

Localization and Structure of Plastidial-Encoded Polymerase

Sub-units

**Andres Eduardo Guerrero Criado^{ab}, Monique Liebers^a, David Cobessi^c,
Noha Youssef^b, Thomas Pfannschmidt^a, Robert Blanvillain^a**

**^a, Laboratoire de Physiologie Cellulaire et Végétale. ^b, OSU. ^c, Institut de
Biologie Structurale.**

Laboratoire de Physiologie Cellulaire Végétale

17 rue des Martyrs 38054 GRENOBLE Cedex 9

Chloroplast biogenesis is a quintessential process in plants for the acquisition of photo-autotrophy. The path towards chloroplast biogenesis begins with the inhibition of skotomorphogenesis (That is the growth program in the dark), followed by photomorphogenesis through phytochrome-mediated red light sensing leading to the initiation of chloroplast biogenesis. We have observed that plants do not develop properly and show an albino phenotype when PAPs are lacking, indicating that chloroplast biogenesis does not occur. Previous experiments have shown that a complex dubbed Plastid Encoded Polymerase (PEP) interacting with Polymerase Associated Proteins (PAP) is the major player in setting a functional chloroplast. Given the semi-autonomous nature of the chloroplast genome, it is established that a cross-talk must occur with the nuclear genome. The white (albino) phenotype is observed with the genetic excision of any single PAP, indicating that the whole complex is no longer able to form or function. Interestingly

both a chloroplastic-transit peptide and a nuclear localization signal have been predicted within the sequences of six PAP proteins. To validate these predictions and gain insight on the role of selected PAP proteins, I have chosen to pursue the subcellular localization and protein-protein interactions of these essential proteins. Moreover structural characterization will help us understand the function of each of these proteins within the PEP/PAP complex.

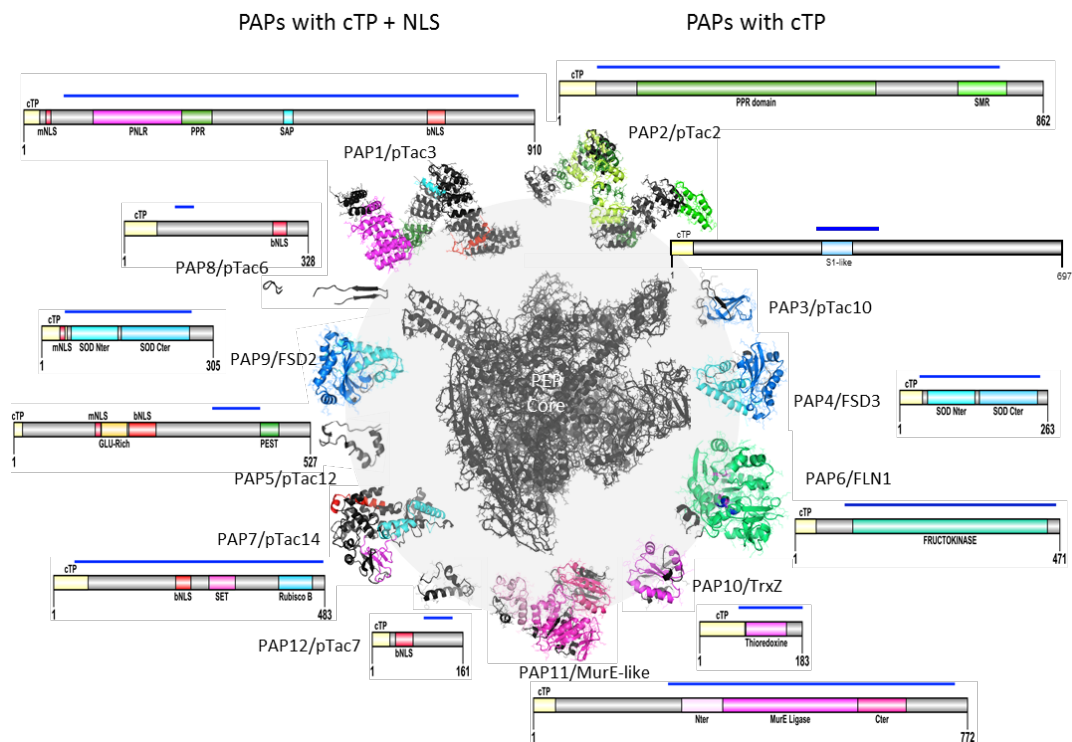
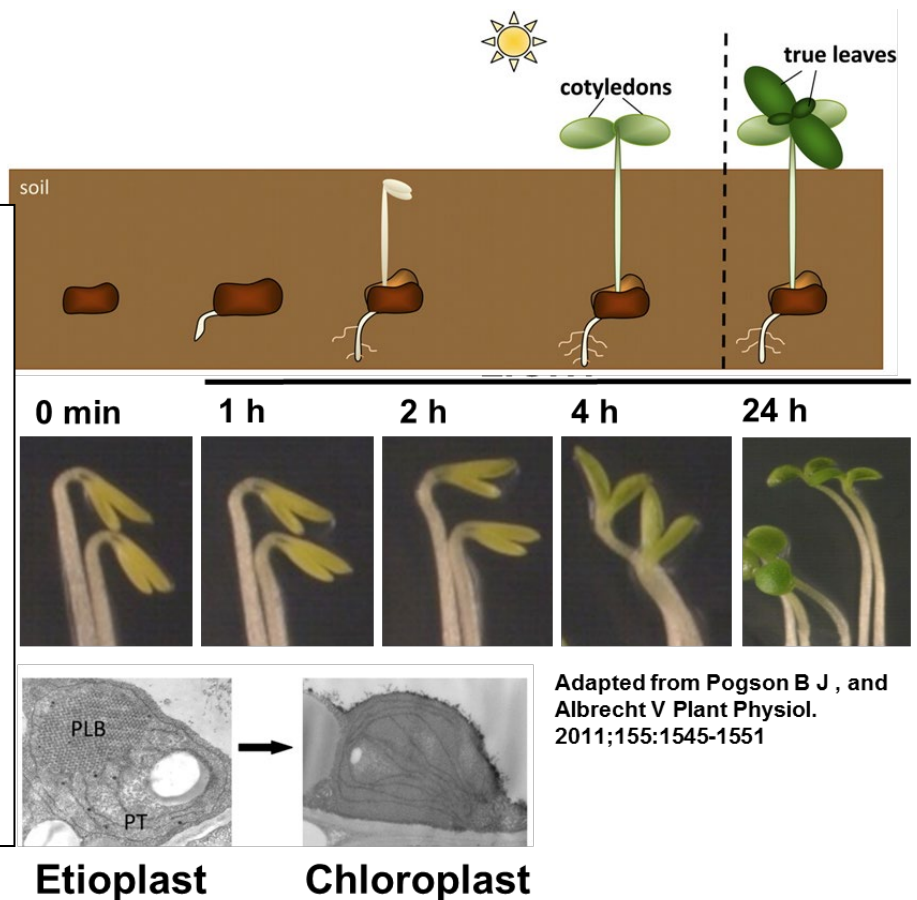


Figure 1. Protein predicted models of the PEP/PAP complex (Pfannschmidt et al., 2015). Note the clear division between proteins that include both a Nuclear Localization Signal (NLS) and a Chloroplast Transit Peptide (CTP). The colored portions of the portrayed proteins depict similarities with known domains in databases.

Introduction

As is common within plants, the initiation of germination could occur in darkness, when seeds are buried in soil. In this case, seedlings initiate a program of growth called skotomorphogenesis. Our model organism *Arabidopsis thaliana* springs forth a hypocotyl which elongates and sprouts its cotyledons. These cotyledons house the precursors to the chloroplast in the forms of proplastids and etioplasts. Once light exposure occurs, photomorphogenesis is launched, transforming these precursor organelles into fully functional photosynthetic chloroplasts.



Due to the nature of the chloroplast with its own semiautonomous genome, most of its constituents (approximately 95 percent of chloroplast proteins) are encoded by the nuclear genome. Therefore we can infer that the chloroplast must communicate its needs and status to the nucleus in order to coordinate its development to the one of the cell. The chloroplastic transcriptional machinery is composed of two polymerases: the nuclear encoded phage type polymerase (NEP) and the Plastid encoded bacterial polymerase (PEP). The PEP is associated with several nuclear encoded Associated Proteins (PAPs) forming a large complex that is supposedly generated within the plastid after light perception NEP and is additionally (PAPs). These PAP proteins develop in a light dependent way and when imported into the chloroplast with their pertinent sigma factors lead to the development of a functional chloroplast.

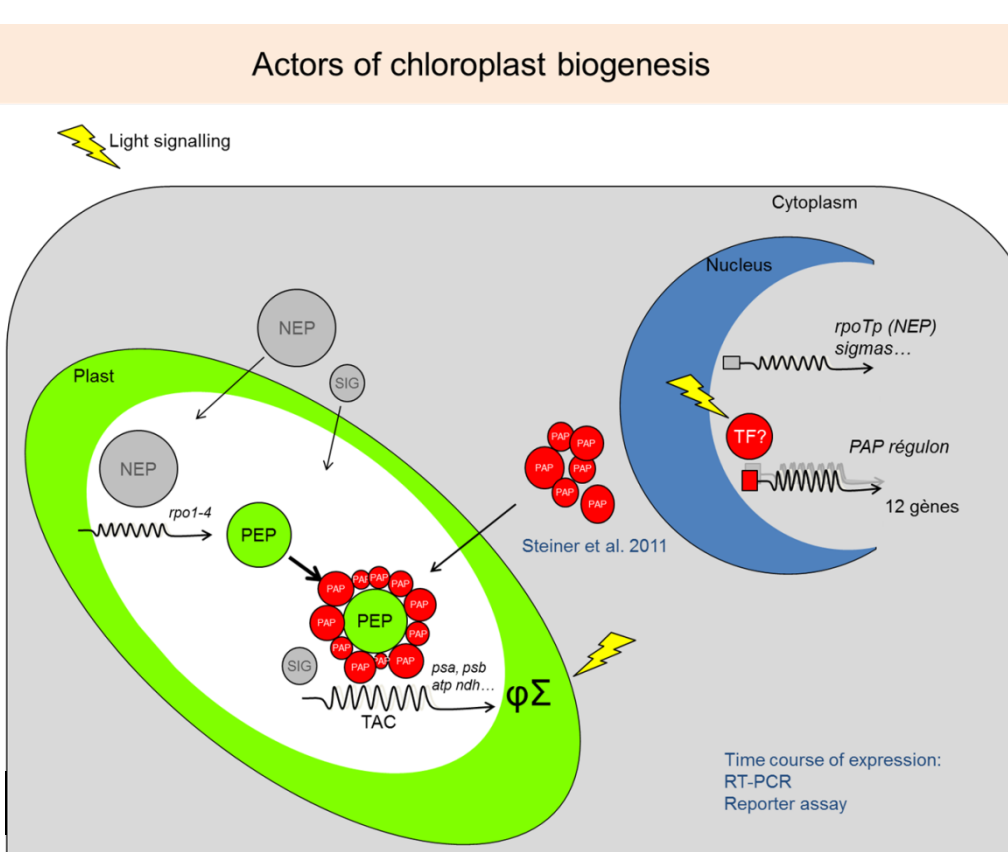


Figure 3. Demonstrates the formation and activation of the PEP/PAP complex. Note the hypothesized need of light for nuclear transcription factor activation. This gives way to the transcription and translation of the vital PEP associated proteins. Following PEP synthesis, PAPs travel to the chloroplast, assemble with the quaternary PEP core to initiate transcription of photosynthetic genes within the Transcriptionally Active Chromosome (TAC).

Isolation of PEP-associated proteins (PAPs) :

Using chloroplasts proteomic analysis

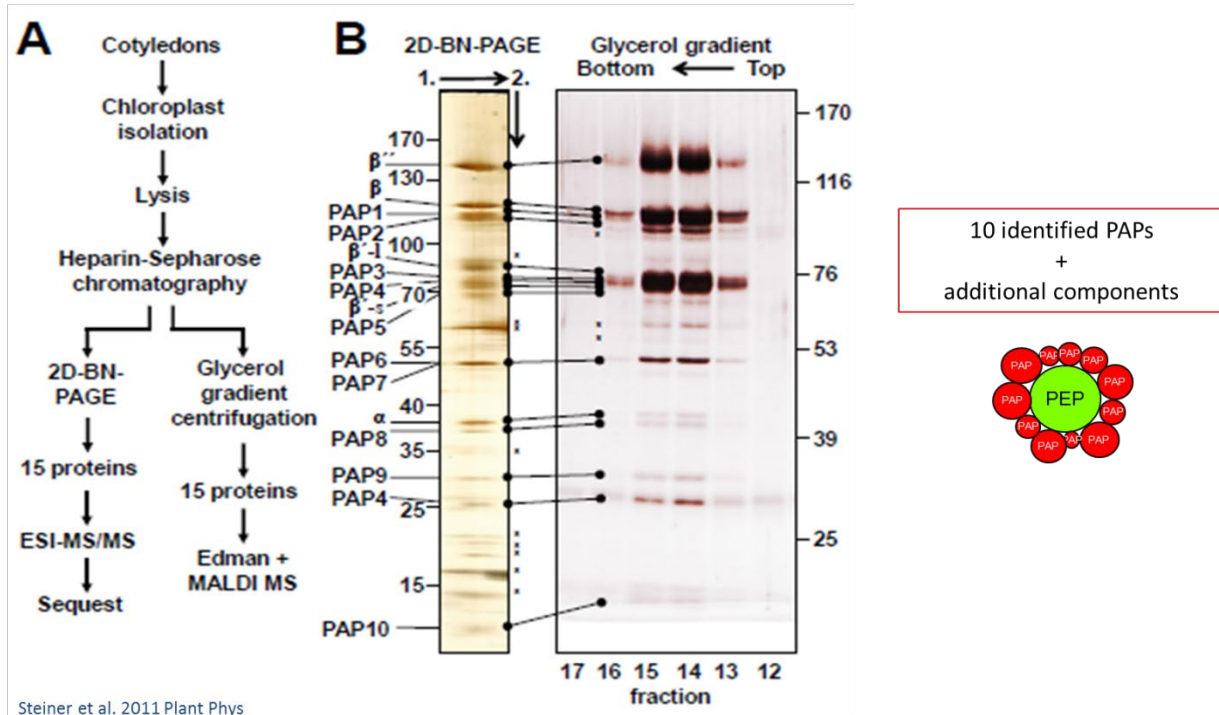


Figure 4. Through the works of Steiner *et al* it has been shown that there are approximately 10-12 PAP proteins.

What is intriguing about this PEP/PAP complex is that the removal of any single component PAP deactivates the complex as a whole and leads the plant to an early death following germination. This is demonstrated at a phenotypic level by a yellow or albino plant that is stunted in growth. We also observe that this falls in line with the typical Mendelian 3-1 segregation.

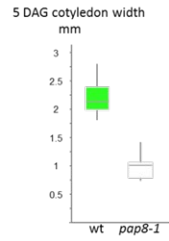
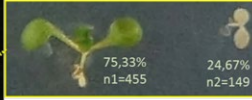
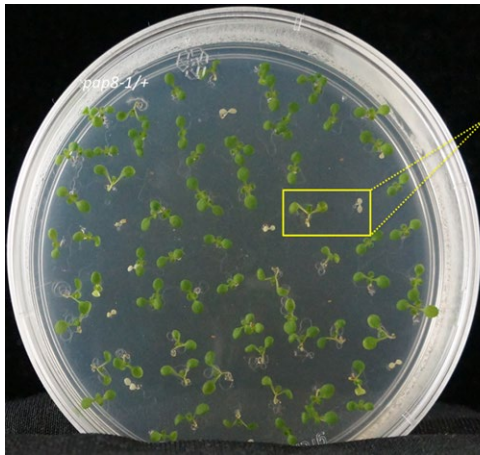


Figure 5. *PAP8-1/+* heterozygous progeny segregates a recessive mutant allele of *PAP8* in a 3-1 ratio. Therefore a bias to Mendel's law has not been detected indicating no defect for gametes. *pap8-1* homozygous plants display smaller albino cotyledons.

METHODS

Harvesting

Harvesting was completed by the drying of our Arabidopsis plants for approximately three days. Roots are removed in addition to the removal of bottom leaves and the siliques are pinched in order to collect the seeds of our mutated or wild type plants.

Sterilization of seeds

To properly avoid any viral or bacterial infection of our mutated seeds sterilization was completed. Making sure to avoid cross contamination, the seed packets collected from previous harvest are soaked in a stock solution of 15ml of bleach, 35ml of Step 2 purified H₂O and 3 drops of (20X) Triton detergent and rested for five minutes. Following this rest period the seeds were centrifuged at 4000G for 2 min. This detergent solution is then washed out with an equal amount of H₂O with the supernatant removed by pipetting. Following this wash, another agitated wash with an equal volume of water is added and subsequently removed. Once again a final wash is completed after pipette removal of wash water followed by light centrifugation to remove excess water.

Medium Preparation

	<u>1L</u>	<u>500ml</u>
<u>MES</u>	.5g	.25g
<u>MS Salts</u>	2.15g	1.075
<u>Vitamins</u>	500ml	250ml
<u>1% Sucrose</u>	10g	5g

(Top Agar)

<u>.5% Agar</u>	8g	4g
------------------------	----	----

Once the mixture has been created the agar plates were poured by utilizing sterilized plates, heating the agar for five minutes, letting it cool and allowing the plates to solidify for approximately 15 minutes.

RNA Quantification

RNA was quantified by Nanodrop through traditional methods

Polymerase Chain Reaction

Addition of 0.5ul of RNA cDNA and blank samples

Addition of 2ul of gDNA sample

Addition of 1ul of Forward Primer

Addition of 1ul of Reverse Primer

Addition of 20ul of {taq polymerase, buffer and dNTPs} (all inclusive)

Addition of 16ul of H₂O to final volume of 40ul

Programming of Thermocycler

The cycling of the thermocycler was controlled to 96 degrees for 2 min for enzyme activation. Denaturation was carried out at 96 degrees for 20 seconds, annealing was set for 52 degrees for 20 seconds, synthesis was set at 72 degrees for 1 min and 15 secs with a final synthesis of 72 degrees for 3 min and 12 degrees for maintenance after 34 cycles completion.

Primer/ Plasmid genesis

For the bioinformatical genesis of experimental plasmids, the application APE was utilized. In order to be able to properly view what a hypothetical plasmid would look like and to be able to have an experimental preview to the weight and genotype of the proper ligated and reformed plasmid. To place our sequence within the plasmid of E.coli, primers were created from around 30bps from beginning of endonuclease sequence into the ORF for the forward primer and for the reverse primer, around 30 bps backwards from end of endonuclease sequence into the ORF finishing within the ORF with a sequence reversal being applied in order to access it later on the opposite chain of DNA.

RNA Extraction

Refer to Qiagen RNA extraction protocol

Polyacrylamide Gel Electrophoresis

For the confirmation gels and extraction gels, 1g of Agarose per 100ml of TAE Buffer was heated and mixed with 10ul of Cybersafe DNA stain, poured into a multiwelled mold and solidified for 5 min. Following the hardening of the gel, the gel was run at 135 volts.

DNA Ligation

Completion of DNA ligation into cloning vectors and protein production vectors was completed utilizing consistent methods with the following mix proportions: .5ul of ligation buffer (2x), 1ul of ligase, 1ul of vector, 3ul insert, Let sit at room temperature for 2 hours (alternatively place in water and cold room for weekend).

Ecoli Transformation

For protein production, sequence analysis and manipulation of our PAP proteins, transformation into E.coli was completed. The addition of 50ul of freshly thawed E. coli, 5ul of ligated vector, .5ul of positive control. Once this mixture was completed it was placed for 5min on ice, 40 seconds in a 42 degree water bath and immediately placed on ice for exactly two minutes.

Restriction Digest

For a 10ul sample, .5ul of forward cleaving enzyme,.5ul of reverse cleaving enzyme, 1ul of buffer (10x), and DNA amounting to at least (800ng) based on nanodrop calculation, with a completion to 10ul with H2O once completed this master mix, run in 37 degree water bath for at least 2 hours

Plasmid preparation and cleanup

Following confirmation and excision of genetic sequences, the GeneClean Kit was utilized to purify and extract our DNA from our analytical/ extraction gels.

Gold Particle Gun Bombardment

Gold particle gun bombardment of genetic material was carried out upon onion cells using the methods described within BioRad Laboratories instructional pamphlet.

RESULTS

Our approach towards a further comprehension of the function and localization of our proteins of interest was two-pronged. We wished to further understand the presence of a possible feedback mechanism within the chloroplast and nucleus of our proteins with a predicted nuclear localization signal (PAPs) and chloroplast transit peptides. To do so, a reverse transcriptase PCR was performed with primers that would include the entire open reading frame of our PAPs (7-8) on Arabidopsis RNA samples. The amplified fragments were then sub-cloned into vectors allowing tagging with GFP or Split YFP to form the constructs in figures 6, 7, 8 & 9.

Subcellular Localization

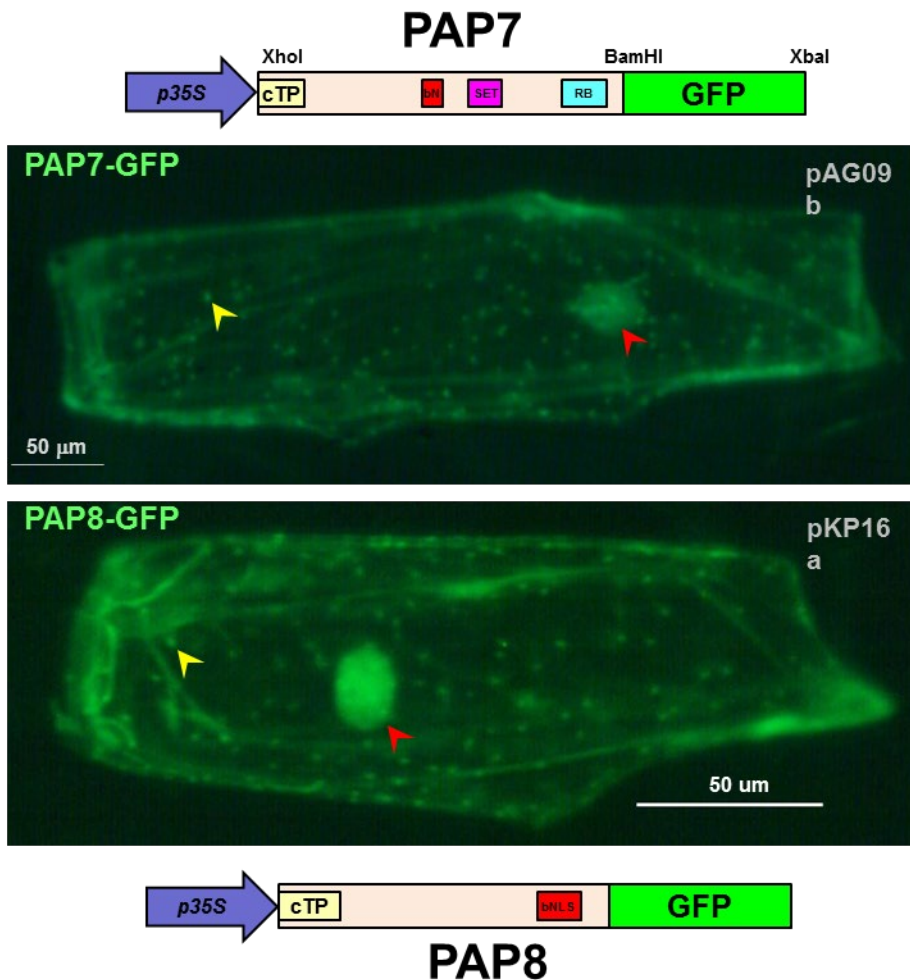


Figure 6. Demonstrates the construct designed in order to be able to observe the subcellular localization of PAP7 and PAP8, a ubiquitous 35s promotor followed by our PAP7orf/ PAP8orf complete with chloroplast transit peptide encoding sequence and Nuclear localization signal fused to GFP for fluorescence detection. The application of our plasmid with the aforementioned construct through particle gun bombardment of onion cells. It is observable that PAP7 and PAP8 seems to reach the nucleus due to the strong signal we perceive in addition to appearing in the proplastids.

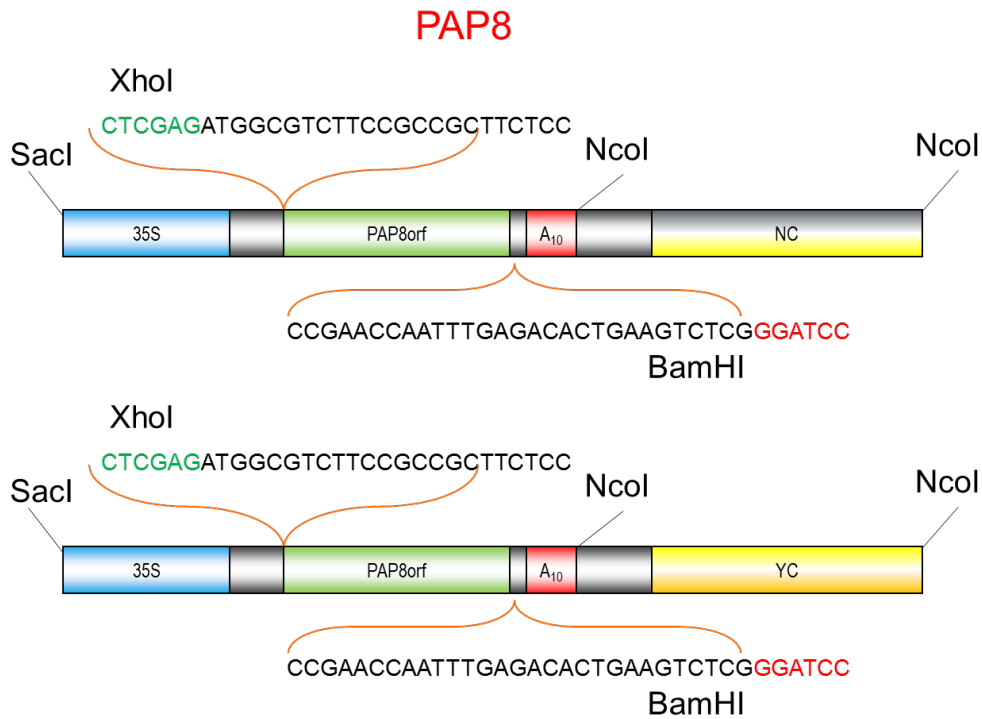


Figure 7. In order to test the possibilities of a homodimer being formed with PAP8, the preceding construct was designed and isolated through the use of split YFP tagging the N and C termini of our protein. This construct can also be applied for further testing of possible heterodimers among PAP7 and other PAP proteins.

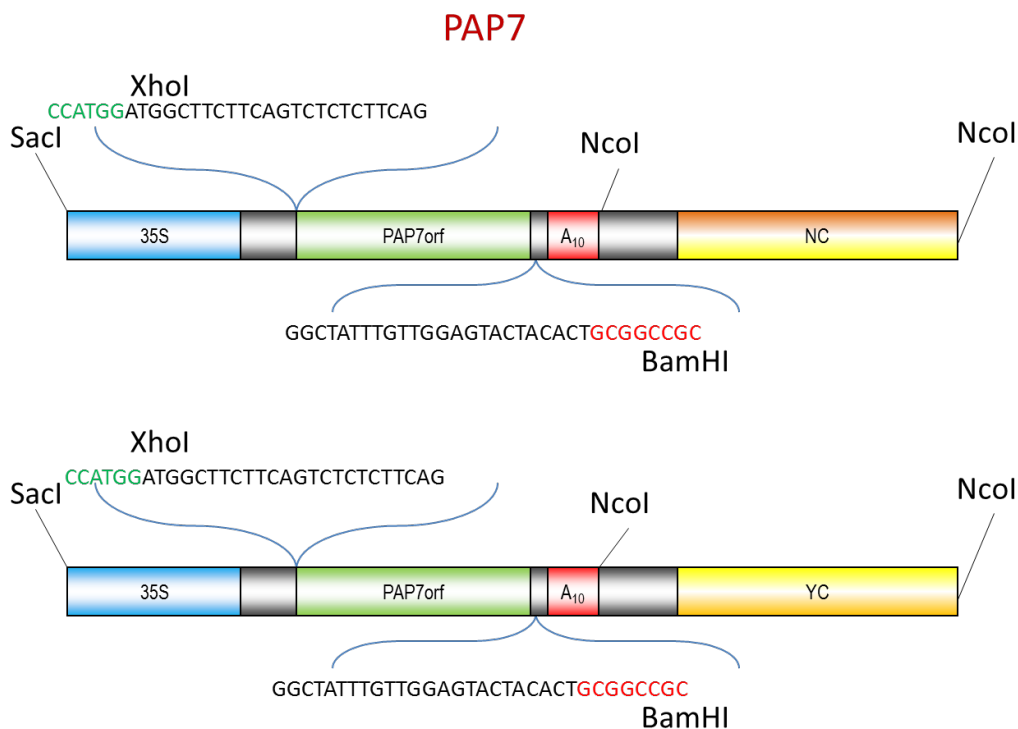


Figure 8. In order to test the possibilities of a homodimer being formed with PAP7, the preceding construct was designed and isolated through the use of split YFP tagging our protein. This construct can also be applied for further testing of possible heterodimers among PAP7 and other PAP proteins.

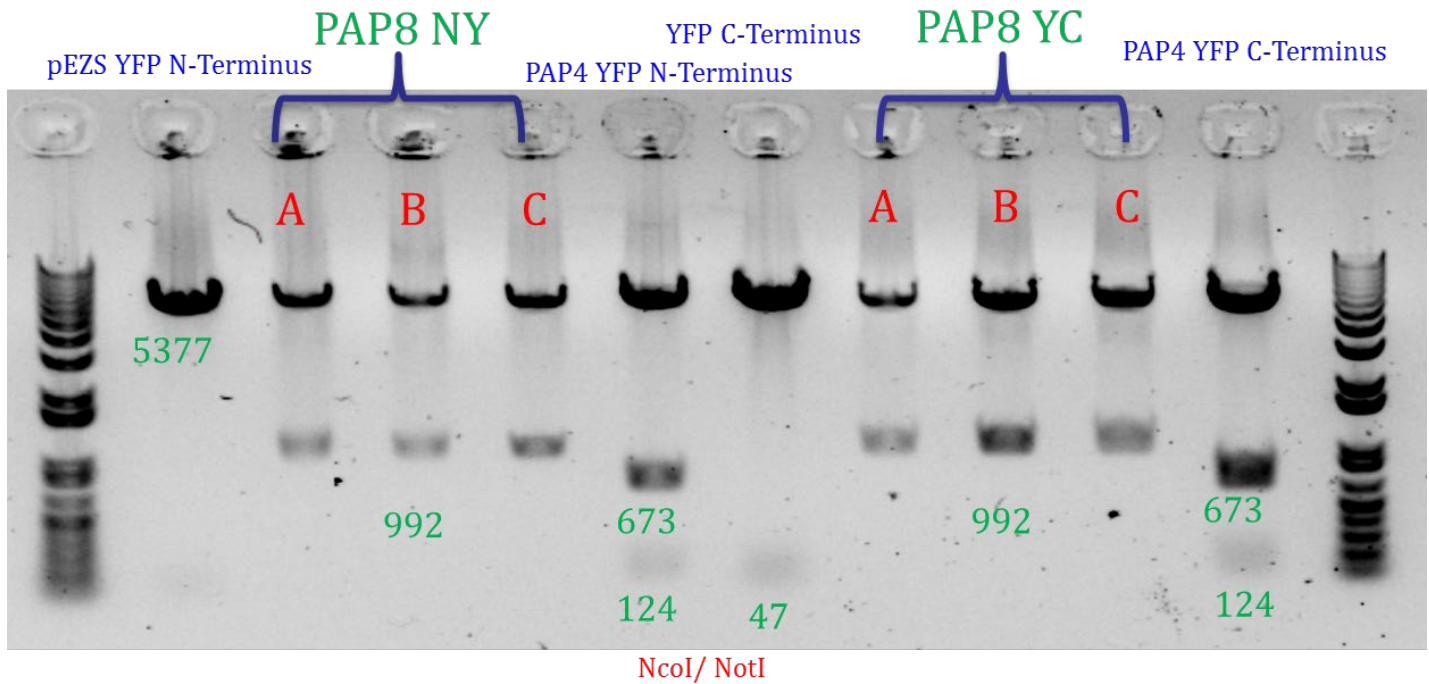
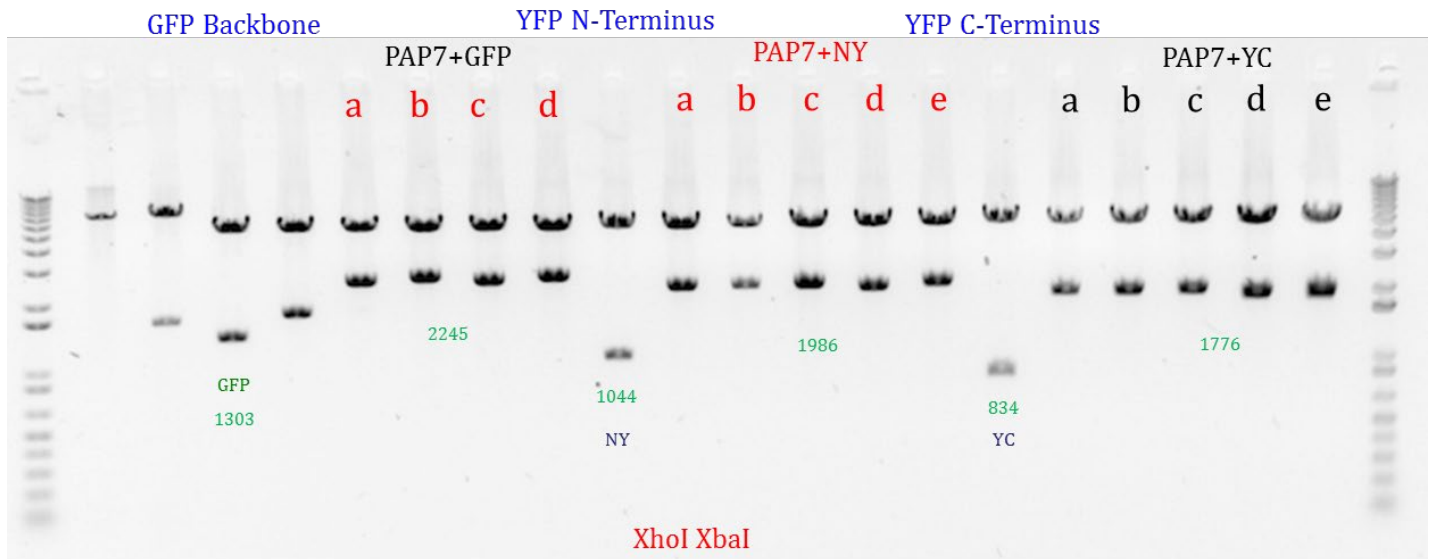


Figure 9. The constructs presented in figures 6,7 and 8 were cloned into a DH5 strain of E. coli and are shown above with their corresponding base pair sizes. Following this gel confirmation further assurance of the proper ligations and cloning's of these new GFP and YFP tagged sequences was done through sequencing.

Structural Characterization

In our attempt to further understand the function of our proteins, it was first necessary to clone ORF in vectors for protein production in *E. coli* and purification using affinity column. We know from previous experimentation and analytical predictions that our PAPs are encoded by the nucleus and transported to the chloroplast as a finished product. Based on this knowledge of the pathway the PAPs take to reach the chloroplast, we developed a strategy to remove the predicted chloroplast transit peptide from our protein sequences. Therefore through primer design and recognition of the sequence we clipped the CTP off of the PAPs in question using NcoI and NotI restriction enzyme digests. These novel recombinant sequences were then inserted into an initial pGEMT vector, transformed into *E.coli* and run through a restriction digest as shown in figure 14.

Once sequence confirmation was obtained we followed our strategy by transferring our inserts of cTp-less sequences to a vector suitable for protein production. This vector is called pETM41. In figure 11 you can observe the pertinent characteristics that make pETM41 an optimal protein production vector such as its Maltose Binding Protein fusion contiguous to the protein of interest (shown here for PAP12). A TEV (tobacco etch virus) proteolytic site used for selective cleavage, both kanamycin and tetracycline resistance, a histidine tag for post-production nickel column affinity and a T7 promoter and terminator site for expression.

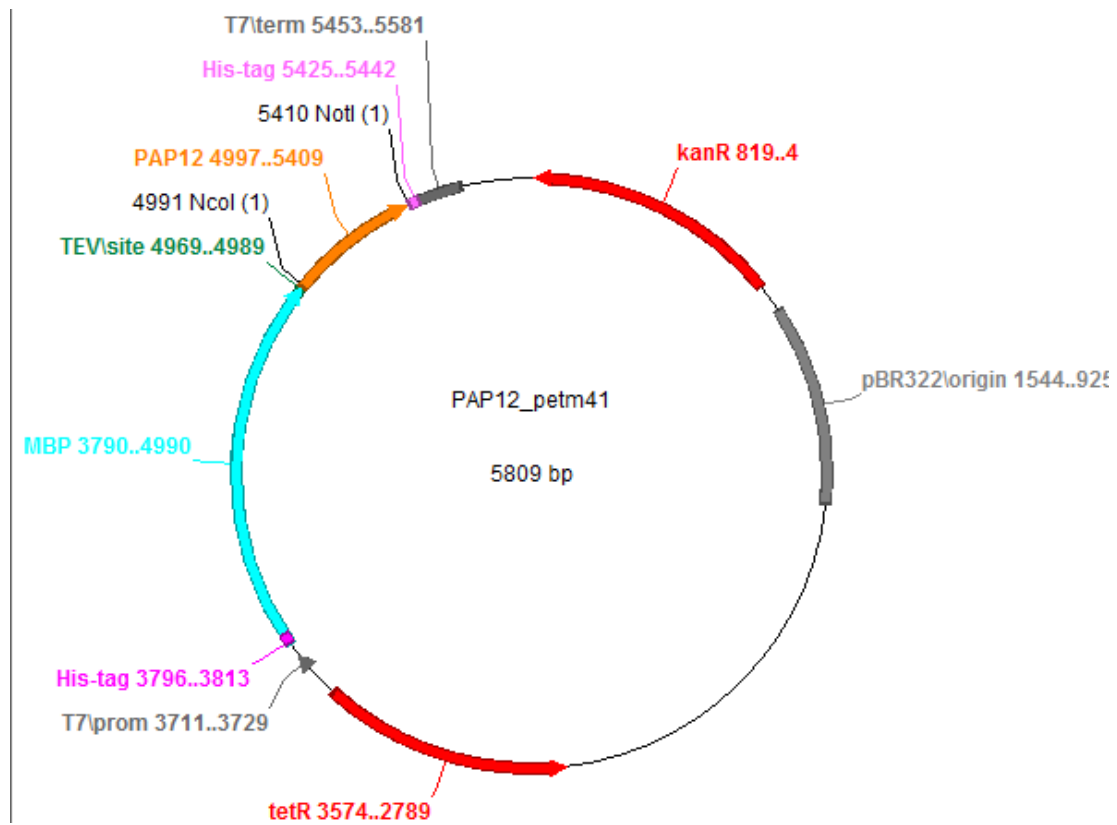


Figure 10. pETM41 vector with PAP 12 insert designed specifically for optimal protein production and further purification techniques and facilitation.

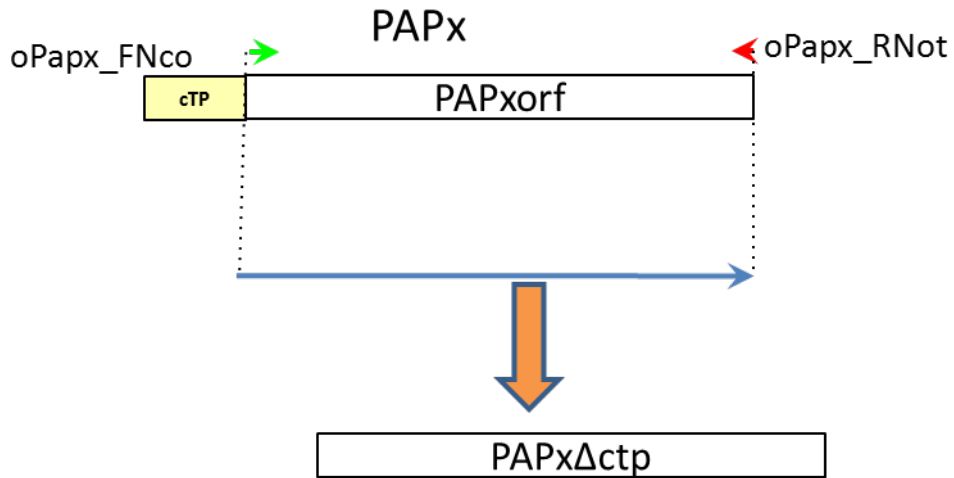


Figure 11. Illustrates the cloning strategy of our PAPx sequences from the *Arabidopsis thaliana* genome using PCR and removal of the unwanted chloroplast transit peptide.

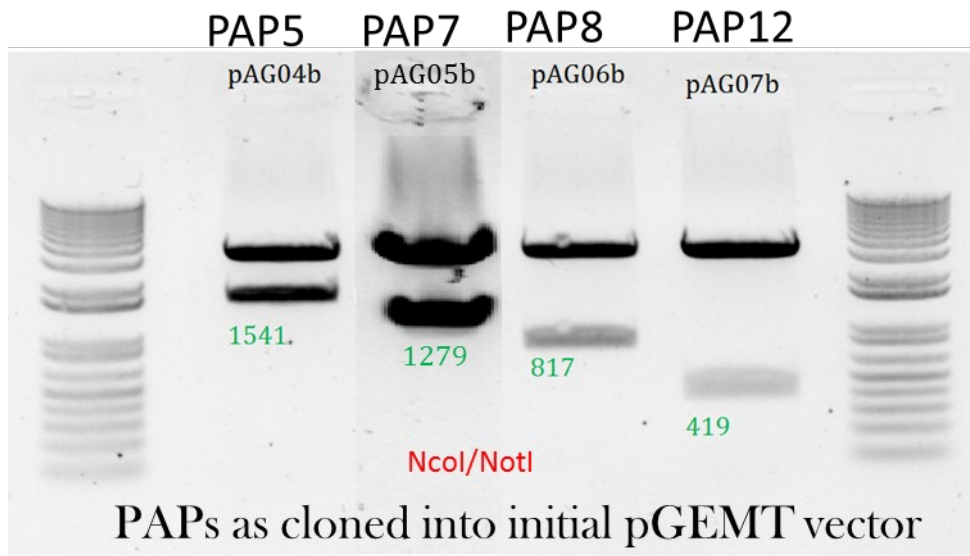


Figure 12. PAPs \square cTP were cloned into pGEMT and digested with NcoI NotI to ensure proper ligation into the expression vector.

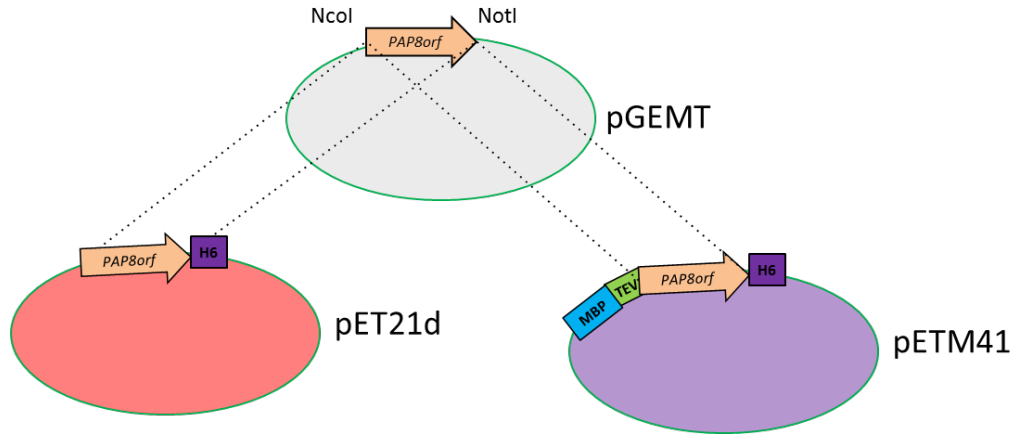
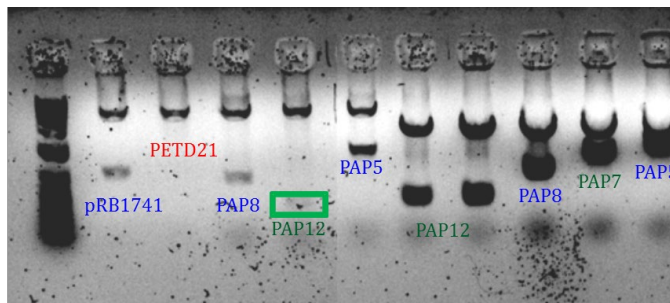
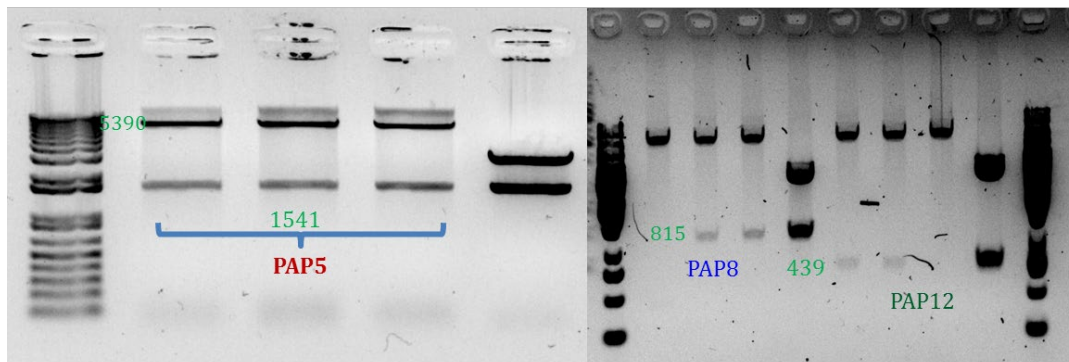


Figure 13. Alternative strategies utilized with PAP8 towards protein production vectors including (pETM41) or not (pET21d) the MBP tag, H6 tag used for Nickel column affinity purification.



Insertion into pETM41



NcoI/ NotI

Figure 14. Shows the proper insertions and isolations of PAPs into vectors pET21d and pETM41.

Discussion

It is thought that the formation of a fully functional PEP complex can only be produced following the activation of light dependent pathways. These pathways in turn would lead to the expression of the PAPs and the generation of a transcriptionally active PEP/PAP complex. Therefore this process of light sensing and activation must precede the activation of any photosynthetic chloroplastic genes. Investigations currently taking place within our lab group have been able to demonstrate that genes encoding for the PEP-PAP complex are already expressed before light exposure (Monique Liebers). This discovery reveals that mechanisms for chloroplast biogenesis may involve post-transcriptional regulations of the PAPs.

Another interesting characteristic of our PAP proteins is the plethora of functions that they seem to cover. Our harnessing of bioinformatics data and tools has permitted us to develop a model of predicted domains within several of our PAPs (Dr. Robert Blanvillain). Particularly of interest are the PAP12 and PAP8 proteins for which the only domains that can be recognized are the chloroplast transit peptide and the nuclear localization signal, for this study we focused heavily on PAP7 and PAP8 for both structural elucidation and subcellular localization. From virtual simulations and sequence analysis we can predict that PAP 7 consists of a SET domain. "SET domain protein methyltransferases catalyze the transfer of methyl groups from the cofactor *S*-adenosylmethionine (AdoMet) to specific lysine residues

of protein substrates, such as the N-terminal tails of histones H3 and H4 and the large subunit of the Rubisco holoenzyme complex.” (Trievel et al.,) .

These proteins can be divided into four classes as typified by their *Drosophila* members E(Z), TRX, ASH1 and SU(VAR)3-9 being characterized in yeast and mammals, except until very recently where it was hypothesized to occur both with green peas and our model organism *Arabidopsis thaliana* (Baumbousch et al., 2001). PAP10/TRXz for instance seems to be involved in dark/light transitions through its thioredoxin activity. Specifically the WCGPC motif predicted within the thioredoxin of PAP10 is a landmark feature of thioredoxin proteins. The cysteine residues of this motif are found predominantly reduced in vivo, allowing thioredoxins to break disulfide bonds in oxidized substrate proteins (Jakob and Reichmann, 2013). This would lead us to confirm findings that its deletion has a significant effect on some nuclear encoded genes (Arsova et al., 2010). HEMERA/PAP5 plays a role in phytochrome signaling in the nucleus and is similar to multi ubiquitin binding protein RAD23 and is also essential for chloroplast development (Chen et al., 2010). We observe through deletion techniques that the elimination of any single PAP protein renders the complex futile and leads to a significant albino/ivory phenotype.

Unfortunately we are not provided foresight on the characteristics of PAP8 or PAP12 from the previously mentioned techniques, and it is for this reason that we hope to produce the protein through *E. coli* within our pETM41 vector. Through our cloning techniques we have in fact been able to produce

Rosetta strains of *E. coli* engineered for protein production with our novel pETM41 plasmids including our protein sequences for PAP5, PAP7, PAP8 and PAP12. These will be tested for optimal protein production conditions, purification conditions and crystallization conditions and X-ray crystallography characterization through our collaboration with Dr. Cobessi at Institut de Biologie Structurale, Grenoble.

Although we do not yet have confirmation of the dual localization of PAP12 and PAP5 (as opposed to our findings with respect to PAP7 and PAP8) the expected cTP and NLS propose similar pathways of action. This will be tested through the same Particle gun bombardment of onion cells that was utilized for PAP7 and PAP8. Using fluorescence microscopy and transient expression in onion cells our results suggest that the NLS is a true one within both PAP7 and PAP8. The obtained results in onion cells will be confirmed in *Arabidopsis thaliana* with their respective promoters after functional complementation using *Agrobacterium tumefaciens* transformation.

Together, these results suggest that the localization of PAP8 is controlled within the cell and the protein is most likely transported into the plastid before its retrograde transport to the nucleus in response to the beginning of photosynthesis. To test this hypothesis, a strategy has been developed in the lab to specifically mark proteins in the chloroplast and test if marked protein could be retrieved in the nucleus indicating protein

movements providing with potential feedback messenger. These results could lead in association with structural studies of PAPs to a better understanding of the PEP complex formation and the activation and of an important and probably complex retrograde communication way from the chloroplast to the nucleus.

References

1. T Pfannschmidt, R Blanvillain, L Merendino, F Courtois, F Chevalier, M Liebers, B Grübler, E Hommel, and S Lerbs-Mache Plastid RNA polymerases: orchestration of enzymes with different evolutionary origins controls chloroplast biogenesis during the plant life cycle *J. Exp. Bot.* first published online September 9, 2015
doi:10.1093/jxb/erv415
2. Trievel RC, Rojas JR, Sterner DE, Venkataramani RN, Wang L, Zhou J, Allis CD, Berger SL, Marmorstein R. 1999. Crystal structure and mechanism of histone acetylation of the yeast GCN5 transcriptional coactivator. *Proc Natl Acad Sci U S A* 96:8931–8936.
10.1073/pnas.96.16.8931.
3. Baumbusch L. O. et al. The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. *Nucleic Acids Res* 29, 4319–4333 (2001).
4. Reichmann D, Jakob U. The roles of conditional disorder in redox proteins. *Curr Opin Struct Biol.* Elsevier Ltd; 2013;23: 436–42. doi: 10.1016/j.sbi.2013.02.006 (Arsova et al., 2010)
5. Chen Z, Zhu X. & Ma H. Phylogenetics and evolution of Su(var)3-9 SET genes in land plants: rapid diversification in structure and function. *BMC Evol Biol* 11, 63 (2011).