Cloning and Expressing Recombinant eIF4G, a Subunit of Arabadopsis eIF4F
Lauren E. Harkin,1 Joceline S. Liu,1 Elizabeth Burks,2 and Karen S. Browning

1) The University of Texas at Austin, Department of Chemistry and Biochemistry
2) The University of Texas at Austin, Department of Pharmacy

Abstract

Eukaryotic translation initiation involves a number of factors that must be present for the small ribosomal subunit to load onto a capped and polyadenylated mRNA. Among these complexes is eIF4F, which includes subunits eIF4G, eIF4E and eIF4A. The large subunit, eIF4G, was cloned as both a full-length gene and a truncated version into a vector for expression in E. coli cells. The full-length version does not express well. For the purpose of raising antibodies to eIF4G, the truncated eIF4G was co-purified with eIF4E. Preliminary results from the rabbit polyclonal antibodies suggest that they are specific for eIF4F.

Introduction

eIF4G is part of a complex of eukaryotic translation initiation factors called eIF4F. eIF4F includes eIF4G, eIF4E, and eIF4A. eIF4G is the largest of these factors and is a docking site for other subunits of eIF4F as well as sites for eIF3, which binds the 40S ribosomal subunit, and PABP, which binds the poly-A tail of the mRNA. eIF4E binds the 5’ cap of the mRNA, and is tightly bound to eIF4G. eIF4A does not co-purify with the other factors but has been shown to associate with them.

Arabidopsis thaliana is a flowering plant in the mustard family whose genome has been sequenced, and is an excellent model system for plants because it can be used for genetic studies.

Wheat, yeast and human have two forms of eIF4G, although the relationship between these two forms in each organism may be different (Gallie and Browning, 2001; Tarun et al., 1997). In Arabadopsis, the two forms of eIF4F are called eIF4F and eIF(iso)4F. These consist of eIF4E and eIF4G or eIF(iso)4E and eIF(iso)4G. Although the nature of the work performed by each is similar, and they are moderately homologous, the subunits may be responsible for translating different mRNAs.

The central location of the eIF4F complex in translation initiation makes it a prime source of translational control. It has even been suggested that the mRNP complex consisting of eIF4G, eIF4E and the mRNA is not temporary and may be able to undergo several rounds of ribosomal loading as a stable complex (Gross et al., 2003). This has implications in that the affinity of eIF4F for a particular mRNA and its stability once bound may affect the rate at which
certain messages are translated, and thus the amount of a particular gene product present in the cell at any one time. Since (probably) all genes in the cell require either eIF4F or eIF(iso)4F to be translated, this control point could affect hundreds or thousands of protein expression levels.

Agriculturally devastating plant viruses have been known to hijack these initiation factors to force the cell to produce viral proteins. For example, the 5’ leader sequence of tobacco mosaic virus was shown to require both eIF4G and eIF3 for message translation (Gallie, 2002). Some plants that have viral resistance have been shown to have knockouts in various subunits of eIF4F or eIF(iso)4F. The loss of eIF(iso)4E has been shown to impart viral resistance to potyviruses (LeLliss, 2002), emphasizing the critical cellular control the eIF4F and eIF(iso)4F subunits possess.

The entire eIF4G gene and a truncated version that was created to optimize expression have been cloned into vectors for protein expression in bacteria. The end goal is to obtain an antibody to eIF4G that can be used in further studies, especially of knockout plants with mutations that effectively remove one or more eIF4F or eIF(iso)4F subunits.

Results and Discussion

Cloning of full-length eIF4G

The Arabadopsis eIF4G gene product is 4203 base pairs, a large size to be cloned into a standard bacterial vector. For this reason, the eIF4G was cloned in two pieces, using engineered restriction sites for Nde I at the 5’ end and Bam HI at the 3’ end, as well as a natural Xba I site in the middle of the gene (bp 2111).

In order for the Xba I site to be used, the Xba I site in the multi-clonal region of the pET 22b vector (Novagen) had to be eliminated. Also, the internal Nde I and Eco RI sites were removed from the eIF4G gene by overlap PCR. The Nde I site (bp 3111) had to be removed so it could be engineered into the 5’ end to clone the gene. The Eco RI site (bp 3081) was removed in case later cloning steps would be facilitated by its removal, since it was in close proximity to the Nde I site.

The C-terminal portion of the eIF4G gene was first cloned into pET 22b using restriction enzymes Bam HI and Xba I. Then the N-terminal portion was added using Xba I and Nde I. The gene-vector construct was confirmed using forward and reverse dideoxy sequencing (Fig. 1).

Cloning of truncated eIF4G

Initial expression studies of the full-length eIF4G produced either no protein or very low yields. To combat the problem, a truncated construct was made. The gene product of the construct would be missing the first 313 amino acids. This part of the protein contained a PEST sequence that tags proteins for degradation (Rogers, 1986). It also contains the region that binds poly-A binding protein (PABP). Since the antibodies to the eIF4G protein do not need to recognize the whole protein, eliminating a portion of it should not affect detection of full-length eIF4G.

The gene was cloned using PCR from the full-length construct, the previously used 3’ end primer with a Bam HI site attached, and a new forward primer with a Nde I site before the internal ATG at bp 952. It was cloned
into Novagen pET22b without the Xba I site as well (Fig. 2).

**Expression of truncated eIF4G**

The truncated version of eIF4G was chosen for expression of protein to be used as antigen for polyclonal antibody production. This assessment was made after comparing expression of full-length and truncated eIF4G constructs (Fig 3).

The truncated eIF4G was co-purified with a previously made construct of eIF4E in pET 28a (Novagen). The eIF4E should bind to the eIF4G after they are both released from cells during sonication. This complexation step may reduce proteolysis since a tighter structure is formed. Also, this aided in protein purification since the final purification column could be an affinity matrix with m7-GTP attached for cap-binding eIF4F purification.

Another interesting aspect of the expression was that instead of using LB broth, 2 x LB broth was accidentally made and turned out approximately double protein expression (Fig 4). 2 x LB broth is broth that contains twice the concentration of LB media powder.

The procedure noted in methods was performed twice. The first preparation was done using LB media and yielded 0.5 mg of purified protein from 24 L of cells. The second preparation used 2 x LB media and yielded 1.3 mg of purified protein from 24 L of cells.

Three bands at 150 kD, 130 kD, and 100 kD on SDS-PAGE gels were sequenced by N-terminal Edmund degradation and were found to have the same amino acids expected from a truncated eIF4G gene product (Fig. 5). This indicates that all three bands are protein of truncated eIF4G origin and that the smaller bands are being degraded at the C-terminus.

**Antibody Production against eIF4F**

Both eIF4F protein preparations were sent to Antibodies, Inc. for rabbit polyclonal antibody production. The pre-immune bleed did not cross-react at all with 10 ng of purified recombinant eIF4F at dilutions more dilute than 1 in 20. The first bleed gave antibodies that had a maximal absorbance starting at a 1 in 800 dilution to 10 ng of purified recombinant eIF4F. These levels were found using ELISA.

Preliminary Western blot analysis shows that the first bleed of the antibody definitely reacts with Arabadopsis eIF4E, cross-reacts with wheat eIF4G, and does not cross-react well with wheat eIF4E (Pavel and Browning, data not shown). However, more Western blots must be done with both the recombinant eIF4F protein and full-length eIF4G-expressing E. coli cells to be sure that the eIF4G is there and is detected by the antibody.

**Conclusions**

*Arabidopsis thaliana* eIF4G has been cloned and expressed in E. coli cells. Protein was purified and injected into a rabbit to produce antibodies that preliminary evidence suggests will be able to detect *Arabidopsis* eIF4F and wheat eIF4G.

Now that the antibody has been made, it can be used along with a library of other antibodies to recombinant translation initiation factors to elucidate the differing roles of eIF4F and eIF(iso)4F by analyzing plants which have various subunits knocked out.
A dicistronic construct will be made with eIF4E in the same vector as full-length eIF4G. This artificial operon should allow for better expression since the two subunits can bind in vivo and the tighter structure may afford better protection from roaming proteases in the bacterial cell. A new purification step could use a Millipore filter centrifugation device that only retains proteins greater than 100 kD, thus purifying just eIF4G. However, this step would need to be done under somewhat denaturing conditions to get the two subunits apart, which could affect protein activity. The full-length eIF4F construct could also be used for in vitro activity assays, to compare the roles of eIF4F and eIF(iso)4F. It could also be compared with the truncated version to investigate the role of PABP-eIF4G interaction.

Materials and Methods

Cloning XpET 22b
XpET 22b was made by digesting pET 22b (Novagen) with Xba I under appropriate buffer conditions, gel purifying and filling in by Klenow. Blunt ends were then ligated to add four extra base pairs between the ribosome binding site and the lac operator, in the middle of the Xba I site, thus eliminating it. Ligations were transformed by electroporation into E. coli DH5α (Novagen). DNA was extracted using the Quiagen miniprep kit.

Cloning full-length eIF4G into XpET 22b
The gene cDNA was obtained by two reverse transcription reactions with appropriate primers for either the N-terminal or C-terminal half of the gene, which engineered a Nde I site at the front of the N-terminal half and a Bam HI site at the C-terminal half. The C-terminal half was engineered by overlap PCR to remove Nde I and Eco RI sites. The codons were changed but code for the same amino acid as before to preserve the protein sequence. The edited C-terminal half was cloned into XpET 22b first by digesting both PCR product and vector with Xba I and Bam HI, gel purifying, ligating, transforming into E. coli DH5α (Novagen) by electroporation, and screening for colonies by digestion with Xba I. The C-terminal end in XpET 22b (also called 3’ovrlp) was sequenced by dideoxy sequencing (MBB Core Facility). The N-terminal end and the engineered C-terminal end in XpET 22b were both digested with Xba I and Nde I. They were gel purified, ligated, and transformed into E. coli DH5α (Novagen). Screening was performed using Nde I and Xba I restriction digests, and the construct was confirmed by sequencing.

Cloning truncated eIF4G into XpET 22b
Primers were used to engineer an Nde I site at the beginning of the truncated version (bp 952) and a Bam HI site at the end of the gene. The gene was cloned into XpET 22b using a digest of both pieces with Nde I and Bam HI, gel purification, ligation, and electroporation into E. coli DH5α. Colonies were screened using Nde I, Xba I, and Bam HI restriction endonucleases. The construct was confirmed using dideoxy sequencing.

Expression of truncated eIF4G and eIF4E
The truncated eIF4G construct was transformed into calcium competent BL21(DE3) cells (Novagen) and the eIF4E in pET 28a construct was transformed into electrocompetent BLR (DE3) cells. Freshly plated cells were inoculated into 10 ml LB cultures containing either 100 mg/L ampicillin for the eIF4G or 30 mg/L kanamycin and 12.5 mg/L tetracycline for the eIF4E and grown at 30°C shaking for 6-10 hours. These cultures were transferred to 100 ml 2 x LB (40 g Invitrogen LB Broth Base per liter of media) cultures with appropriate antibiotics. After an overnight incubation (16-20 hours) at 30°C shaking, the cells were transferred to 1 L of 2 x LB. The cells grew to an OD of 0.8 to 1.0 at
30°C shaking, and were induced with either 0.4 mM IPTG for the eIF4G or 0.2 mM IPTG for the eIF4E. After three hours of induction at 30°C shaking, the cells (6 Liters) were harvested in a JS-4.2 rotor (Beckman) in a J6-MI centrifuge at 4.2 RPM for 20 minutes. Pellets were removed into two 250 ml bottles (Beckman), each containing cells from 1 L of eIF4E expressing cells and 2 L of truncated eIF4G expressing cells for centrifugation at 5,000 g. Pellets were weighed and flash frozen using liquid nitrogen, and stored at -80°C. The harvested cells were resuspended in 80 ml sonication buffer [10% glycerol (Fisher), 20 mM Hepes pH 7.6 (USB), 0.1 mM EDTA, 2mM DTT (Roche), 0.6 M KCl (Fisher), 1.5 Complete Protease Inhibitor Tablets (Roche), 0.1 mg/ml phenylmethylsulfonyl fluoride (Sigma), and 0.1 mg/ml soybean trypsin inhibitor (Sigma)]. The resuspended pellet was sonicated in a salt-water-ice bath using a Vibra Cell sonicator (Sonics & Materials, Inc.) 5 mm probe for two times at 70% for 30 sec., two times at 40% for 30 sec., and five times at 90% for 30 sec. The cells were centrifuged at 50 kRPM for 1 hr in a Ti50.2 rotor (Beckman Coulter) in an Optima LE-800K Ultracentrifuge (Beckman Coulter) at 4°C. The supernatant was diluted to 0.1 M KCl with dilution buffer [10% glycerol, 20 mM Hepes pH 7.6, 0.1 mM EDTA, 2 mM DTT] and loaded onto two 5 ml phosphocellulose columns. Proteins were eluted with dilution buffer containing additionally 0.3 M KCl. 1 ml fractions were collected and their composition was checked by SDS-PAGE. Concentrations were determined by the Bradford method. The previously described process was performed four times. Then fractions containing eIF4F were combined and run over another 5 ml phosphocellulose column and eluted as before, collecting 1 ml fractions. SDS-PAGE and the Bradford method were used to determine fractions to load onto an affinity matrix containing m7 GTP linked to Sephadex 4B (Amersham) and eluted with 200 μM m7 GTP in dilution buffer plus 0.1 M KCl. SDS-PAGE and the Bradford method were used to determine which fractions to concentrate. Fractions containing eIF4F were concentrated using a Millipore filtration-centrifugation concentrator. SDS-PAGE and the Bradford method were used to determine the concentration of the sample.

**Antibody Analysis**

Antibodies were raised to two of the complete preparations described above which yielded together 1.8 mg of protein. Antibodies, Inc. (Davis, CA) injected the antigen into a rabbit according to their standard protocol. ELISA was performed using 10 ng of purified recombinant protein that was used as antigen per well. The first bleed was sequentially diluted 1:2 from 1 to 25 until the dilution reached 1 in 6400, and the assay was performed in duplicate. The preimmune bleed was also tested at concentrations of 1 to 10, 1 to 20, 1 to 40 and 1 to 80 in duplicate. The protein incubation step was 1 hr., the primary antibody incubation was 1 hr. 5 min., and the secondary incubation step was 1 hr. 5 min.

*This work was made possible by the Beckman Foundation, the National Science Foundation, and the University Co-operative Society.*

**References**


conformational coupling between eIF4G and eIF4E. Cell. 115, 739-750.


Figure 1. Gene map of full-length eIF4G in XpET 22b.
Figure 2. Gene map of truncated eIF4G in XpET 22b.

Figure 3. SDS-PAGE stained with Coomassie blue. Truncated eIF4G runs at approximately 150 kD, while eIF4G runs at approximately 200 kD. Lane 1 and 8 contain Precision Plus Protein™ Standard. Lanes 2-7 show fractions of truncated eIF4G eluted from a phosphocellulose column. Lanes 9-11 show fractions of full-length eIF4G eluted from a phosphocellulose column. Truncated eIF4G expresses better than full-length eIF4G.

Figure 4. SDS-PAGE stained with Coomassie blue of the truncated eIF4G band. The band on the left is from a phosphocellulose column loaded with protein from cells grown in LB media (20g Invitrogen LB Broth Base in one liter of media). The band on the right is from a phosphocellulose column loaded with protein from cells of which half were grown in LB media and half were grown in 2 x LB media (40g Invitrogen LB Broth Base in one liter of media). Both bands are the peak fractions of different phosphocellulose columns each loaded with cells expressing truncated eIF4G and eIF4E. The gels were run concurrently, so staining procedures were a similar as possible.
Truncated eIF4G (150 kD)
Degraded Truncated eIF4G (130 kD)
Degraded Truncated eIF4G (100 kD)

eIF4E (26 kD)

Figure 5. SDS-PAGE stained with Coomassie blue of purified and concentrated truncated eIF4G/eIF4E. The degradation products and the non-degraded truncated eIF4G were all sequenced by N-terminal Edmund degradation and confirmed to be eIF4G. 13.5 µg of eIF4F protein was loaded into lane 1; lane 2 is blank; Precision Plus Protein™ Standard is in lane 3.

Supplements

Supplement 1:

*Arabidopsis thaliana* eIF4G DNA sequence:

ATGATGCATCAGGCTCAGGGACATGGGTGCTCAGATTATCATCCTCAGTTAGGCCATGTGGTGTTGAGCCCCGATTTGGGAAATGTTG
GGGCCCCACAAGACCACCCCTGCTAAAAGATTACACATCCTGACACACGAAGAGCTGAGGCT
TGATCGATGTGGTTGCCCGTATTATCGAGAGGCGATTAAACACATTCTACCC
ACCTCCAGATCGAGACAGCTCAGCATCTACCATTTGCTCAAAGAGGCAGTCTTTGTGCAACTCCC
ATATAAATCCCTAATACCATATAATCTCCCGGGTTGTAACCCATATTGGTTGCAACTGTC
ATCGGCTCAGGACCGAATACATTTAACAGTTATTTGATGGGTCAGAGATACATTAT
CAACCAACCTGTGCTCAGACATGTTTTCTCTCTGCTCCAGACGCGCTCCTGACATCTTTCTCTC
GATTCGACTTCCTGCTTTAAGACAGCTAGAAAAATGTAATATGATCTGCTTCCGAGCT
AATGCGAAGTCTAGATGGACGCGAATGGTTCTGAAACAGCTTTGAGAAGAGCTC
TTCCAGCAAGCCAAACCTAGCTTTGCTCTGGAGTACCAAAATTCGCTTGACGAGCT

Supplements

Supplement 1:

*Arabidopsis thaliana* eIF4G DNA sequence:
**Arabidopsis thaliana**

eIF4G protein sequence:

```
MMHQAQGHGFATPMGAQIHPQLGHVGVGLSPQYPQQQGGKYGGARKTTPVKITHPDTHEELRL
DRRGGDPYSEGDSTALKPHSNPPSPVSSAFAPPVNLVQPSYNSNTMTIYPPVSPVLPNSPMSSAQ
APRYHPVIIQSGQVRQVLINPAHTAPQLRPAAPAHLSSDSTS5VKARRAQNVMSALPYNALKSV
KAPAGYKLSKGPSKRSDSHEVNGNOSLOOKVNEALSSQSSQKPKSFVSGVPNNSAPPASKPSVTEPL
ASSVETPPVKSSTVAPVTAPVTTTEIIARRAEUVESIESVEDTCTKVEPHNLTERNGQTMPSDLSVDPE
TAVAAKENLSLPTANGFRKLIVSTTSDDSDTSDISDSTKGESSHASSIEGSSIPQDEKLDCEDN
RTASDCLKDERSVIADKHALGLGQAEQNEVGDATDVCPYSEKLAVTDDTDSDLPHSTHVLSTV
PLGHSETHKSAVETNTRRTSTGKKEIKIKEQKAADAAGTTSDDLYMAYKPEEEKKESNVVHVDSS
NQNLLPAIPQAAEVPAEVDPVKNEDPVEDAADVSTPKLECHTNSVAKGRSSSDESDNCTINEK
KYSRDFFMKFAIDCALTPEFDPVDIAEALIVAYMGAASHHHEHDSYPTPGKVMDRQASAGARLDRR
PSNVAGDDRWTKNQQLSAPAGYNVGFRPGQGGNSGVLRNPRMQPGPHSPQRYVPGPMGGMGR
NTDPLERWQRGNSQKGLFPSHTPMQVMHKAERKYQVGTIADEEQAQKRLQSLKNLTPQNF
FEKFLEQVKSVDNIAVTLSGVISIQDFKDALMEPTFCEMYADFCFHLSGALPDENEGIKFTRLL
LNKQCEEFEGERKEESEEASRVAEEGQVETEEEREKRLQVRRLMGNLRIELYKCRMTEKI
MHACIQKLLGYQNDPEHELIAVCLKGTGGIVDMDHNKAKFQMDGYFEKMKMLCKQELSSSRYR
FMLNAINDLRKNQWERMKVEQPGPKIIEEVRDAAQERAQTOANRLSRGPSMNSSRGRHMFESSP
RGGGGMLSSPPAAMQNSGYHPGRRFQNSDQRFDPRSPYEMPVMQPSVCEPITLGPGGLGG
QMGRSMRPAVANQYDSDATQPGGDSRPPGAPGLNGHSRHSPAPVTHGRSPOGRTAYVHREF
ASLSRASDLSPSVSSARQVLQPSATVNSPREALSEQLENSLSAIKEYYSARDENEMGCMKD
MNSPAPYEPMISLWMLWHSRRRSPSRPPRPRRPLPPRQVRRRLRNGLMILYKLLMLTKEI
```

Supplement 2:

Truncated *Arabidopsis thaliana* eIF4G DNA sequence

```
ATGCCAGACTCTCTGGTCTCTGTATCTCTGAAACAGCAACGCTTGAAGGTGTGTACTACACTCTGTAGCTCCAA
ACCTCTGACTACAGCCGAGTCCAGTGGTGATAGAGATGGACAGAATCAGGGAGTTCTCTTTCCAGCAGGAAAGGTGACC
```

---

**Supplement 2:**

Truncated *Arabidopsis thaliana* eIF4G DNA sequence
Arabidopsis thaliana eIF4G protein sequence:

MADSLVSDPETATVAKENLSLPAITNGFRKQLLLKVSSTDAPTSVDSTIDKSTEGSSSHASSEISG
SSPEKDLDKCDNTASDILKDRSVDISADAIKHTLSVLEKAQNEVDGTACVPSVEKLAVTDDTSS
DLPHTHVLSTVPLHSETHSKAVETNTRRRKKKKEEILOKADAAGTTSDLYMAKYP
KEKESNNVHVDSNSQNLLPAIPQAVEAIVDTEPSKNEPEDSEADVSTPKLETADNSVNAKRSGS
DEVSDNSCNTEKKSRSFDLLLKFADLCATLPEGLFDVDPSDIANALIVYMGASHHEHDSYTPPGKVM
DQAGARLDRPSNAVDGSLPLAPAGYGGNFGRPGQSGVLRNPRMQGIPVSRMPQ
QPVPGPPGMGMRNPDRPLEWQRSGNFQKGLFSPSHPMPQVVMHAKERKYVQTVIADEEYQAKQRQ
LKSILNKLLTPQNFKEIQKVEQVSNIDNATLSGVISQIFDKALMEPTCMVEADCFHLSGALPDFN
ENGKJKTRLLNKCQFEQERKEEEESRVEAEQVQFTFRNRRLLLMLRLGNLIGEL
LYKKMLTEKIMHACIQLLGSQYNDPHENIELACKLIMSTIGMIDNKHAKFQMDGYFEMKMKML
SCKQELSSRVRMLAINIDLRKNDKQERMKVEEGPKKIEEVRDAAQETQANLRSRPMSNSSG
RRGHMEFSSPRGGGMGLSSPPAQGMSYHGPQPGRQFNQIDFDRSPYERMVPMQPSVCEEP
ITLPQPPQGLGQGMSIRPHAAPASNTYQSDATQAGGDSRRPAGLGLNFGSRHPAPVTHGRSSPQE
RTGALEYHRFASSLRSRLASDLPEVSSARQVLQGSPATVNSPRENALSEQEQLNLSSAIKEYSARDE
NEIIGCMKDMNSPAVLWTVDDPGKDDRLAKLLVNLVSKADNLNEVQLVKGFRS
VLKTLEDAVNDAPKAAFLGRIFGKSVEKVLTEIGRLIQEGGEEPGSLIEFGLGGDVLSGVLME
IKTEAGEETLVEIRRSSGLRIENVKHPAPRNSKILEKFT

Truncated