggga
6301 gtttctctat tttattatat agcaaaaaac ttccaaaatt tctcatttct gtgtttaacc
6361 ttttcgtaca tcagttttaa agcacagaga gctcaatgtt cttccaatat cgttattaat
6421 aaattttgat taaattcaat caaatcggag ttatattacc acatgaggtt aaagggccat
6481 attaaaaagt ctgcacttca tatgagcaac aaggctttta tgtctttatg gttgatttga
6541 tggcccatat aagatagttc aaaggcccat attaaaaaat gccctaacat gattcgattc
6601 gtatataaag acgcttaatg tgacaagcga acttgctcat tagggtttct catctacgac
6661 ggcgtggtgt tcctccttcc tgctctgaaa aatggtgaaa aacctaatcc tttcttttg
6721 ttattgttct tgaatctctg taattccata gcaagttatt cgagtatact tccttctgtc
6781 tcgttctagc tttcgttttg agatttcaag tttcatagtt gatctgtgga ttctgataga
6841 agtctagttt tgtgattttg tcttttgact ctattactct ttttcgatgt ttcacaggcg
6901 aagagaacga agaaggttgg aatcgtcggc aaatacggta atgttcctga ctcatctctt
6961 attgaatcgc ttaggatgtt aatgctatcg aaaataagga taggatctgc aaatttgcgt
7021 agtcttgtag ctcattttaa tggagtatat taacatgtat cgttcatttt tgattacatg
7081 ttgtcaatat gtaatggatc tgtctacact ggtttcatat tgcaattcgt tcttcatttt
7141 ggggtctaag ttggttatta tatgtattgt gagttctctt cctaacgtat gtatggttat
7201 gtgattcagg aacacgttat ggtgcgagta tcaggaagca gattaagaag atggaggtca
7261 gccagcacag caagtacttc tgtgagttct gtggcaaggt atgaaagaga ttaacattca
7321 gttactgttt tcatttctaa atctcttttg gttttggatt attgtggtaa ctgtggtttg
7381 atcaaaatga ttgcagtacg gagtgaagcg aaaggctgtt ggtatctggg gttgcaagga
7441 ttgtggcaag gtcaaggcag gtggtgctta cacaatgaag taagtgattt ctgagttata
7501 tttgtattag ctttgtgttg gttatcaatc aatgatttgt tttctgatga aattgttttc
7561 gtttgtgaat tgcagcaccg ccagtgcggt cactgttaga agcacgatca gaaggttgag
7621 ggagcagatc gagggttaaa agtctgctgg ctttttatat ttggtttcct tgttttgaca
7681 atttaagttt tgcaaactct atgcttgtga ttttgaatca agacttatac attgagcagt
7741 ttaagcaggt ttttattttt tgttgagtta attttcaaag tttcattacc aatagttcta
7801 gctatataac ctaaagttag gcttgaggtt ttgaatgttt gatcctattc aacacctaat
7861 tataaatcct tgaatcctga tctaatgtga tgtttcttat actttatttt gatcagcact
7921 tctttaaccc tataacttgg taacagtatt cgcacgagtc tttacagatc aaaatttta
7981 gagtttgatg tgtgttttca aagtgcttta ctagaaatcc ttacgataat acaatctata
8041 actgagtgta ctaagtttaa ttacacaaaa gtcggtgaat aaacaaatac taacgttact
8101 ggtttttatg tctggcaatg atctcttata aaatctatat attgacttgt ttatatgctt
8161 ttgtgtcgta atcatcctaa gatattggtc ccttgagtgt attttaatgg ccatacatgt
8221 ctcattatga tctttgatga aattgattga tctaagttat acttcagttg aaatttatga
8281 gttttacatg aaattagaaa aacgaataat aatccaacag ttcaaaacaa aatccaacta
8341 actggggaaa atcttttcgg agacaacaac aagacccaag tgtgtcgttc ttttccccca
8401 atgagccagg ctgttgcact tctaagtttt cttggaactt catgagattc tttcaagatt
8461 ctatctttca acttcagctc tcatcctcta acccaatggt tgatgtatgc tccagtcacc
8521 gcagctatcg tttcagttgc tcatccggat gcaattacag agccactaaa tgtatcacca
8581 ttgagaatgc agagcactag tatcatggtt tgtgtacctt gaagccaaag atttttggga
8641 catagctttg agtaggcttc tgcggcttta agttcccgta agtcatcaag aacaagtgag
8701 ggaatgtggt tccaaagtac gctccatcaa tatctgcaca agtcatttaa ggaaaacgct
8761 aagatacaat gacactcaag atcagagtgt ctcaactgaa gcaaaatcag acctaagctt
8821 tcactgaaat attgaaatgg caaagccaaa attatggatt atggaacaga gaaaaaggat
8881 actgccttgg aatttagatc gcggggtaag atatatcctc gcatttaggg cagtatatct
8941 tcacagtact cgatctcggg atatcggatt gtccaactgg aagacaagac tgaccgcaac
9001 agaaaactct cgggcatctc ccgaaatcac agttcttgta cttctcagtc tacagagatc
9061 aataaacaag agacgtcgtt aatcaaataa cgaagctaaa agcaccaaaa cataaacaaa

9121 agagtagaga gattctctaa ctaaccattg cagccattcc tttagtagtc aaaatgtaac
9181 gaacatgaat aagaccatat agcatctcag cagctgattc caccatttca tgctgttcat
9241 cagtaaacat ctcacctaaa catagataca agatttcaga ccaaacttta ctagctcctt
9301 attaacaaaa aaagaaagaa gaagaagaag agagatgact tactgttgga agcatcaaca
9361 tctaaaatga gatcaagtgc ataatcatag taagggactt gaccacttaa accacaaaga
9421 ttgaaatcat cttgaatata atcttcatcg acttcacaga agaaatcatt ccctctcaaa
9481 ttacaaaacc aagagatcca cgacgtatca tcaccctccg atccactcac atcagaccat
9541 tcactatcag tttccgattc cactgcaaca ccaaaaaaaa aaaatcatca gaaagaaagc
9601 aattaaccaa caaaaagact cacagatgtg aaaaagagag gaacaaacca tcaggaagct
9661 gagttttagt agatgtgaag gaaaagggat ctttgtcttt agaagagaaa accctagatg
9721 tggtggtgga agttgaagat ttctctagtt tcttattgag tgcatcgttg attcgtttcc
9781 gatcaatagc tccaccgagg atctctgatc tcgatgaccc accaccacct cctccactac
9841 gttccttgta cataattcga ttccttctcc aaaaagacaa gaaatttgga gattagattc
9901 tcgaaagtgt cagactttct gagaagcttg agagagaaag gaaattagag agacggagaa
9961 atctagggtt ttcttctttt taattttaat ttttttaaa aaattatata taattttggt
10021 aacagagtgg aggacaggac ataggcgatg taaagatggt ctctctttct cctctcaact
10081 gttttgtttt tttttgcttt ccagtttttg ttctttctat tttacaatgg gcttctataa
10141 ttgggccttg tgttagctaa agtcagttta gtctcattgc ttgcttccat cagaaacgac
10201 caaaacagat gggccacagc ccacataaga tgaaatccat aaatcgtgaa caaacatgac
10261 aaccaacaag attacgattt ttggaattag agtacatata catgttggaa cttggaagat
10321 gcattggtca acggagctag acctgacaaa ggtcttcgga taataataat atcgtaaaca
10381 ttaaggaaac tttttaattt taaaaaatcg gacatgaaga gcacgaattg aactctatac
10441 aaaatctaga gttgaatcgt gtttcagaaa ctgaagtaca cgagcttgtg gtggcgtggt
10501 gcacgactct gcataaaact cattatatta tctattcaaa atctaagcat ttgtttcaga
10561 caacacaagt ttcattttta aaagatctcg acttacgaca aaaaaaaaaa aaaaagatct
10621 cgacttttt tgttagttaa atactacaat agtttgagaa atattttgta tttattttat
10681 ttttttagt ttgagaaata ttttgtattt cttcgtttgt gttcaagtaa agtttgtgat
10741 ccacgtttat atacaaaatt gatagataaa aattggactg gcatatcatg taatatacac
10801 acccaaacat aatttttggg atctagaaac atttgcttgc agaacggagc ctttgtcgga
10861 gaaagtagtc atgcctcgtt acaacgactt ttgacaattt cccaaaagtt ttatagaagt
10921 tgtaaattta atactagtag ttgtatattg aaattttatt tttgcattaa acggtaagga
10981 tcgagtcga

Appendix C

DNA restriction digest

1. Incubate plasmid DNA, restriction enzymes and appropriate buffers at 37ºC for 1 hour.

2. Add 1 µl of shrimp alkaline phosphatase (SAP) to the vector restriction digest.

3. Restriction digests are allowed to continue for an additional 30 minutes at 37ºC.

Ethanol precipitation of DNA

1. Cut desired gel band out and place in a 1 mL eppendorf tube.

2. Add 0.1X (X being the volume of the DNA sample) 3M NaOAc and 2X cold absolute ethanol.

3. Contents are mixed and pulsed in the table-top centrifuge and liquid is removed from the pellet.

4. 200 µl of 70% cold ethanol is added and the sample is spun down.

5. Once again, the liquid is taken off. The sample is spun down again and any remaining liquid is removed.

6. Then, the open tube is placed in a 37ºC incubator to dry for 10 minutes.

7. The pellet is resuspended in 5.5 µl of de-ionized water.

Ligation of DNA

1. Combine 3 µl 5X ligase buffer, 1 µl T4 DNA ligase, 5.5 µl DNA of gene insert, 3 µl DNA of vector and 2.5 µl de-ionized water.

2. This 15 µl ligation reaction is incubated at 4ºC overnight.

Ligation of DNA

1. 20 µl of competent cells are thawed on ice and 1 µl of ligation DNA is added.

2. This mixture is incubated on ice for 30 minutes.

3. Heat shocked for 1 minute in a 44ºC water bath.

4. The mixture is then placed back on ice for 2 minutes.

5. 1 mL of LB is added to the mixture and is incubated at 37ºC, shaking, for 1 hour.

6. This mixture is then plated on an appropriate antibiotic LB plate and incubated at 37ºC overnight.

Screening transformed colonies for the correct DNA insert

1. Perform plasmid DNA prep by starting 5 mL LB cultures, containing 100 µg/mL of the appropriate antibiotic, for the colonies.

2. These cultures are incubated at 37ºC shaking overnight.

3. Also, a patch plate is made of all colonies tested on an LB antibiotic plate, and incubated at 37ºC overnight.

4. The next day, the 5 mL cultures are used to perform a Mini-prep according to the manufacturer's instructions (QIAgen) and eluted with 40 µl de-ionized water.

5. Each of the plasmid DNA preps were tested by a restriction digest to determine whether the DNA was successfully inserted into the vector suing appropriate restriction enzymes according to the procedure above for restriction digest.

6. The completed restriction digest is run on a DNA agarose gel.

7. Check for correctly sized DNA fragments.

Basic protein expression procedure

1. Using a single colony, 250 mL LB cultures containing 100 μg/mL appropriate antibiotic are started at 37ºC shaking overnight.

2. The next day, the 250 mL culture is poured into a 1 L LB flask containing the antibiotic, and is left shaking at 37ºC until OD_{600} reaches around 0.4-0.6 absorbance.

3. At this OD_{600}, most of the bacteria are in log phase, and are in a good state to express proteins when induced by IPTG, which acts to activate the lac operon that controls the recombinant gene.

4. The induction phase is left for 3 hours, shaking, at a temperature of 30ºC.

5. The cells are harvested by being centrifuged in the JA-10 rotor at 5.5K rpm at 4ºC for 15 minutes.

6. The bacterial pellet is transformed into a 50 mL test tube and stored at -80ºC.

12.5% SDS-page gel

1. Clamp a 12.5% SDS-page gel in the upper chamber of a gel apparatus and seal around the gel using 1% melted agarose.

2. Place the gel comb in between the two plates of the gel and pour stacking gel into this opening.

3. Let the stacking gel polymerize for about 20 minutes, or until solid.

4. Pour 1X running buffer into the upper and lower chamber.

5. Remove the gel comb.

6. Load protein samples and molecular weight marker into the formed wells.

7. Finish assembling the gel apparatus and electrophorese at 30 mA for 45-60 minutes.

8. Remove the gel and stain in Coomassie Brilliant Blue stain for 10 minutes.

9. Place gel in destain solution for about 1-2 hours.

Western blot

1. Run a 12.5% SDS-page gel according to procedure above.

2. In the membrane blotting apparatus, two pieces of blotting membrane, cut to the size of the gel, wet with Anode 1 buffer (see **Appendix A**), are arranged.

3. Then, two more pieces, wet with Anode 2 buffer (see **Appendix A**), is laid on top.

4. Then, the PVDF membrane is first wet with methanol and then Anode 2 buffer, is placed on top of the blotting membrane.

5. The gel is also wet with Anode 2 buffer and stacked on.

6. Two more pieces of membrane are wet with Cathode buffer (see **Appendix A**) and laid on the very top.

7. The blotting apparatus is run at 100 mAmps for 1.5 hours.

8. The PVDF membrane is placed protein-side facing up in a clean container with 5% milk/HNAT solution either shaking at room temperature for 2 hours, or at 4ºC overnight.

9. The solution is poured off the membrane and 5% milk/HNAT containing 1:10,000 concentration of the primary and secondary antibody, in our case, anti-His C-terminus-HRP antibody (Invitrogen), is added and is left shaking at room temperature for at least 1 hour.

10. This solution is poured off and 3-5 rinses of 5% milk/HNAT for 5 minutes each are done, with the last few washes performed using HNAT only.

11. The PVDF membrane is developed using equal parts of the two reagents provided in the chemiluminescence kit (Pierce).

12. This solution is left on the membrane to react for 10 minutes.

13. Using film, luminescence can be picked up on the membrane, indicating a protein that bound the anti-His tag antibody successfully.

Preparation and Care of the Ni-NTA column

1. After every use, the Ni-NTA column is stripped with 3X the column volume of strip buffer (see **Appendix A**)

2. The column is stored in binding buffer with 30% ethanol.

3. The day before the next use, wash with 3X the column volume with charging buffer (see **Appendix A**).