COVER SHEET

TITLE: Testing a Putative Species-Independent Translation Enhancer: “3 Hairpin” Motif

AUTHOR’S NAME: Ting Zhang

MAJOR: Biochemistry

DEPARTMENT: Biochemistry

MENTOR: Brian G Fox

DEPARTMENT: Biochemistry

MENTOR(2): Michael A Goren

~~DEPARTMENT(2): Biochemistry~~

YEAR: 2010

ABSTRACT

Testing a Putative Species-Independent Translation Enhancer: "3 Hairpin" Motif

Ting Zhang, Michael A. Goren and Brian G. Fox
Department of Biochemistry and Center for Eukaryotic Structural Genomics
University of Wisconsin-Madison

Funded by NIH U54 1U54 GM074901. J.L. Markley, PI. B.G. Fox and G.N. Phillips, Jr., Co-PIs.

Abstract

Cell-free (CF) protein synthesis enables rapid production and engineering of recombinant proteins. Existing CF systems differ substantially from each other with respect to efficiency and scalability. We examined the possibility and mechanism of recently reported "Three Hairpin" (3HP) motif included in mRNA to enhance translation using a high-efficiency wheat germ (WG) CF extract. In this work, enhanced GFP expression plasmids constructed with and without 3HP were transcribed and translated. The differences in levels of transcription and translation from the 3HP and control constructs were detected by DNA and GFP fluorescence, respectively. Although 3HP reportedly gave universal ~10-fold enhancement of cell-free translation, we observed no comparable increases in either mRNA or protein yield associated with 3HP. We tentatively assign this result to the higher efficiency of the WG CF translation system used in our laboratory compared to the coupled-CF protein production system described previously.

TING ZHANG/BIOCHEMISTRY
Author Name/Major

Author Signature

BIOCHEMISTRY
Mentor Name/Department

B G Fox
Mentor Signature

5/7/10
Date

Testing a Putative Species-Independent Translation Enhancer: "3 Hairpin" Motif

Ting Zhang, Michael A. Goren and Brian G. Fox
Department of Biochemistry and Center for Eukaryotic Structural Genomics
University of Wisconsin-Madison

Funded by NIH U54 1U54 GM074901. J.L. Markley, PI. B.G. Fox and G.N. Phillips, Jr., Co-PIs.

Abstract

Cell-free protein synthesis enables rapid production and engineering of recombinant proteins. Existing cell-free systems differ substantially from each other with respect to efficiency, scalability and the ability to produce functional eukaryotic proteins. We examined the possibility and mechanism of recently reported "Three Hairpin" (3HP) motif included in mRNA to enhance translation using a high-efficiency wheat germ cell-free extract. 3HP is believed to melt the region downstream of AUG during advance of initiation complex and consequently increase the residence time of the pre-initiation complex in the vicinity of start codon. In this work, enhanced GFP expression plasmids constructed with and without 3HP were transcribed and translated. The differences in levels of transcription and translation from the 3HP and control constructs were detected by DNA and GFP fluorescence, respectively. Although 3HP reportedly gave universal ~10-fold enhancement of cell-free translation, we observed no comparable increases in either mRNA or protein yield associated with 3HP. We tentatively assign this result to the higher efficiency of the wheat germ cell-free translation system used in our laboratory compared to the coupled-cell-free protein production system described previously.

Introduction

Although integral membrane proteins (IMP) account for almost 25% of open reading frames in fully sequenced gnomes, progress on understanding their structure and function has lagged behind their soluble counterparts *(1)*. In part, this is due to the difficulty in obtaining sufficient quantities of homogenous protein for *in vitro* studies. Current commonly used platforms for IMP expression are C41/C43 *E. coli* single cell production system, insect cells based platform, prokaryotic *E. coli* cell based cell free production system and wheat germ extract based production system. The utility of *E. coli* for IMP production is limited by available membrane space, the toxic effects of competition for the insertion machinery, and incorrect lipid profiles for proper folding *(2)*. Furthermore, even if the protein is produced and properly incorporated, this can dramatically alter the host membrane, and can induce host stress responses that activate proteolytic systems to degrade the IMP *(3)*. An insect cell based platform is one alternative to *E. coli*, but it is expensive, and has the potential problem of unintentional viral genome integration *(4)*. Hence, wheat germ extract based cell-free (CF) translation systems have emerged as alternatives to cell-based methods for IMP production. This approach is attractive because it decouples the synthesis of IMP from the constraints of living hosts. Furthermore, it has been suggested that the eukaryotic translation and folding machinery has been optimized through evolution to facilitate co-translational domain folding, which is crucial in *in vitro* translation of functional eukaryotic proteins that require post- translational modifications. Using this technology, functional G-protein coupled receptors, ion channels, and transporters have been produced in high-yield (mg/ml of reaction) *(5)*.

Although CF protein synthesis enables rapid production and engineering of recombinant proteins, existing cell-free systems differ substantially from each other with respect to efficiency, scalability and the ability to produce functional eukaryotic proteins. A major bottleneck for IMP structural analysis is the limited availability of sufficient amounts of protein samples since wheat germ CF protein translation system typically yields 0.5 to 1 mg/ml. As more protein is needed for NMR and x-ray crystallography analysis, the process can be a significant financial investment as the traditional route to increase yield is to use more extract. Various alternative approaches have been made by many labs to improve the efficiency of this viable membrane protein translation platform. For instance, Endo's lab at Ehime University, Japan, significantly increased wheat germ extract based CF translation efficiency by developing dialysis mode compared to the traditional batch mode *(6)*; Luo's lab at Cornell generated a cell-free proteinproducing gel to improve the efficiency of the cell-free transcription by decreasing the diffusion rate, so that the reaction can proceed much faster due to the close physical contacts of all the components involved in the transcription machinery *(7);* Swartz's group at Stanford engineered the energy source in translation buffer system to improve translation efficiency *(8)*.

In this work, I examined the possibility and mechanism of recently reported "Three Hairpin" (3HP) motif, as shown in figure 1, to enhance translation using a wheat germ cell-free system. The 3HP motif has been proposed to be capable of melting the region downstream of AUG upon advancing of initiation complex and increase the residence time of pre-initiation complex in the vicinity of start codon. This is expected to facilitate AUG recognition by small ribosome subunit and subsequent recruitment of the large ribosome subunit. This model is supported by the observation that the hairpin motif substantially enhanced TL efficiency of the poly (A) leaders *(9)*. Kim et. al showed that the primary and secondary structure of mRNA were important to the expression level of proteins. In other words, if it was a function of the ribosome itself, and not a chaperone or accessory protein, the location of translation was irrelevant. Given that the literature supported a specific type of mRNA 5' enhancing translation, we thought that this 3HP motif might be playing the role *(10)*. In this work, in order to test whether 3HP structure improved protein production through transcription or translation reaction, enhanced GFP (eGFP) expression plasmids constructed with and without 3HP were transcribed and translated. Specifically, two approaches were taken: one was DNA plasmid test, the other one was digested linearized DNA test. These two approaches were also taken in the original 3HP publication by Mureev's group. Linear DNA was tested due to the pEU-HSCB-eGFP polymerase terminator only had 50% efficiency (unpublished result). It was assumed that the mRNA post terminator was not translated. The differences in levels of transcription and translation from the 3HP and control constructs were detected by DNA Nanodrop and GFP fluorescence reader, respectively. Although 3HP was reported to be the active portion stimulating cell-free protein expression, I observed no comparable increases in either mRNA or protein yield associated with 3HP. The higher efficiency of the decoupled wheat germ cell-free transcription/translation system used in our laboratory might have masked the effect of the putative translational enhancer 3HP structure used in the coupled-batch mode CF protein production system reported by Mureev et al.

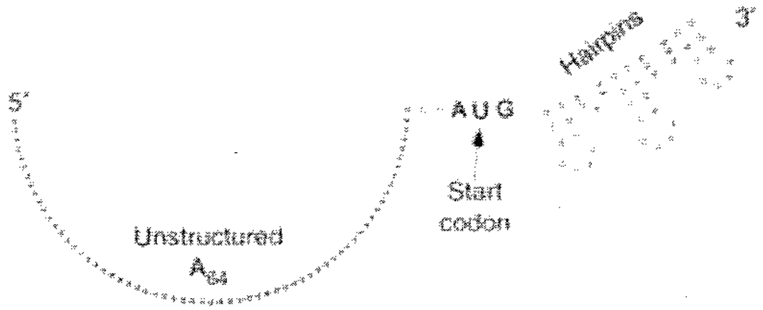


Figure 1. Mureev S, et al (Nature Biotechnology, 27, 2009, 747-752) predicted secondary structure of $(A)_{64}$ 3HP-EGFP mRNA calculated with Mfold. Only the first 112 nucleotides are displayed.

Materials and Methods

Materials

Routine reagents were of the highest grade available from standard vendors. Distilled and deionized water was used for all reagent preparations. Reagents for FlexiVector cloning were from Promega (Madison, WI). The eGFP gene was designed and synthesized by the Center for Eukaryotic Structural Genomics *(11)*. Care must be taken to exclude ribonucleases from all reagents and equipment used for preparation of the cell-free translation reactions, including micropipeters.

Cloning vector

The vector pEU optimized for wheat germ cell-free translation contains 5′-SgfI and 3′-PmeI restriction sites, N-terminal His6 purification tag, and a toxic selection cassette in the multi-cloning site (figure 2), and it is available from the NIH Protein Structure Initiative Material Repository (http://www.hip.harvard.edu/PSIMR/index.htm). Genes encoding for eGFP was amplified by PCR using the primers (5' TCC AAA GGT GAA GAA CTG, 3' GTG TGT TTA AAC ATG TTT GTA CAG TTC GTC CAT AC). A 5'-Tev site was also cloned in using a two step PCR to create pEU-HSCB: ATG-SgfI-TEV-eGFP-stop-PmeI (figure 2). The two primers used were (1st: 5'AAC CTG TAC TTC CAG TCC+ORF, 3'GCT CGA ATT CGT TTA AAC TA +ORF; 2nd: 5'GGT TGC GAT CGC CGA AAA CCT GTA CTT CCA GTC C, 3'GTG TGA GCT CGA ATT CGT TTA AAC). To add 3HP to pEU-HSCB: ATG-3HP-SgfI-TEV-eGFP-stop-PmeI, a two step around the circle PCR strategy was used to create overlapping ends (1st 5' Primer CAC GTA AGT GAA ACC GGA CAT CAT CAT CAT CAT CAC GC; 1st 3' Primer CAG ACT TTA TAC ATT ACT GTC ATG GCT GTA GTT GTA G; 2nd 5' Primer GTC TGT AAA GAC ATA AAA CAC GTA AGT GAA ACC GGA C; 2nd 5' Primer GTG TTT TAT GTC TTT ACA GAC TTT ATA CAT TAC TGT CAT G), followed by Dpn digestion to digest the methylated parental strands. The newly synthesized strands with overlapping ends were directly transformed into Top10 chemically competent *E. coli* for ligation independent cloning. The transformation was platted on Ampicillin and chloramphenical for selectivity screening and incubated at 37 °C for 18 hours. Eight colonies from the transformation were grown up in LB and LB with 15%

sucrose in parallel for sucrose sensitivity testing. Those that were sensitive to sucrose had their 5' end sequenced at the UW Biotech Sequencing Facility to verify insertion and sequence fidelity of the 3HP region. All PCR-amplified genes were also sequenced to confirm their fidelity. Plasmid DNA for transcription reactions was purified using Marligen maxi-prep kits (Marligen Biosciences, Ijamsville, MD). Two methods were used to obtain the linearized DNA to circumvent the limited terminator efficacy, PCR and restriction digest. PCR amplification of Tev-GFP and 3HP-Tev-GFP (3'GCA GAG GGC AGA TAC CAA ATA CTG; 5'CAC ATA CGA TTT AGG TGA CAC). PCR amplification yielded a 2.5 kb linear fragment as shown by figure 5. The other method used endonuclease restriction enzymes NgoMIV and Fsp I (New England) to cut out the 3370 bp DNA segment of interest (figure 6). Following successful acquisition of the linear piece, transcription and translation reaction were taken out following exactly the same procedures described below.

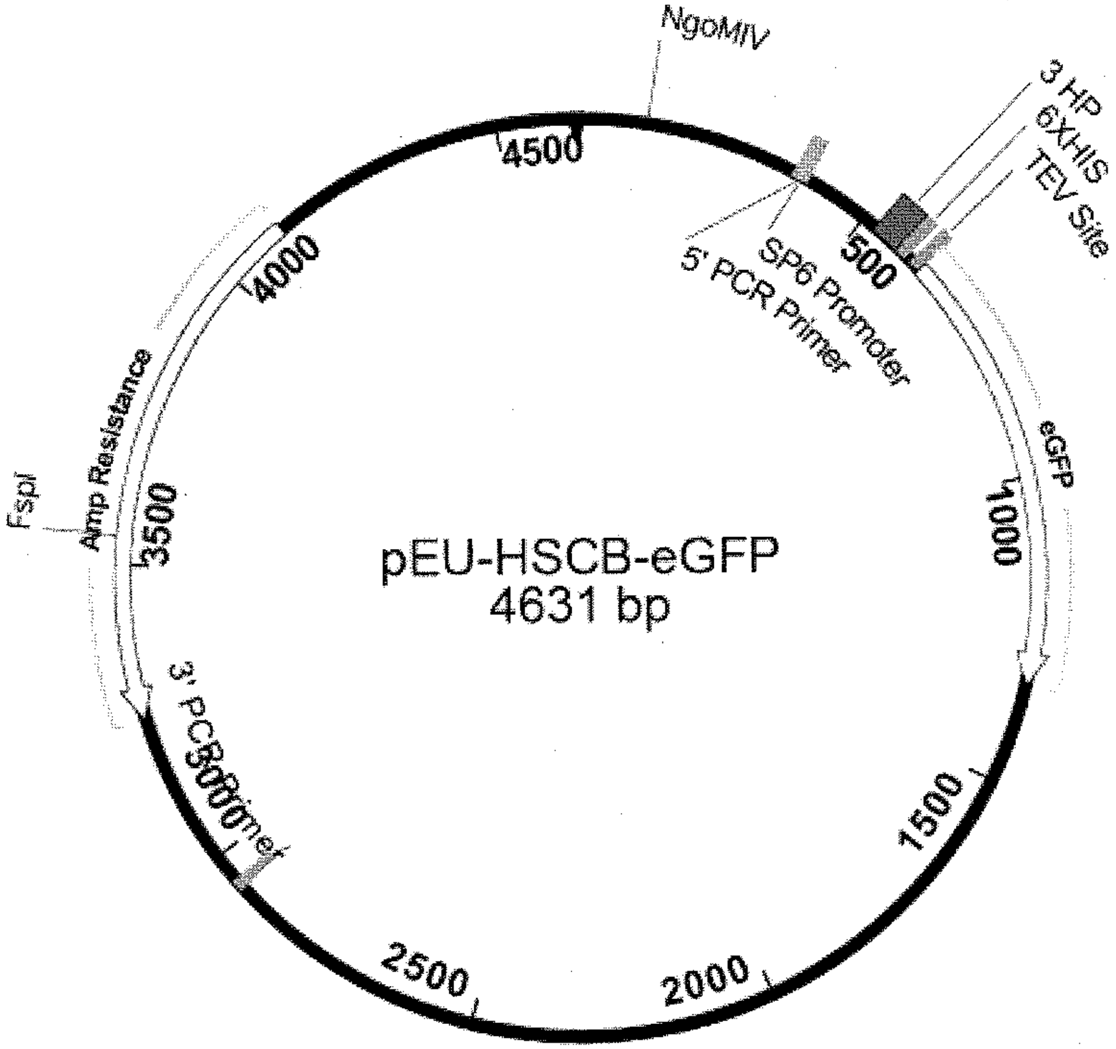


Figure 2: The vector pEU contained 5′-SgfI and 3′-PmeI restriction sites and a toxic selection cassette in the multi-cloning site. This schematic illustration contains the position of 3' and 5' primers (shown in grey boxes) and the two restriction sites used to obtain the linearized DNA, FspI and NgoMIV; SP6 promoter (shown in green box) and the cloning position of 3HP, which is only in the 3HP variant (shown in red box) His-tag and TEV site.

Transcription

The transcription reaction has a total volume of 500 µL, and contains 100 µg of purified plasmid DNA, 80 mM HEPES, pH 7.5, 16 mM magnesium acetate, 2 mM spermidine, 10 mM dithiothreitol, 2.5 mM of each nucleotide triphosphate (ATP, UTP, GTP, CTP), 25 units of RNasin (Promega), 30 units of Sp6 RNA polymerase (Promega). The remainder of the total volume was from deionized and sterilized water. The reaction was incubated at 37 °C for 3 h, at which 10 ul of the reaction was purified using RNeasy kit (Qiagen, CA), and the mRNA concentration was determined by absorbance on a Nanodrop. Based on the concentration of mRNA provided by Nanodrop, the translation reaction was programmed with varying amounts of mRNA (5, 10, 25, 50, 100, 200 ug). Then the tubes were centrifuged at 15,000 rpm in an Allegra 21R centrifuge (Beckman Coulter, Fullerton, CA) and F2402H rotor to remove the magnesium pyrophosphate by product. The supernatant, containing mRNA, was transferred to a 1.7 mL centrifuge tube containing 1/5 transcription reaction volume of 6 M ammonium acetate, and the mRNA was precipitated by addition of 3 times the reaction volumes of 100% EtOH. After missing well, the sample was centrifuged at 15,000 rpm for 20 min at 4 °C, and the mRNA pellet was washed with 600 µL of 70% ethanol. After another centrifugation the supernatant was carefully removed and the pellet was air dried.

Cell-free translation and protein quantification

The translation mixture has a total volume of 50 µL, and contains 30 mM HEPES, pH 7.8, 100 mM potassium acetate, 2.7 mM magnesium acetate, 1.2 mM ATP, 0.25 mM GTP, 16 mM creatine phosphate, 0.4 mM spermidine, 0.3 mM of each amino acid, 0.8 mg/mL of creatine kinase and 24 units of RNasin. Wheat-germ extract (15 µL, Cell Free Sciences, Yokohama, Japan) was added from a concentrated commercial preparation to a final OD600 of 60 and the remainder of the total volume was from deionized and sterilized water. The purified mRNA pellet was dissolved in the translation mixture and the reaction was placed into a 12 MWCO dialysis cup (Biotech International, Perth, Australia) suspended in a buffer reservoir containing all of the above reagents except mRNA, creatine kinase, RNasin and wheat germ extract. The reaction was incubated at 26 °C for 16 h. Protein levels were determined by eGFP fluorescence level with excitation wavelength at 488 nm and emission wavelength at 509 nm, using GFP standard prepared by the Center for Eukaryotic Structural Genomics. A standard fluorescence versus concentration curve was generated, in order to quantify eGFP production.

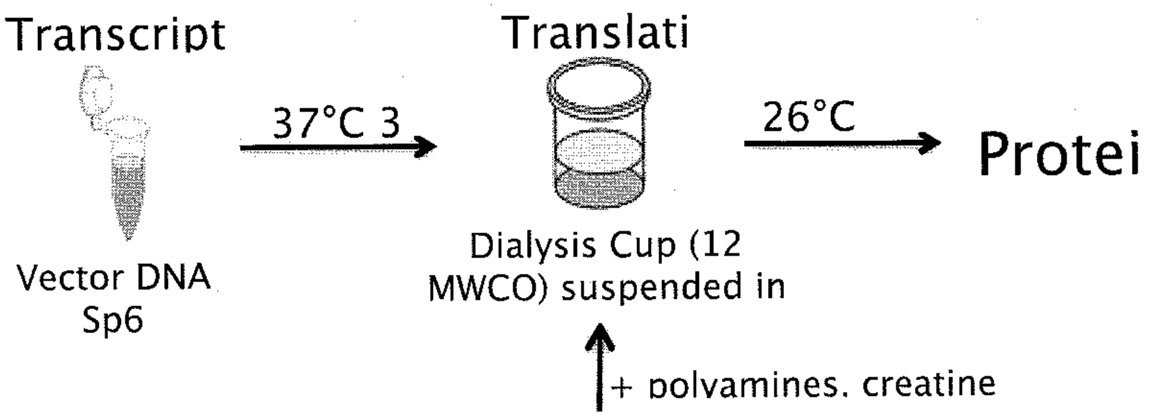


Figure 3: Schematic illustration of dialysis method. The translation mixture has a total volume of 50 µL, and contains 30 mM HEPES, pH 7.8, 100 mM potassium acetate, 2.7 mM magnesium acetate, 1.2 mM ATP, 0.25 mM GTP, 16 mM creatine phosphate, 0.4 mM spermidine, 0.3 mM of each amino acid, 0.8 mg/mL of creatine kinase and 24 units of RNasin. Wheat-germ extract (15 µL, Cell Free Sciences) was added from a concentrated commercial preparation to a final OD600 of 60 and the remainder of the total volume was from deionized and sterilized water. The purified mRNA pellet was dissolved in the translation mixture and the reaction was placed into a 12

MWCO dialysis cup suspended in a buffer reservoir containing all of the above reagents except mRNA, creatine kinase, RNasin and wheat germ extract.

Results

The decoupled wheat germ transcription/translation system made it possible to test whether 3HP motif enhanced protein production through the transcription or translation reaction. In the original report, Mureev's group used coupled transcription/translation reaction, batch mode. Therefore, it was not possible to determine where the enhancement came from. I took two approaches to test these separate processes. I monitored mRNA production as well as eGFP yield based on plasmid and linear DNA. One approach tested how plasmid DNA affect transcription and subsequent translation level, the other approach tested how linearized DNA piece affect transcription and translation level in wheat germ extract based CF translation system. The reason why we did linear DNA test was because pEU-HSCB-eGFP polymerase terminator only had 50% efficiency (unpublished result) as shown in figure 4. It was assumed that the mRNA post terminator was not translated.

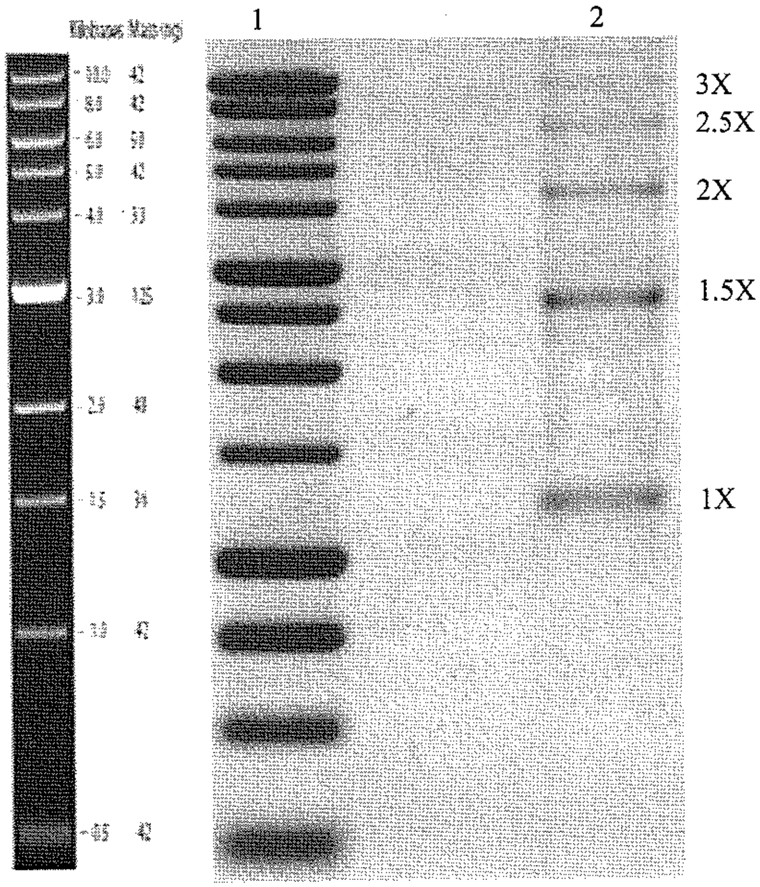


Figure 4: Lane 1: 1kb DNA ladder, BioLab; Lane 2: 1X means mRNA was transcribed around the circle once, 1.5X means mRNA was transcribed around the circle 1.5 times passed the terminator, etc. Due to the usage of DNA ladder, this was only a qualative illustration.

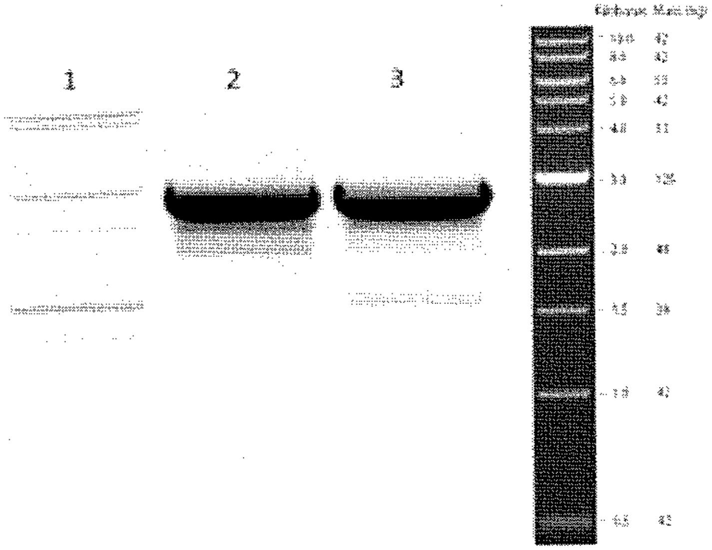
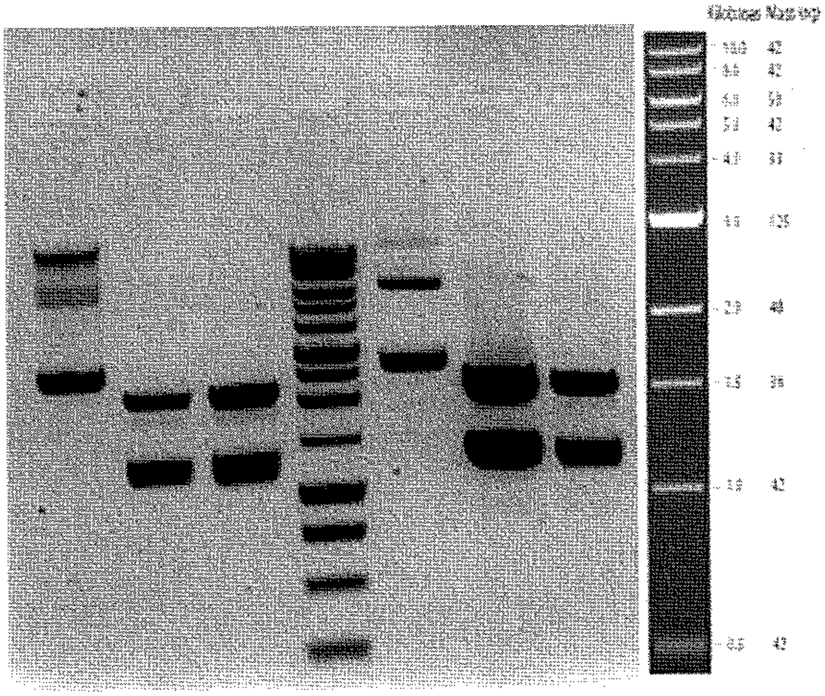


Figure 5: PCR result from amplification of linear piece of DNA according to the methods. 1: 1kb DNA ladder; 2: GFP; 3: 3HPG, 2500 bp. Due to the incompatibility of Qiagen PCR purification kit and Nanodrop, the amplified product was not detected properly. Hence, no further experiments were done with these products.


Kilobases Mass (ng)
10.0 42
8.0 42
6.0 50
5.0 42
4.0 33
3.0 125
2.0 48
1.5 36
1.0 42
0.5 42

100	466	588

Table 1: mRNA of both 3HPG and GFP were transcribed. The concentration of mRNA post transcription was recorded using Nanodrop after purification with commercially available RNA purification kit. 3HP was observed to enhance mRNA transcription at lower DNA input but not as significant as DNA input increased. All values are in ug.

Table 1: Visualization of data in a ±1 s.e.m. of three replicates, plotted in log scale.

However, with linearized DNA, obtained using enzymatic digestion described above, we did not observe same level of increase in transcription as shown in table 2.

DNA input	**GFP**	**3HPG**
100	324±71	287±138
100	465±71	476±138
100	400±71	556±138

Table 2: Expression was tested three times with linearized DNA, obtained using enzymatic digestion described above. No significant level of transcription increases was described. All values are in ug.

We did not observe significant enhancement of 3HPG in translation reaction in the case of DNA plasmid or in the case of linearized DNA as shown in table 3 and table 4.

mRNA Input	**GFP 1**	**GFP 2**	**3HPG1**	**3HPG2**
5	4.4	6.0	5.7	6.4
10	8.1	10.5	8.6	10.4
25	8.6	8.1	8.8	8.1
50	7.4	4.0	8.8	3.9
100	3.4	1.3	3.0	1.2
200	0.2		0.1	

Table 3: DNA plasmids were transcribed and then translated to protein. eGFP fluorescence level was detected using fluorescence plate reader. Expression was tested twice and the data is labeled 1 and 2 to differentiate. We did not observe significant enhancement of 3HP in translation reaction. All values are in ug.

mRNA Input	GFP 1	GFP 2	GFP 3	3HPG 1	3HPG 2	3HPG 3
5	4.5	15.8	15.8	5.2	14.6	15.4
10	6.6	16.6	19.4	5.1	15.2	14.8
25	6.8	16.5	14.5	7.5	13.3	15.2
50	3.9	9	4.2	7.5	7.9	9.3
100	1.6	3.7	4.5	7.5	2.5	3.6
200	1.3	1.3	1.3	2	1.5	2.8

Table 4: Comparison of GFP and 3HPG protein production in linearized DNA, obtained using enzymatic digestion described above. Linearized DNA was transcribed then translated into protein. eGFP fluorescence level was detected, and protein expression was calculated using the same procedures as decribed in the case of plasmid DNA. Expression was tested three times and the data is labeled 1, 2 and 3 to differentiate. All values are in ug.

Figure 6: Visualization of data in ±1 s.e.m. of three replicates. Although there were discrepancies in expression levels among the three runs, the trend of expression level was consistent: 3HP did not significantly increase protein production. The discrepancies could be attributed to human error and reagent preparation efficiency. The decreases in protein expression level as mRNA input increased may be attributed to significant loss of reaction volume due to solvation of mRNA.

Discussion:

In this work, I tested whether 3HP motif enhanced protein production by examining both transcription and translation in a decoupled wheat germ CF system. I took two approaches to test these separate processes. One approach was taken to test how DNA plasmid affect transcription and subsequent translation level, the other approach was taken to test how linearized DNA piece affect transcription and translation level in wheat germ extract based CF translation system. We observed differences in transcription levels with DNA plasmid but not with linearized DNA, which might be due to the problem associated terminator efficiency. pEU-HSCB-eGFP polymerase terminator only had 50% efficiency. It was assumed that the mRNA post terminator was not translated. In terms of translation level, we did not observe significant enhancement by the putative species-independent translational sequences (SITS) 3HP with either linear or plasmid constructs. One thing to be noted was although there were discrepancies in expression levels among the three runs, the trend of expression level was consistent. The discrepancies could be attributed to human error and significant reaction volume loss. We observed that when mRNA concentration was too high, much less protein was produced. We assigned this observation to the significant loss of reaction volume due to solvation of mRNA molecules because when mRNA input exceeded 100 ug per 50 ul translation reaction, mRNA was very hard to get dissolved in translation mixture and a 40% of reaction volume loss was observed. In short, we observed no comparable increases in either mRNA or protein yield associated with 3HP, which may be attributed to practical reasons listed below. In the publication noting that 3HP increased protein production, the lab used the coupled Promega wheat germ CF transcription and translation system which runs in batch mode. Batch mode has been reported to yield consistently lower in protein production compared to the dialysis wheat germ CF transcription and translation system, 10-80 ug/ml versus 0.5-1 mg/ml *(12)*. Batch mode only offers very limited nutrient supply, compared to the dialysis method used in this work since translation reaction in the dialysis cup is constantly supplied by the buffer underneath. Also, to run a coupled wheat germ CF reaction, the concentration of the key component, magnesium, needs to be compromised since transcription reaction requires magnesium concentration specifically to be 11.0 mM, whereas translation reaction requires magnesium concentration to be 2.7 mM, which is inhibitory to transcription reaction *(13)*. Hence, coupling reduces yield by compromising both steps. The higher efficiency of the decoupled wheat germ cell-free transcription/translation system used in our laboratory might have masked the effect of the putative translational enhancer 3HP structure used in the coupled-batch mode CF protein production system reported previously.

Acknowledgement

Professor Brian Fox

Michael Goren

Dr. Russel Wrobel

All lab members at Fox lab

Center for Eukaryotic Structural Genomics

References:

1. Wagner S, Bader ML, Drew D, de Gier JW. (2006) Rationalizing membrane protein overexpression. *Trends Biotechnol, 24*, 364–371.

2. Blommel, P. G., Martin, P. A., Wrobel, R. L., Steffen, E., and Fox, B. G. (2006) High efficiency single step production of expression plasmids from cDNA clones using the Flexi Vector cloning system, *Protein Expr Purif 47*, 562-570.

3. Y. Chang, G. Wesenberg, C.A. Bingman, B.G. fox (2008) In vivo inactivation of mycobacterial integral membrane stearoyl-CoA desaturase DesA3 by a C-terminal specific degradation process, *J. Bacteriol., 190,* 6686-6696.

4. Ezure T, Suzuki T, Shikata M, Ito M, Ando E, Utsumi T, Nishimura O, Tsunasawa S. (2010) Development of an insect cell-free system. *Curr Pharm Biotechnol. 3,* 279-284.

5. Endo, Y., and Sawasaki, T. (2006) Cell-free expression systems for eukaryotic protein production, *Curr Opin Biotechnol 17*, 373-380.

6. Takai, K., Sawasaki, T., and Endo, Y. (2010) The Wheat-Germ Cell-Free Expression System, *Current Pharmaceutical Biotechnology 11,* 272-278.

7. Park, N., Kahn, JS., Hartman, MR., Funabashi, H., Xu, J., Um, SH., Luo, D. (2009) High-yield cell-free protein production from P-gel, *Nat Protoc 12*, 1759-1770.

8. Jewett, MC., Calhoun, KA., Voloshin, A., Wuu, JJ., Swartz, JR. (2008) An integrated cell-free metabolic platform for protein production and synthetic biology, *Mol Syst Biol 4*, 1038-1047.

9. Kim, KH., Yang, JK., Waldo, GS., Terwilliger, TC., Suh, SW. (2008) From no expression to high-level soluble expression in *E. coli* by screening a library of the target proteins with randomized N-termini. *Methods Mol Biol, 426,* 187-195.

10. Mureev, S., Kovtun, O., Nguyen, UT., Alexandrov, K. (2009) Species-independent translational leaders facilitate cell-free expression, *Nat Biotechnol 8*, 747-752.

11. Frederick, RO., Bergeman, L., Blommel, PG., Bailey, LJ. et al. (2007) Small-scale, semi-automated purification of eukarotic proteins for structure determination. *J Struct Funct Genomics, 4,* 153-166.

12. Cell-Free with Increased Solubility, Michael R. Slater, Ph.D., Robin Hurst, M.S., Becky Pferdehirt, B.S., Doug White, M.S., Andrew Niles, M.S., Natalie Betz, Ph.D., and Elaine Schenborn, Ph.D., Promega Corporation

13. Goren, M. A., and Fox, B. G. (2008) Wheat germ cell-free translation, purification, and assembly of a functional human stearoyl-CoA desaturase complex, *Protein Expr Purif 62*, 171-178.6.