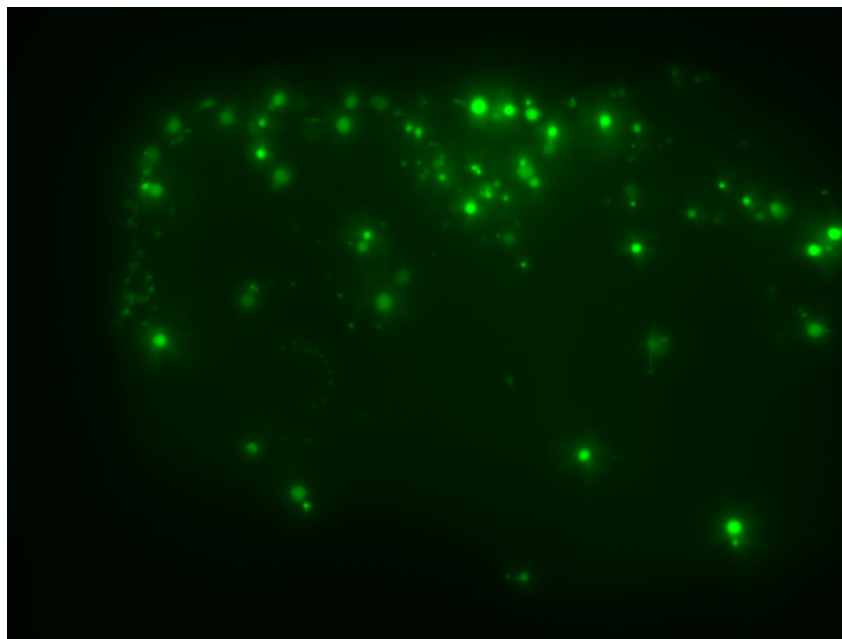


Identification Of A General Targeting Motif Guiding Interaction Between Nuclear-Encoded Plastid Precursors And The Translocon Of The Plastid Outer Membrane.

Erika Sanders
Senior Thesis
May 8th, 2017



BCMB Department
Dr. Barry Bruce, Dr. Kristen Holbrook
The University of Tennessee, Knoxville

ABSTRACT

Plant cells are defined by a collection of organelles and the membranes that surround them. When new proteins are synthesized in the cytosol, they must be transported to their destination at a specific organelle. Proteins targeted to the chloroplast contain a short N-terminal region called the *transit peptide (TP)*. This region acts as a “zip code” recognized by translocons on the outer membrane of the chloroplast (Toc) as the first step of chloroplast protein import. The TP is particularly complex due to its lack of consensus sequence, but there are several physicochemical motifs that contribute to targeting and import. One of such motifs is a region that interacts with the Toc receptor proteins, which appear to promote TP binding, but cannot support import alone. This region is termed ‘FGLK,’ defined by the character of the amino acids it contains. To investigate the universality of this motif we designed a heuristic bioinformatics approach to identify other precursors containing this motif in the *Arabidopsis* genome. The consolidated output from localization prediction tools was integrated into a scoring algorithm to select the final 7 most confidently predicted chloroplast proteins, while avoiding potentially dual-targeted proteins to the mitochondria. We developed an *in vivo* assay to assess the import efficiency of precursor protein constructs by using a chimeric transit peptide-fluorescent protein (FP) construct. The FGLK domains have been substituted in each model TP to analyze their subcellular localization using fluorescence microscopy. The resulting fluorescence will be quantified and analyzed to further the understanding of the functional motifs in TPs.

INTRODUCTION

The eukaryotic cell is characterized by its organelles, that coordinately regulate the cell’s functions. Organelles require a constant supply of specific proteins in order to maintain their biological role within the cell and the organism as a whole. According to the Central Dogma of biology, proteins are transcribed in the nucleus, and translated in the cytosol. Then, they are delivered to their site of functionality. The study of how proteins move from one cellular compartment to another is called **protein translocation**. For double-membrane bound organelles such as mitochondria and chloroplasts, it is especially difficult to discern the recognition and translocation steps for protein import. In mitochondrial-

targeted proteins, scientists have discovered a conserved domain in the primary sequence termed the Signal Recognition Peptide. Similarly, chloroplasts-targeted proteins contain a region called the Transit Peptide (TP) to be recognized by the nanomachinery at the outer membrane of the chloroplast. In both cases, these N-terminal extensions are cleaved upon entering the membrane complexes. While this process has been well-studied in mitochondria, there is much to be learned about protein translocation in the chloroplast (Schleiff, 2010).

As mentioned above, the TP acts a sort of “zip code” to localize the protein. Translocons on the Outer Chloroplast membrane (Toc) are responsible for recognizing the TP. The divergent nature of TP sequences makes it difficult to extend the results of various domain studies in a global manner to all transit peptides (Holbrook et. al., 2016). However, recent studies have focused on the functionality of physiochemical motifs rather than focusing on the primary sequence. Studies about the structural domains of TP generally include mutagenesis, deletion, and biochemical *in vivo*, *in vitro*, and *in organello* assays to determine the utility of each region in import (Holbrook, K., 2016; Pilon, M., 1995). Recent studies have identified two important N-terminal regions to preprotein recognition. The first is the N-terminal uncharged region. The second is the FGLK motif, which is a physicochemical motif loosely conserved across species.

FGLK region is defined as having the following: an aromatic a.a.; a turn-inducing or helix breaking a.a.; a small nonpolar a.a.; a basic a.a., and lacking any negatively charged a.a. Results from Holbrook’s extensive studies of the FGLK motif in the small subunit of rubisco and ferredoxin implicate the importance each characteristic type of amino acid involved in the FGLK motif (2016). Using her work as a model, we have further developed the *in vivo* assay using onion epidermal cells in biolistic transformation to assess the global implications of this motif in other chloroplast-targeted proteins.

Our results utilize a proven method for testing the global implications of the FGLK motif in seven carefully selected chloroplast-targeted proteins. We have also developed a complex cloning strategy to analyze the

preproteins in an *in vivo* assay using *allium cepa* epidermal cells. These methods will highlight the importance of the FGLK motif in a group of rigorously selected preproteins resembling the small subunit of rubisco studied extensively in previous reports (Holbrook, K., 2016). These import assays will provide insight to the recognition of the physicochemical motif FGLK in *arabidopsis thaliana* and provide insight into the mechanistic details of the Bimodal Import Model of transit peptides.

RESULTS

Bioinformatic Selection

In order to expand information about the universality of the FGLK motif, we chose proteins that allowed us to control for as many similarities to ssTP as possible. We still needed proteins that were confidently predicted to the chloroplast, belonged to the *arabidopsis thaliana* species, were of a similar TP length, and contained at least one FGLK motifs. However, we wanted to test preproteins with different functions inside the chloroplast. To make our results as useful for as many researchers as possible, we also chose a variety of well-studied proteins, based on the number of hits each of the finalists had in NCBI Blast and PubMed database.

Analysis of the *Arabidopsis* genome revealed 912 highly confidently predicted plastid precursors, of which, 327 proteins contain a similar length of transit peptide at ~50-60 a.a. and 231 precursors contain at least one 'FGLK' motif. We selected a subset of 24 proteins based on the presence of FGLK motif around residues 28-39, the functional annotation (Plant Proteome Database and Phytozome), proteomic localization (Plant Proteome Database), level of expression (*Arabidopsis* eFP Browser), and coexpression analysis (ATTED-II). These precursors were then further analyzed using localization predictions (TargetP, ChloroP 1.1, iPSORT, Predotar, PredSL, ProtComp, Protein Prowler). The output from these prediction tools was integrated into a mathematical scoring algorithm to select the 7 most confidently predicted chloroplast proteins, while avoiding potentially dual-targeted proteins to the mitochondria.

ChloroP 1.1	Chloroplast Localization	[Chloro loc]
	Cleavage Site (CS-Score)	
	"cTP"	
	"cTP-length"	
TargetP	Length	[cTP – (mTP + SP + other)] – TargetP [RC]
	cTP	
	mTP	
	SP	
	Other	
	RC	
iPSORT	Localization	[-1 for mito or +1 for chloro]
	Target Sequence	
Predotar	Chloroplast Localization	[chloro - (mito+ ER+ other)]
	Mitochondria Localization	
	ER Localization	
	Other Localization	
PredSL	Chloroplast Localization	[chloro loc- (mito loc + SP loc)]
	Mitochondria Localization	
	Secreted Protein	
	Cleavage Site	
ProtComp	Chloroplast Localization	[chloro loc- (mito loc + PM loc + nuc loc+ perox loc)/9]
	Mitochondria Localization	
	Plasma Membrane	
	Nuclear Localization	
	Peroxisome Localization	
Protein Prowler	Chloroplast Localization	[chloro loc- (mito loc + SP loc + other loc)]
	Mitochondria Localization	
	Secreted Protein	
	Other Localization	

Figure 1. Seven bioinformatic localization programs were used to predict the localization of 24 proteins to the chloroplast.

$$[\text{Chloro loc}] + [\text{cTP} - (\text{mTP} + \text{SP} + \text{other})] - [\text{RC}] + [-1 \text{ for mito or } +1 \text{ for chloro}] + [\text{chloro} - (\text{mito} + \text{ER} + \text{other})] + [\text{chloro loc} - (\text{mito loc} + \text{SP loc})] + [\text{chloro loc} - (\text{mito loc} + \text{PM loc} + \text{nuc loc} + \text{perox loc})/9] + [\text{chloro loc} - (\text{mito loc} + \text{SP loc} + \text{other loc})]$$

Equation 1. Integration of bioinformatic localization results (Fig. 1). Combining these results, we created the algorithm depicted in Equation 1. The seven predictions were weighted such that each prediction to the chloroplast was a positive. Due to frequent localization to both the chloroplast and the mitochondria, each prediction to the mitochondria and all other cellular compartments was assigned a negative score.

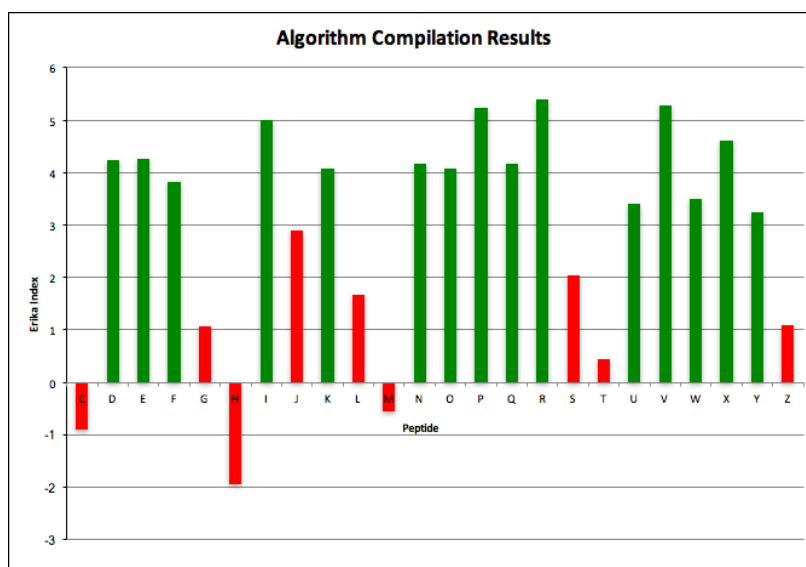


Figure 2. Localization predictions to the chloroplast. The range of the values was -1.94 to 5.41, while the average localization prediction was 2.93. The proteins that had a prediction value for greater than this average are colored in green. Those falling below the average are graphed in red. Each of the 24 proteins are denoted by a capital letter on the x-axis.

Equation 1 allowed us to assign each of the finalist preproteins with a value. The highest seven scored proteins were chosen to be studied in our *in vivo* analysis. With protein final selections made (Figure 3), attention turned to the complicated cloning strategies necessary to get our proteins from their concatemerized state into individual YFP expression vectors.

TP 1	Glutamine-5-phosphoribosylpyrophosphate amidotransferase	MAATSSISSLSLNAPNKLNNNNNNKPHRFLRNPFLNPSSSSFSPLPASISSSSSPSFPL
TP 2	Glucose-6-phosphate dehydrogenase	MATHSMIIPSPSSSSSLATAASPFKETLPLFSRSLTFPRKSLFSQVRLRFAEKHSQLD
TP 3	CHLI subunitol magnesium chelatase	MASLLGTSSSAIWASPSLSSPSKPSPPICFRPGKLGSKLNAGIQIRPKKNRSRYHVSVMNVATEINS
TP 4	2C-methyl D erythritol 2,4-cyclodiphosphate synthase	MATSSTQLLLSSSLFHSQITKKPFLLPATKIGVWRPKKSLSLSCRPSASVSAASSAVDVNE
TP 5	Beta-carotenoid hydroxylase	MAAGLSTIAVTLKPLNRRSFSANHPISTAVFPPLRFNGFRRRKILTVCFVVEERKQSSPMD
TP 6	Allene-oxide Cyclase	MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSRSFHNQRISSNSPGLSTRARSTTSSTGG
TP 7	Rubisco activase	MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTFLGKKVTVSRFAQSNKKSNNGSFKVLAVKEDKQTDG

Figure 3. Final preproteins selected. The black amino acids indicate the residues within the predicted TP length, as predicted by ChloroP 1.1 in Figure 1. The red amino acids depict the first ten amino acids of the mature domain. Note the functional variety of each of the selected preproteins in *arabidopsis thaliana*. Proteins will further be referred to by the number in the far-left column TP1-TP7.

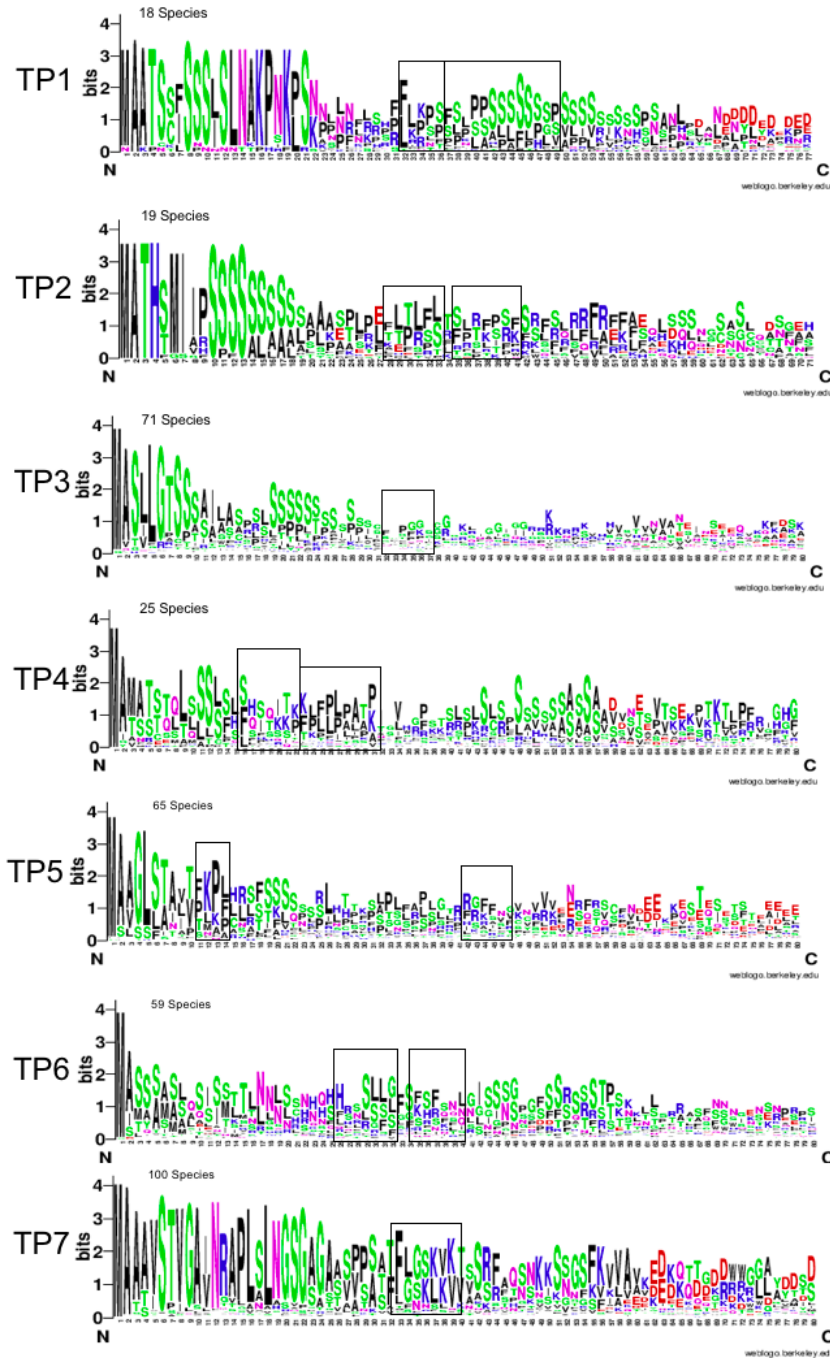


Figure 4. Logo plot of transit peptides according the chosen preproteins. Highly conserved FGLK regions are boxed. Notice that most of the preproteins have two FGLK motifs. In some instances, even more boxes could have been drawn if you wanted to include FGLK motifs more loosely. This logo plot gives our results a more universal application than in previous studies.

Mutation Strategy

Previous *in organello* work on the FGLK motif has demonstrated the importance of both FGLK domains in import in ssTP (small subunit of rubisco precursor protein). However, *in vivo* analyses suggested that mutation of one FGLK motif in the TP region was not significant enough to abrogate import of the protein, suggesting some redundancy to the motif *in vivo* (Holbrook, K., et. al, 2016). These results led us to create mutations in FGLK domains individually (Fig. 5, Rule 2 - 3) and combined (Fig. 5 - Rule 4).

Further *in vitro* studies confirmed the importance of all of the defined residues in the FGLK motif in ssTP. One of the most prominent trends found in *in vivo* analysis of ssTP was in the positive charge of N-terminal basic residues to the FGLK motif (Holbrook, K., et. al., 2016). Rule 5 was designed to further investigate this finding through mutation of the FGLK-proximal RK residues. Rule 6 was generated to test a region of unusual amino acid content, labeled “Miscellaneous” in Figure 5, above.

Regions of interest were chosen using a series of multiple alignments for each of the selected proteins across species. These regions were replaced with Serine residues, rather than the typical Alanine residue. Substitution was the mutation of choice, because it maintained the length of each wild type transit peptide (Bionda et al., 2010). The choice to substitute domains with Serine was two-fold. First, substitutions retain the length of the TP. Second, because Serine is much more abundant in the TP region of chloroplast-targeted proteins than Alanine (von Heijne, et al., 1989). By careful selection of preproteins and their mutations, we will be able to further investigate the role of FGLK motif in *arabidopsis thaliana* proteins.

Mutation Rules:

Rule 1: WT peptide sequence

Rule 2: N-terminal FGLK sequences

Rule 3: C-terminal FGLK sequences

Rule 4: Combined N/C FGLK sequences

Rule 5: Internal (near FGLK domain) RK sequences

Rule 6: Miscellaneous

each of the TP sequences were placed into a fluorescent YFP vector. This cloning process is on-going. All CloneJet vectors have been created. Additionally, 27 of 42 TP-YFP constructs have been confirmed by Colony PCR Screening, followed by sequencing. Considering the 84 total CloneJet and TP-YFP constructs, cloning is 82% complete.

	Rule 1	Rule 2	Rule 3	Rule 4	Rule 5	Rule 6
TP 1						
TP 2						
TP 3	X	X	X	X	X	
TP 4	X	X	X	X	X	X
TP 5	X	X	X	X	X	X
TP 6		X	X	X	X	X
TP 7	X	X		X	X	X

Figure 6. Ligation progress of the TP-YFP pAN 187 vectors. Note that initial cloning into CloneJet vectors has already been confirmed.

Biolistic Transformation of TP-YFP Constructs

Using well described methods in other scientific works, we have tested the importability of many of our TP-YFP constructs in an *in vivo* assay (Holbrook, K., et. al., 2016; Chotewutmontri et al., 2012; Chotewutmontri and Bruce, 2015; Nelson, et. al., 2007). Onion epidermal layers were biolistically transformed using a Bio Rad gene gun. The location of the protein was tracked using a YFP tag under fluorescence microscopy. This protocol involves the conjugation of DNA to tungsten (M-17) particles. The particles are then placed in a high pressure vacuum, where the tungsten-DNA particles are shot into an onion sample. Onion samples are stored in a dark drawer overnight, and then analyzed the following morning using fluorescent microscopy using either a traditional fluorescent microscope or digital microscopy. Results between the two different techniques were interesting, and there is hope that in the future, quantification of import can be accomplished using both microscopy techniques.

Using the digital BZ-X710 fluorescent microscope has allowed us to further optimize this assay. This instrument is a recent biological application of a microscope typically used in industry. Images were taken at a 2X magnification using the GFP setting. The ratio of transformants to the total number of cells was collected to give us the transformation efficiency of varying DNA concentrations. We also introduced a co-localization marker labeled with CFP, that is known to localize to the peroxisomes. This separation of the YFP and CFP channels allowed us to confirm transformation in each of the cells analyzed for transformation efficiency. Results are summarized in Figure 7, below. We also found that refreshing the spermidine stock, increasing DNA concentration, and sonication of tungsten particles prior to dispersion on macrocarriers were crucial steps of biolistic transformation (Fig. 7).

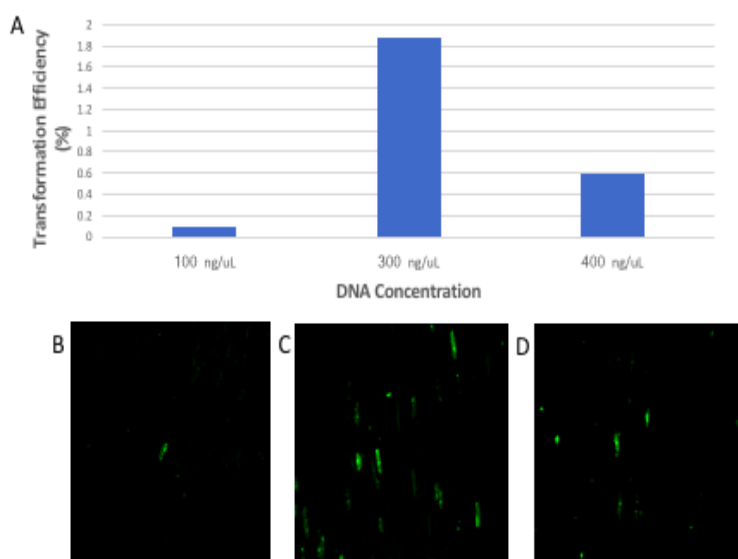


Figure 7. An intermediate DNA concentration (~2.5 ug of DNA) yielded the best transformation efficiency. Cells imaged using Keyence BZ-X710 microscope at 2X magnification. Quantification performed in ImageJ. Results displayed in Microsoft Excel (Panel A). Panels B, C, and D provide representative images for 100 ng/uL, 300 ng/uL and 400 ng/uL DNA concentration.

In addition to optimization of the biolistic transformation, we were also able to explore additional imaging capabilities of the transiently expressed proteins. A series of images taken at 60X magnification is shown in Figure 8. With the imaging capabilities of the Keyence BZ-X710, we were able to capture the entire cell in one frame, using the stitching feature. Further, we were able to take hundreds of images in different planes of the cell at 60X magnification using the Z-stack feature. Imaging at 60X magnification Z-stack took about 10 minutes per cell, versus the hours it would have taken to image each chloroplast using traditional epifluorescent microscopy. Further tests remain to determine whether the Z-stack autofocus feature is sufficient for the rigorous quantification scheme used in this thesis, as well as in previous Bruce lab papers (Holbrook, K., et. al., 2016).

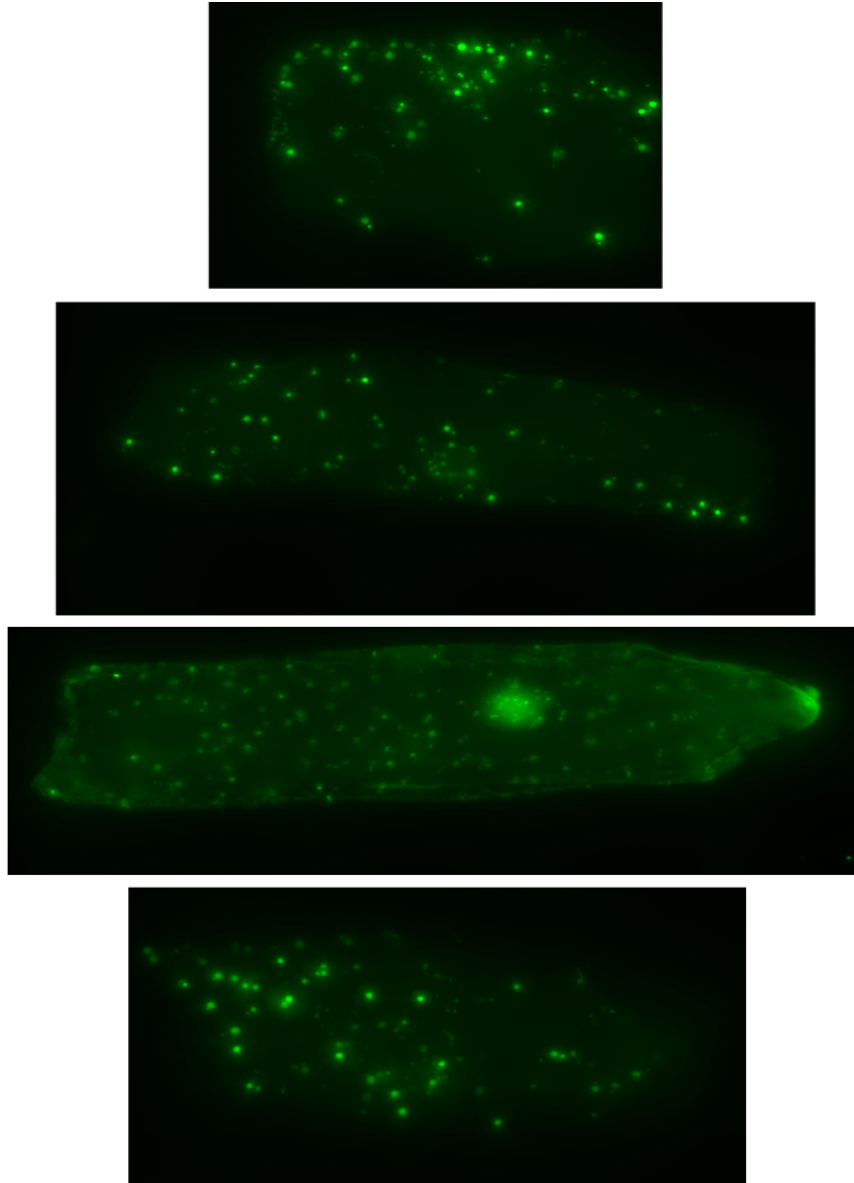
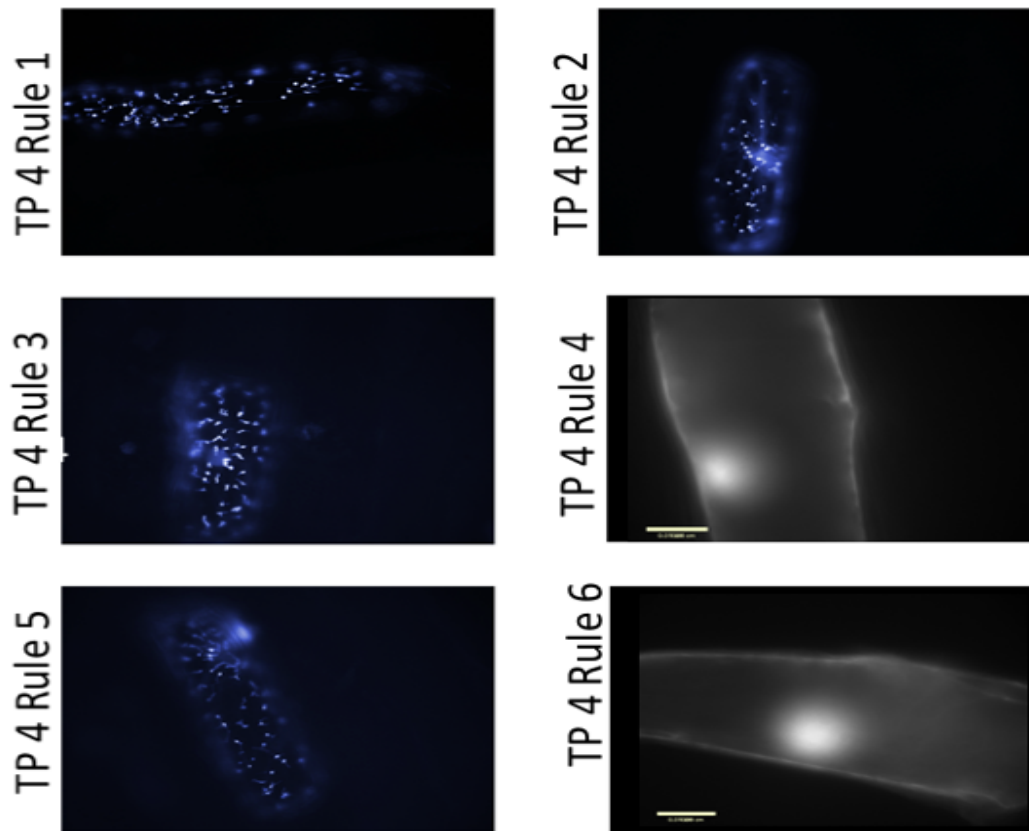


Figure 8. *In vivo* analysis of WT (ssTP-YFP) used as a control for future high magnification imaging using the Keyence BZ-X710. Images taken at 60X magnification of the entire cell.

In Vivo Analysis of TP-YFP Import: Representatives and Quantification

A series of representative images have also been taken of our findings using an epifluorescent microscope at the 40X magnification. The following panels summarize the most important domains to import, by highlighting the most crucial regions in yellow (Figure 9A - 8D). For many of the panels, cloning is ongoing. Quantification of the transient expression has been developed as a fluorescence ratio (plastid:cytosol); the ratio is depicted in Figure 10. This method allows us to account for cell to cell differences in expression, by comparing the fluorescence of the plastid to its surrounding cytosol (Holbrook, K., et. al., 2016). A sample quantification graph has been provided for TP 6, Allene Oxide Cyclase (representatives depicted in Figure 9C, quantification in Figure 10).



2C-Methyl-D-erythritol 2,4-cyclodiphosphate synthase: At1g63970 (EI= 5.007; c-TP= 52)

Rule 1: MATSSTQLLLSSSSLFHSQITKKPFLLPATKIGVWRPKKSLSLSCRPSASVSAASSAVDVNE

Rule 2: MATSSTQLLLSSSSLSSSSSSSPFLLPATKIGVWRPKKSLSLSCRPSASVSAASSAVDVNE

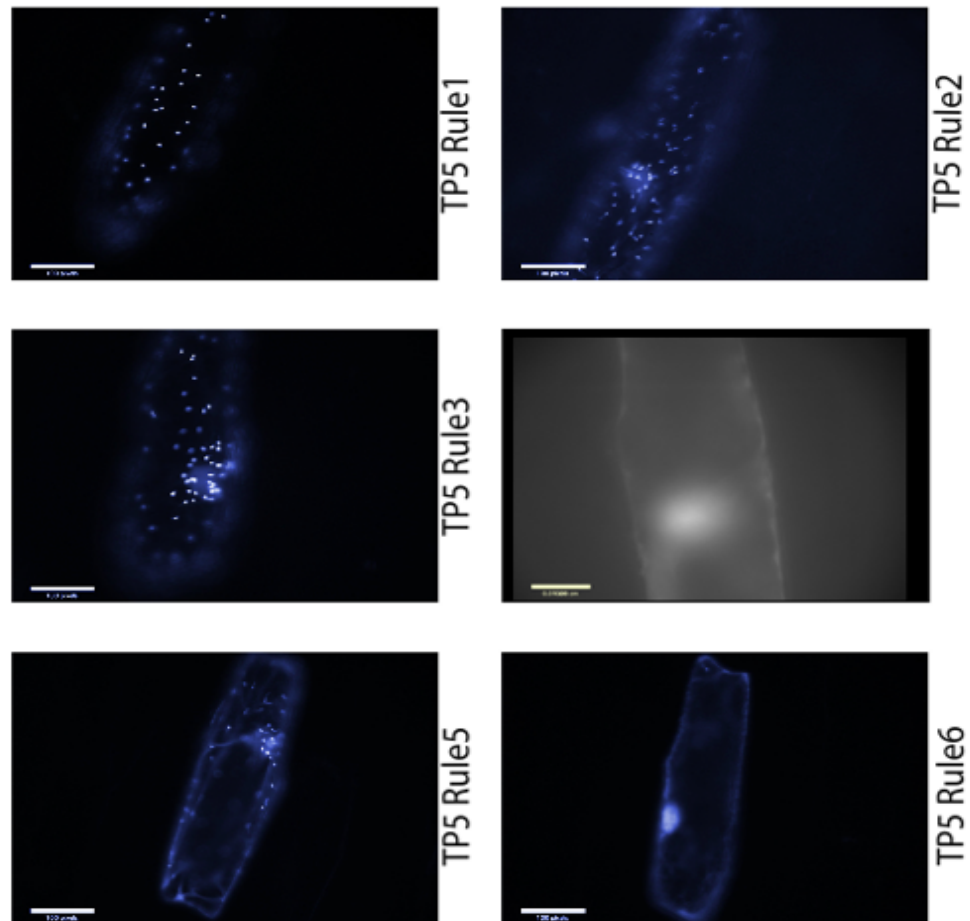
Rule 3: MATSSTQLLLSSSSLFHSQITKKSSSSSSSSIGVWRPKKSLSLSCRPSASVSAASSAVDVNE

Rule 4: MATSSTQLLLSSSSLSSSSSSSSSSSSSSSSIGVWRPKKSLSLSCRPSASVSAASSAVDVNE

Rule 5: MATSSTQLLLSSSSLFHSQITSSPFLLPATSSIGVWSPSSSLSLSCRPSASVSAASSAVDVNE

Rule 6: MATSSTQLLLSSSSLSSSSSSSSSSSSSSSSIGVSSSSSLSLSCRPSASVSAASSAVDVNE

Figure 9 A. These images are representatives for 40 X microscopy using the epifluorescent microscope in the Transit Peptide (cTP = 52) for 2C- Methyl-D-erythritol 2, 4 cyclodiphosphate synthase (TP 4).



Beta Carotene Hydroxylase 2: At5g52570 (EI= 4.611; c-TP= 52)

Rule 1: MAAGLSTIAVTLKPLNRSSFSANHPISTAVFPPSLRFNGFRRRKILTVCFVVEERKQSSPMD

Rule 2: MAAGLSTIAVTLKPLNRSSFSANHPISTAVSSSSSSSGFRRRKILTVCFVVEERKQSSPMD

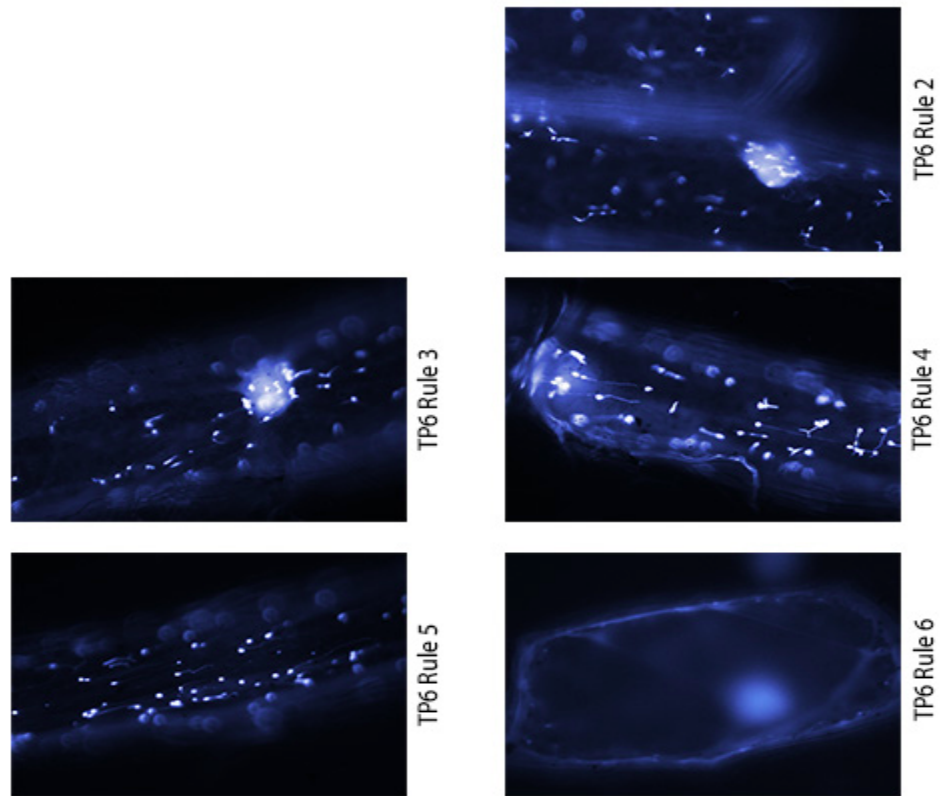
Rule 3: MAAGLSTIAVTLKPLNRSSFSANHPISTAVFPPSLRFNSSSSSILTVCFVVEERKQSSPMD

Rule 4: MAAGLSTIAVTLKPLNRSSFSANHPISTAVSSSSSSSSSSSSSILTVCFVVEERKQSSPMD

Rule 5: MAAGLSTIAVTLKPLNRSSFSANHPISTAVFPPSLSFNGFSSSSILTVCFVVEERKQSSPMD

Rule 6: MAAGLSTIAVTLKPLNRSSFSANHPISTAVSSSSSSSSSSSSSILSSSSSEERKQSSPMD

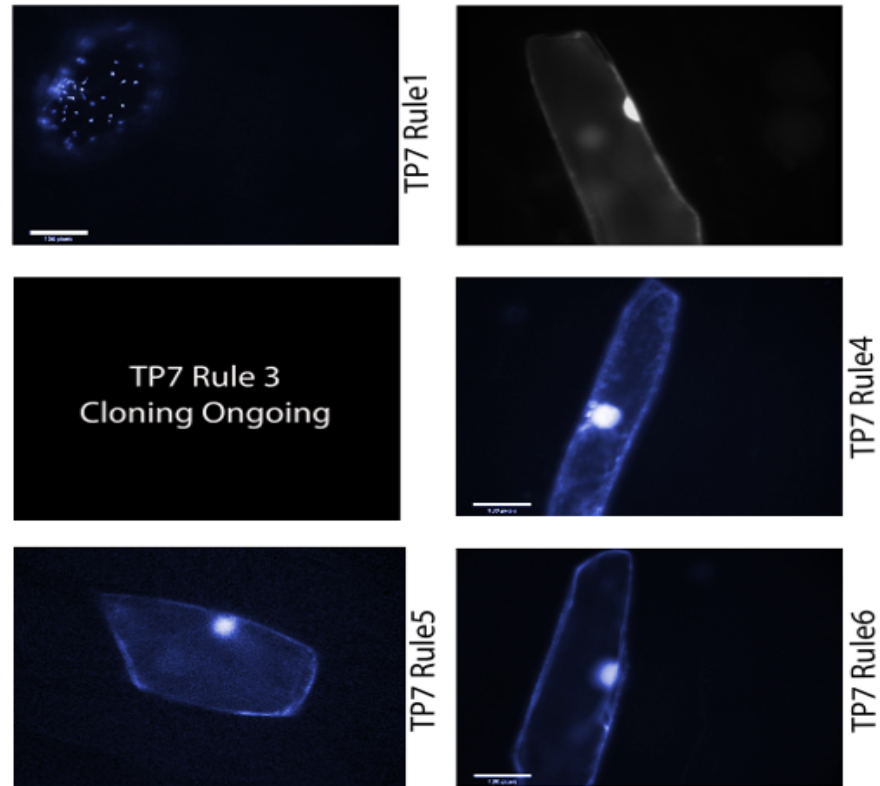
Figure 9 B. These images are representatives for 40 X microscopy using the epifluorescent microscope in the Transit Peptide (cTP = 52) for Beta Carotene Hydroxylase 2 (TP5).



Allene Oxide Cyclase (At1g13280) (EI= 4.274; c-TP= 52)

Rule 1: MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSRSFHNQRISSNSPGLSTRARSTTSSTGG
 Rule 2: MIMASSAAASISMITLRNLSRNHQSHQSTSSSSSSSFHNQRISSNSPGLSTRARSTTSSTGG
 Rule 3: MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSRSSSSSSISSNSPGLSTRARSTTSSTGG
 Rule 4: MIMASSAAASISMITLRNLSRNHQSHQSTSSSSSSSSSSSSISSNSPGLSTRARSTTSSTGG
 Rule 5: MIMASSAAASISMITLRNLSRNHQSHQSTFLGFSSSFHSSS ISSNSPGLSTRARSTTSSTGG
 Rule 6: MIMASSAAASISMITLSSLSSSSSSSSSTFLGFSSSFHSSS ISSNSPGLSTSSARSTTSSTGG

Figure 9 C. These images are representatives for 40 X microscopy using the epifluorescent microscope in the Transit Peptide (cTP = 52) for Allene Oxide Cyclase (TP 6). Additionally, results have been quantified for biolistically transformed transit peptides.



RuBisCo Activase: At2g39730 (EI= 4.073; c-TP= 58)

Rule 1: MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTFLGKKVTVSRFAQSNKKSNGSFKVLAVKEDKQTDG

Rule 2: MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTSSSSSVTVSRFAQSNKKSNGSFKVLAVKEDKQTDG

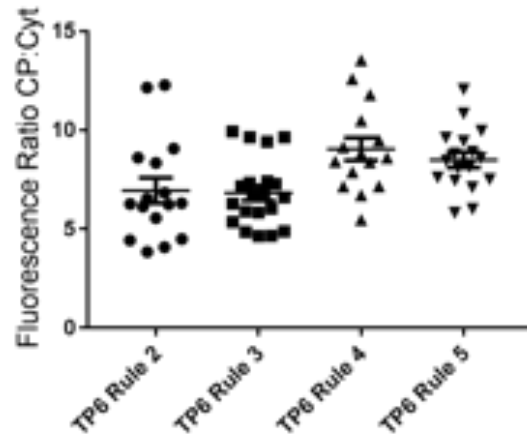
Rule 3: MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTFLGKKVTVSSSSSSSSSNGSFKVLAVKEDKQTDG

Rule 4: MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTSSSSSVTVSSSSSSSSSNGSFKVLAVKEDKQTDG

Rule 5: MAAAVSTVGAINRAPLSLNGSGSGAVSAPASTFLGSSVTVSSFAQSSSSSNGSFKVLAVKEDKQTDG

Rule 6: MAAAVSTVGAISSAPLSLNGSGSGAVSAPASTFLGSSVTVSSFSSSSSSSGSFVLAVKEDKQTDG

Figure 9 D. These images are representatives for 40 X microscopy using the epifluorescent microscope in the Transit Peptide (cTP = 58) for Rubisco Activase (TP 7).



$$\frac{\text{Pixel Intensity } (C_p)}{\text{Pixel Intensity } (C_{\text{cyt}})} = \text{Plastid Localization Ratio}$$

Figure 10. Quantitative representation of TP 6 (Allene Oxide Cyclase). based on analysis of the plastid versus cytosolic pixel intensity. A higher ratio denotes a higher degree of plastid localization. TUsing fluorescence microscopy, we can visualize the sublocalization of an SStp-FP chimeric construct in plant cells, and use imaging software to analyze the pixel intensity of the plastid versus the cytosol. The top panel shows expression of TP-YFP in Mutations 2, 3, 4, 5, and 6. TP 66 did not import; therefore, no ratios are shown.

DISCUSSION

Using a bioinformatic approach, we have identified a series of seven preproteins in which to use in an *in vivo* assay. Comparing our chosen preproteins across species, our logo plot has suggested the conservation of FGLK sequences in *Arabidopsis* and other species. These preproteins are all confidently predicted to localize to the chloroplast, while avoiding the mitochondria. We are confident with this approach that we have taken every possible precaution in our selection process. The variety of functions these proteins have in the chloroplast is also a step toward understanding the universality of the FGLK motif in the recognition step of preprotein translocation.

We are hesitant to draw conclusions about *which* amino acids of the FGLK domain and surrounding areas are universally to import until more data is collected. In the meantime, preliminary representative images

show the varying degree of sensitivity that each preprotein shows to Serine-substitution in the FGLK region. For some preproteins, there does not appear to be an obvious change in import according to our predictions concerning the role of *both* FGLK motifs (Figure 9C). However, there seems to be some evidence for the importance of basic residues R and K near the FGLK motif, given that Figures 9 A - D demonstrated a reduction of import when the basic residues in the miscellaneous mutations were made. Perhaps these basic residues are comparably important in the recognition step of chloroplast protein translocation. In contrast, other preproteins have exhibited extreme sensitivity to *any* mutation of the TP sequence, such as the case for Rubisco Activase (Figure 9D). Thus far, any mutation to the FGLK sequence in Rubisco activase has abrogated import. This is of particular interest given the relationship between Rubisco activase and the best-studied preprotein, ssTP. Both of these proteins are among the most abundant proteins in plants and the most important in photosynthesis. Their relationship is bound to be important for future studies.

Implications for the Bimodal Import Model

Given what we are in the process of discovering, we anticipate that our results will support the model devised in recent literature (Holbrook, K., et. al, 2016). This Bimodal Import model for chloroplast translocation was developed after extensive *in vivo*, *in vitro*, and *in organellar* research results were compiled. The Tic and Toc complexes are illustrated here, consisting of their functional components. The mature domain of the protein is depicted as an orange twisted region, while the TP is depicted using various significant colors. The green regions indicate the FGLK domains, which are recognized by Toc34 and Toc159 components. After recognition, Toc34 and Toc 159 shift apart to grant the TP access to the Toc75 channel protein, spanning the outer and inner membrane of the chloroplast. The TP is then recognized by proteins in the Tic complex and is “pulled” into the stroma of the chloroplast. This movement across the chloroplast membranes is known as translocation, and it is facilitated by the stromally located ATPases (HSP70). We look forward to the contribution our results will soon bring to scientific area of chloroplast protein translocation.

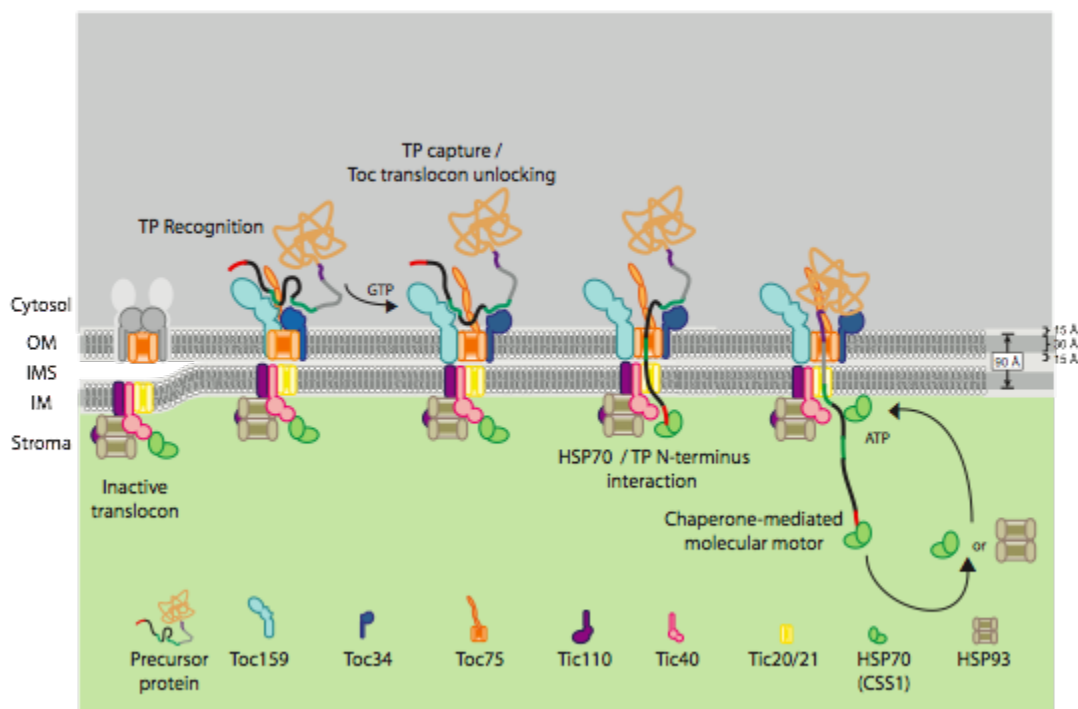


Figure 11. Depiction of the best-supported import mechanism for the translocation of chloroplast-targeted proteins. This model accounts for the specificity in the secondary structure of TP preproteins, as well as the nucleotide factors and other membrane components required to successfully import proteins to the stroma of the chloroplast (Figure from Holbrook, K., et al, 2016).

METHODS

Cloning Strategy

To expand upon the cloning strategy previously described, Figure 12 may be helpful. Preproteins were packaged in concatemerized vectors, cut and ligated into a working vector according to CloneJet protocol available from the manufacturer.

First, each preprotein was cut from the individually, preproteins were lifted from each “Rule” Vector using primers designed to amplify exactly one preprotein from the Rule Vector. Because the quantity of preprotein insert was much higher than the parental Rule vector, the insert was ligated into a CloneJet vector.

In order to get the preproteins into a fluorescent vector, the 42 CloneJet vectors were cut using Restriction Enzymes Nco1 and Nhe1. The parental construct, WT pAN 187-YFP, was cut using the same restriction enzymes. The ends

of the parental vector were then dephosphorylated. Both insert and parental vector were run on a 1 % DNA Agarose Gel. On the DNA agarose gel, the parental CloneJet DNA and the preprotein insert were clearly distinguished in the gel by use of a 100bp DNA Ladder. Preprotein inserts were excised from the gel under UV light and purified using a Qiagen Gel Clean Up Kit. Finally, the purified insert was ligated into the dephosphorylated parental pAN 187-YFP construct to create a chimeric fluorescent preprotein construct.

All CloneJet and ligation products were screened using sequencing techniques at the Molecular Biology Core Facility at the University of Tennessee, Knoxville.

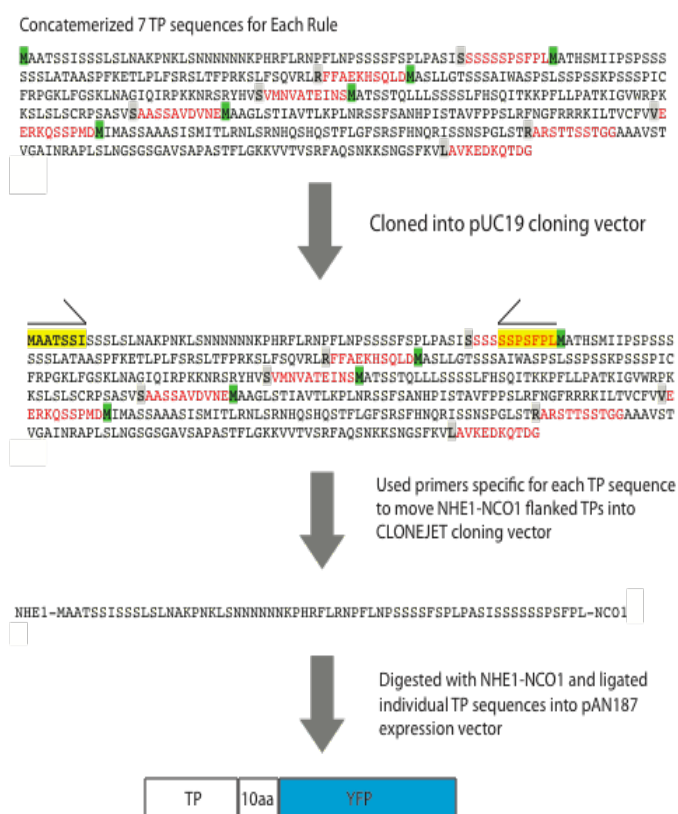


Figure 12: Each transit peptide was ordered in a concatemerized vector per rule. In order to lift each TP from its Rule vector, we designed primers corresponding to each TP within a rule vector. Each rule was stored in a CloneJet vector. Then, TPs were incubated with restriction enzymes prior to their placement in an expression vector to ensure unidirectional import. After the TP for each rule (42 correct CloneJets!!) was selected, we were able to attach it to Tungsten particles and bombard onion epidermal cells. Then microscopy was used to determine the efficiency of import for each TP.

Biolistic Bombardment

Tungsten Preparation

30 mg of Tungsten (M-17 particles) were measured into a 1.5- mL eppendorf tube. Tungsten particles were cleaned by adding 500 uL of (FRESH) 70% Ethanol to the tungsten particles. Ethanol and tungsten were vortexed together for 1 minute. The suspension was allowed to sit for 1 minute on the bench. Then the suspension was spun for ≤ 5 seconds (not exceeding 3000 x g). Particles were forced to the bottom of the eppendorf tube, so that the ethanol could be pipetted out as a supernatant. Next, the pellet was resuspended by adding 500 uL of distilled water to the tungsten pellet. The particles were spun again for < 2 seconds (not exceeding 3000 x g) and the supernatant was removed again. This water wash step was repeated twice. Finally, after the third water wash step, 500 uL of 80% glycerol was added to the suspension. The suspension was vortexed briefly and then stored at 4°C for up to 1 month.

Macrocarrier Disc Preparation:

Tungsten-glycerol stock was vortexed on the highest setting for at least 10 minutes. Sample tubes were set up for each “shot” of DNA. Label tubes. The tungsten-glycerol stock was sonicated in a water bath sonicator for 10 minutes before distributing the tungsten stock into reaction tubes. (You can skip this first sonication step if tungsten was prepared fresh right before making the discs).

Tungsten in 10 uL-aliquots into each labeled eppendorf sample tube, vortexing briefly in between aliquots.

The following reagents were added in rapid succession:

~2.5 ug of DNA

25 uL of MgCl_2

5 uL of 0.2 M spermidine (hygroscopic material; should be prepared fresh every month & aliquoted; Store in -20°C)

Mix samples occasionally by tapping them. Place the tubes on ice for 20 minutes.

Spin samples for 1-2 seconds. Remove supernatant. Maintain pellet for wash steps. Add 100 uL 70% ethanol. Remove supernatant without disturbing the pellet. Then add 100 uL 100% ethanol, & remove it without disturbing the pellet. Repeat the 100% ethanol wash step twice.

Spreading the Macrocarriers:

Vortex each sample tube for 1 minute. Then pipet the DNA-tungsten mixture up and down a few times. Sonicate for 1 second three times in a water bath sonicator. Pipet the mixture a few more times. This ensures more evenly spread tungsten particles across each macrocarrier. Tungsten should be spread over about $\frac{1}{3}$ of the middle of the macrocarrier. Tungsten spreads most evenly when it is “dropped” onto macrocarriers without actually touching the macrocarrier.

Bombardment

Rupture discs of 1100 psi were used; Distance from the macrocarrier to the onion tissue below was kept constant for all experiments. MS agar plates were placed with the “shoot” side of the onion tissue toward the back of the biolistic gene gun and the “root” side of the tissue toward the door of the gene gun. The detached cells were facing up on agar and therefore the cells that were bombarded with DNA-coated tungsten particles.

Onion Preparation

Organic yellow onions were used in this assay. The best onions were medium-sized and round. Cuts from the third layer of the onion toward the “shoot” end of the onion were used in these experiments. These regions of the onion were selected due to the development of an onion as it grows from the inside out and from the roots toward the shoot. The epidermal layer was removed in the hood using gloves. The recently detached cells were placed face up on the MS agar plates. The orientation of onion epidermal layers (“root to shoot”) was tracked and standardized across experiments.

Onions were used for this experiments, because they have non-green leucoplasts, eliminating the additional fluorescence of chlorophyll common to pea plants and other green leucoplast containing species.

Fluorescent Time Scale

Following bombardment, onion samples were imaged at an 18-hour time point, determined to be the most fluorescent time frame by previous assays by Kristen Holbrook in her dissertation completed in 2016.

ImageJ Quantification for Transformation Efficiency Experiments

Transformation Efficiency was assessed using a 2X objective lens in the Krishnan lab. This experiment was among the first (if not the first) conducted using a biological digital imaging system in the system of the Keyence BZ-X710 model of microscope. This system was able to take a series of pictures of the entire tissue sample and “stitch” them together. Transformation efficiency was quantified using ImageJ.

Using ImageJ, crop the image to include the minimum square size of the tissue. Take an average pixel count for 10 individual cells. Divide the entire pixels of an onion sample into the number of pixels for one cell to get a proportion correlating to the transformation efficiency.

Epifluorescent Light Microscopy (Hesler, Room 239)

Using the light microscope, I was able to capture a series of representative and quantification images. Representative images were taken using a 20X objective in order to capture the entire cell in one frame. These images are a good visual representation to help understand and precede the quantification data.

We used peroxisomal-localized plasmids labeled with CFP, so that we could confirm transformation of each cell we used for quantification. Because the chloroplast-localized proteins were labeled with YFP, they were visible in another channel.

REFERENCES

- Bionda, T., Tillmann, B., Simm, S., Beilstein, K., Ruprecht, M., and Schleiff, E. (2010). "Chloroplast import signals: the length requirement for translocation in vitro and in vivo." *J. Mol. Biol.* 402. Pp. 510 - 523.
- Chotewutmontri, P., and Bruce, B.D. (2015). "Non-native, N-terminal Hsp70 molecular motor recognition elements in transit peptides support plastid protein translocation." *Journal of Biological Chemistry*. 260: 7602 - 7621.
- Chotewutmontri, P., Reddick, L.E., McWilliams, D.R., Campbell, I.M., and Bruce, B.D. (2012). "Differential transit peptide recognition during preprotein binding and translocation into flowering plant plastids." *Plant Cell*. 24: 3040 - 3059
- Holbrook, K., et al. (2016). "Functional Analysis of Semi-conserved Transit Peptide Motifs and Mechanistic Implications in Precursor Targeting and Recognition." *Molecular Plant* 9(9): 1286-1301.
- Nelson, B.K., Cai, X., and Nebefuhr, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J.* 51, pp. 1126 - 1136.
- Pilon, M., et. al. (1995). "Functional Domains of the Ferredoxin Transit Sequence Involved in Chloroplast Import." *The Journal of Biological Chemistry* 270(8): 3882-3893.
- Schleiff, E., Becker, T. (2010). "Common ground for protein translocation: access control for mitochondria and chloroplasts." *Nature Reviews Molecular Cell Biology* Vol 12, pp. 48-59
- Von Heijne, G., Steppuhn, J., and Herrmann, R.G. (1989). "Domain structure of mitochondrial and chloroplast targeting peptides. *Eur. J. Biochem.* Vol 180, pp. 535 - 545.