Chapter 1. Introduction

1.1 General Introduction

Magnesium chelatase (Mg-chelatase) is a tetrapyrrole biosynthetic pathway enzyme located at the branch point between heme and chlorophyll biosynthesis (Reid and Hunter 2002). It catalyzes the first committed step in chlorophyll biosynthesis, producing Mg-protoporphyrin IX (Mg-Proto) by insertion of a divalent magnesium ion into the substrate protoporphyrin IX (Proto)(fig. 1). Alternatively, Proto IX can be used as a substrate by the enzyme ferrochelatase, which inserts a ferrous ion in the formation of heme b (Ferreira et al. 1995).

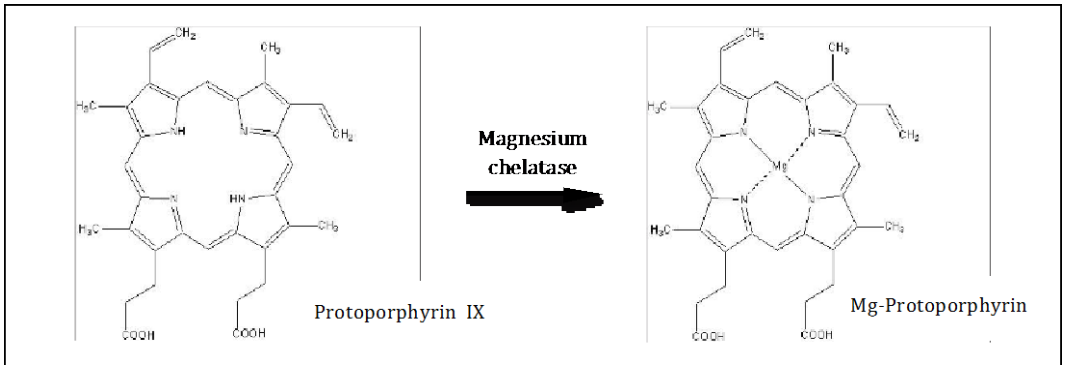


Figure 1. Magnesium Chelatase enzymatic activity

Mg-chelatase functions to insert a divalent magnesium ion into the substrate protoporphyrin IX, creating Mg-protoporphyrin IX

This branch point at which heme and chlorophyll biosynthesis diverge (fig. 2) is of particular interest, in part because products heme and chlorophyll are both vital plant compounds, but also because regulation at this point seems likely (Willows 2003). Furthermore, there are indications that the key to cross talk between the chloroplast and the nucleus resides in some aspect of tetrapyrrole biosynthesis,

and this enzyme has been implicated as having a role in the process. Lastly, it has been reported that one of the subunits of this enzyme is a receptor of the plant hormone abcisic acid (ABA) (Shen et al. 2006b). The link between the role of this subunit as an ABA receptor and chlorophyll biosynthesis, if in fact one exists, has not been elucidated. Thus, greater understanding of the components of this system, and in particular this key enzyme, may help illuminate processes involved in intra-cell signaling.

1.2 The Tetrapyrrole Biosynthetic Pathway

1.2.1. The Enzymatic Steps of Tetrapyrrole Production

In higher plants the first nine steps of heme and chlorophyll synthesis occur in a shared pathway (fig. 2). The first step, catalyzed by glutamyl tRNA synthetase, activates glutamate by ligating it to its cognate tRNA to form $tRNA^{Glu}$ (Schon et al. 1986). Glutamyl tRNA reductase (GluTR) reduces glutamyl-tRNA, producing glutamate-1-semialdehyde (GSA) (Wang et al. 1984), which is in turn converted into 5-aminolevulinic acid (ALA) by glutamate 1-semialdehyde aminotransferase (GSA-AT) (Hoober et al. 1988) in the second and third steps of the pathway.

The fourth enzyme, ALA dehydratase, converts two molecules of ALA to porphobilinogen, a pyrrole molecule (Jaffe 1993). Subsequently, four porphobilinogen molecules are joined to form a linear tetrapyrrole, 1-hydroxymethylbilane. The fifth enzyme, uroporphyrinogen III synthase, converts hydroxymethylbilane to uroporphyrinogen III, the first closed tetrapyrrole in the pathway (Beale 1999).

In the first step of the siroheme branch of the pathway, S-adenosyl-L-methionine-dependent methyltransferase methylates uroporphyrinogen III to form dihydrosirohydrochlorin. Subsequently, siroheme ferrochelatase converts dihydrosirohydrochlorin to siroheme (Tanaka and Tanaka 2007).

Alternatively, uroporphyrinogen III decarboxylase converts uroporphyrinogen III to coproporphyrinogen III, which is in turn decarboxylated to form protoporphyrinogen IX by the eighth enzyme, coproporphyrinogen III oxidase (Beale 1999). The ninth step in the pathway, accomplished by protoporphyrinogen IX oxidase, generates Proto IX, the substrate common to both the heme and chlorophyll branches (Lermontova et al. 1997).

The tetrapyrrole biosynthesis branch point is a significant position in the pathway, as allocation of the Proto IX substrate into heme or chlorophyll biosynthesis is determined here (Tanaka and Tanaka 2007). In addition, from uroporphyrinogen III on, pathway intermediates can be oxidized by molecular oxygen to become the corresponding porphyrin. Porphyrins, in turn, are capable of absorbing light energy, forming highly reactive singlet oxygen (Tanaka and Tanaka 2007), and damaging the plastid membranes in a process called photo-oxidative damage (Mock et al. 1998). Thus, flux through this pathway, specifically at the divergence of the chlorophyll and heme branches, must be carefully controlled.

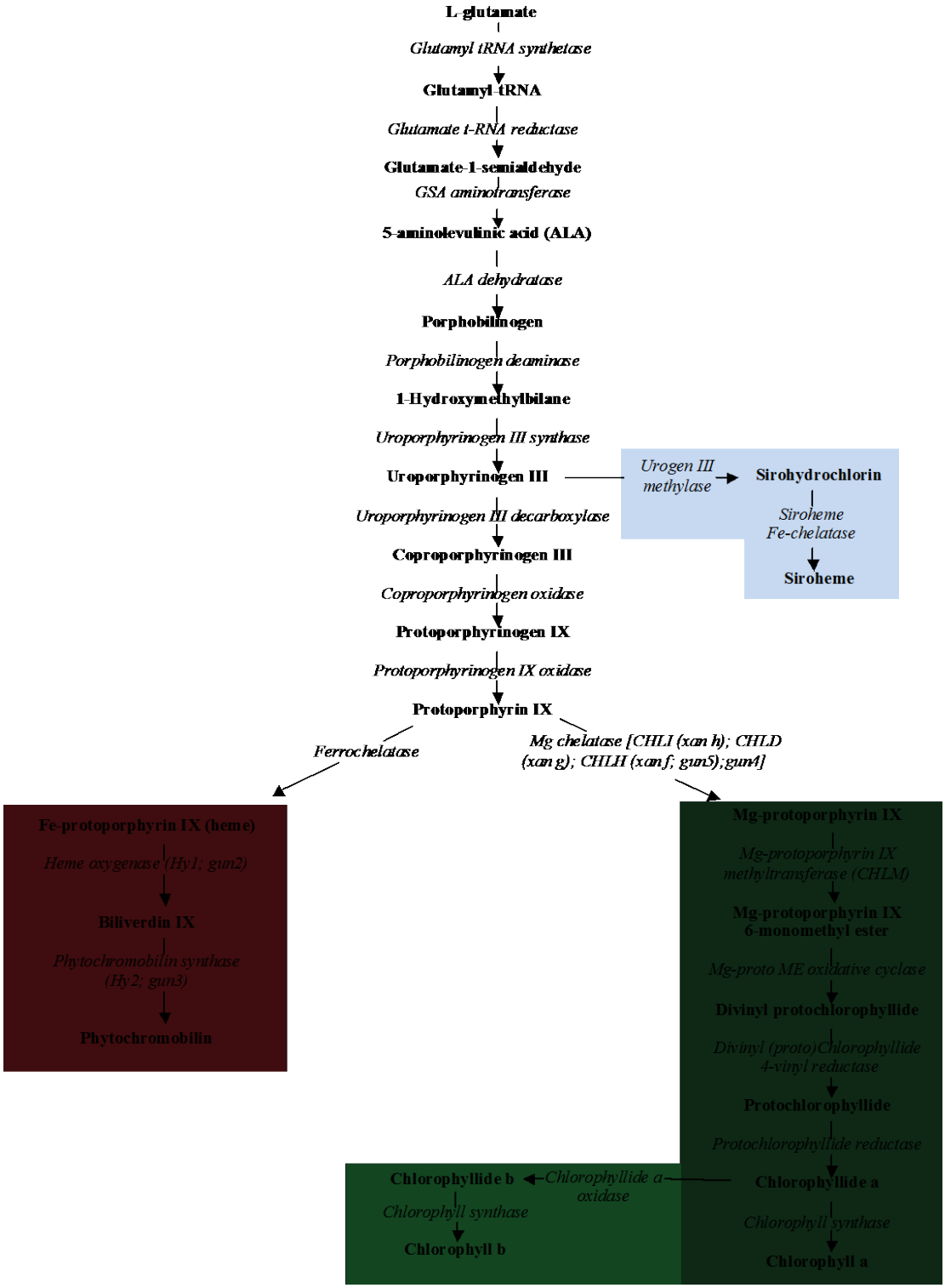


Figure 2. The Tetrapyrrole Biosynthetic Pathway in Plants

This figure shows the chlorophyll biosynthetic pathway, with Heme, Siroheme, and Chlorophyll branches outlined in red, blue and green, respectively.

1.2.2. Regulation of Tetrapyrrole Biosynthesis

Tight regulation of tetrapyrrole biosynthesis is important for plants, for numerous reasons. While all cells need heme for respiratory cytochromes and essential heme co-factored proteins, in photosynthetic tissues the requirement for chlorophyll greatly outweighs that of heme. The plant must regulate the distribution of Proto IX at the branch point of chlorophyll and heme biosynthesis in order to maximize its resources. Furthermore, accumulation of pathway intermediates including and after uroporphyrinogen III are potentially phototoxic. To prevent this possible toxicity, the plant must coordinate production of the tetrapyrrole and its associated apoprotein (Cornah et al. 2003; Meskauskiene et al. 2001).

The tetrapyrrole pathway is controlled at multiple levels. Regulation includes activities such as transcription, post-transcriptional RNA processing and modification and post-translational protein import and protein stabilization (Thompson and White 1991; van Grinsven and Kool 1988). For example, the diurnal transcriptional regulation of GluTR, the Mg-chelatase H-subunit (CHLH), and the Mg-chelatase co-factor GENOMES UNCOUPLED 4 (GUN4), has been demonstrated (Papenbrock et al. 1999; Peter and Grimm 2009).

1.2.2.1 Regulatory factors of Glutamyl-tRNA reductase

In higher plants, the flux of tetrapyrrole biosynthesis is thought to be controlled primarily via regulation of GluTR, for two reasons. First, application of exogenous ALA causes accumulation of porphyrin and magnesium-porphyrin intermediates, and second, the rate of ALA formation corresponds to chlorophyll

accumulation (Beale and Weinstein 1990). GluTR is encoded by a small family of genes made up of two to three members (Bougri and Grimm 1996; Ilag et al. 1994; Kumar et al. 1999; Kumar et al. 1996; Tanaka et al. 1996; Tanaka et al. 1997).

In plants, the gene encoding GluTR was named after the homologous *E. coli* gene, *HEMA* (Ilag et al. 1994), and is thus referred to as *HEMA1*, while the protein is always referred to as GluTR. *Arabidopsis*, *HEMA1* has been shown to be highly expressed throughout seedling development, and is light regulated via the photoreceptors phytochrome and cryptochrome (McCormac et al. 2001; McCormac and Terry 2002), as well as by feedback regulation by heme, as described below.

A second GluTR gene in *Arabidopsis*, *HEMA2*, is expressed only in roots and flowers, in a light-independent fashion (Kumar et al. 1996). While *HEMA1* is strongly transcriptionally up-regulated in white light, the transcript level of *HEMA2* shows little to no change in white light (McCormac et al. 2001).

1.2.2.2 Regulation of the pathway by heme

Organisms without a chlorophyll branch in the tetrapyrrole biosynthesis regulate this pathway primarily at the level of the initial precursor, ALA, which is negatively controlled by feedback from heme in bacteria (Wang et al. 1999), animals (Andrew et al. 1990), and yeast (Labbe-Bois and Labbe 1990). Likewise, ALA synthesis was suggested to control the flux through the pathway in plants, as the lag phase of chlorophyll accumulation in the light could be eliminated by the addition of ALA (Beale and Weinstein 1991).

Additional work provided further evidence that feedback regulation by heme controls the activity of Glu TR, at the step of delta-aminolevulinic acid synthesis. This enzyme was shown to be inhibited by the addition of heme *in vitro*, while the addition of other intermediates that might control feedback regulation, including Mg-Proto and protochlorophyllide (Pchlide), had much less effect (Beale and Weinstein 1991; Castelfranco and Zeng 1991; Pontoppidan and Kannangara 1994; Vothknecht et al. 1998; Weinstein and Beale 1985). The N-terminal 30 amino acids of GluTR are necessary for heme-mediated inhibition of this enzyme, but do not ablate enzymatic activity (Vothknecht et al. 1998).

In intact plastids, the use of apoperoxidase to artificially deplete the heme pool resulted in a 32% increase of ALA synthesis (Thomas and Weinstein 1992). Also, tomato mutants defective in phytochromobilin synthase and heme oxygenase show pale phenotypes with reduced chlorophyll, though these mutations occur in the heme branch of the pathway, which is distinct from the chlorophyll biosynthetic branch (Terry and Kendrick 1999). Use of the iron chelator 2,2'-dipyridyl to remove free $Fe^{2+/3+}$, and thus block ferrochelatase, has been shown to decrease in the heme level and increase the level of Pchlide, indicating that the usual control of Pchlide and other chlorophyll intermediates is compromised with the prevention of heme synthesis (Duggan and Gassman 1974). These studies support the proposition of heme as a regulator of tetrapyrrole biosynthesis.

1.2.2.3 Regulation of the chlorophyll branch by FLU

Although numerous studies support the concept of heme as a negative feedback regulator of tetrapyrrole biosynthesis, the studies done to demonstrate this control do not exclude the possibility that other control mechanisms of ALA synthesis exist (Yaronskaya et al. 2003). Regulators of tetrapyrrole biosynthesis can be identified by screening for mutants defective in the control of chlorophyll biosynthesis, such as the inability to restrict Pchlide accumulation, as with the mutant *tigrina d*. This mutant of *Hordeum vulgare* (barley) accumulates 10-15 fold higher amounts of Pchlide than wild type when grown in the dark (Nielsen 1974). The gene responsible for this phenotype was found to be orthologous to the *Arabidopsis thaliana FLU* gene (Lee et al. 2003), a novel repressor of tetrapyrrole biosynthesis.

a. Characterization of *FLU*

The *FLU* gene was detected in *Arabidopsis thaliana* using ethyl methanesulfonate (EMS) mutagenized seeds in an attempt to discover regulators of tetrapyrrole biosynthesis. Individuals with impaired tetrapyrrole regulation were identified by screening etiolated seedlings for an inability to restrict Pchlide accumulation in the dark, a mutation detected by the emission of red fluorescence after illumination with blue light. As with *tigrina-d* mutants, homozygous plants transferred into the light bleached and died, but could be rescued by germination in constant light (Meskauskiene et al. 2001).

Upon cloning, sequencing, and verification by complementation, the predicted FLU protein included features possibly functional in feedback control, in addition

to a potential chloroplastic transit peptide. The hydrophobic central part of FLU (AA125-146) was thought to be important for anchoring the protein within a membrane, while the C-terminus was predicted to be hydrophilic (Nielsen et al. 1997).

The hydrophilic C-terminal region of FLU contains two distinct domains implicated in protein-protein interactions, the first consisting of two tetratricopeptide repeats (TPRs), one of which contains an alanine residue that leads to the inactivation of FLU when replaced by valine. The second putative protein binding domain is a short coiled-coil motif adjacent to the hydrophobic membrane anchor (Meskauskiene et al. 2001).

Coiled-coil motifs are frequently involved in protein-protein interactions and have been implicated in regulation, as with the heat shock transcription factors from the yeast *Kluyveromyces lactis* (Peteranderl et al. 1999), and the plants *Arabidopsis thaliana* (Nover et al. 2001) and tomato (*Lycopersicon peruvianum*) (Scharf et al. 1998). Protein expression studies were used to demonstrate the interaction of FLU with GluTR, and to show that this interaction is due to the TPRs, rather than the coiled-coil motif (Meskauskiene and Apel 2002). Further investigation shows that FLU interacts with HEMA1 (GluTR expressed predominantly in leaves, in a light dependent fashion) but not with HEMA2 (GluTR expressed only in roots and flowers, in a light-independent fashion), and that this interaction requires the coiled-coil domain at the C-terminal part of GluTR (Goslings et al. 2004).

b. FLU-like Proteins (FLPs)

In *Chlamydomonas*, two FLU-like proteins (FLPs) were identified as alternative splicing products from a single gene. These proteins were found to have similar

functional domains to the *Arabidopsis* FLU protein, with a chloroplastic transit peptide, a hydrophobic region and two TPR motifs at the C-terminus. A notable difference in these FLPs, compared to FLU, is the presence of an additional coiled-coil motif; two are present in the *Chlamydomonas* proteins, as opposed to only one coiled-coil in FLU. The splice alternates differ from one another in that the short form (s-FLP) lacks 12 residues present in the long form of the protein (l-FLP); this insert resides between the first coiled coil motif and the hydrophobic region (Falciatore et al. 2005).

To test complementation of FLU with these proteins, *Arabidopsis flu* mutants were transformed with *Chlamydomonas FLP* constructs inserted into a binary vector driven by the 35S promoter. Most of these lines were able to complement the *Arabidopsis* mutation under 7 day 12h light/12 dark cycles, however, the FLP proteins were not able to complement the *flu* mutation when seedlings were grown for 4 days in the dark, then transferred to continuous light. This suggests that the FLP proteins only partially complement FLU (Falciatore et al. 2005).

c. The model for FLU activity

It was theorized that increased levels of Pchlide in *flu* mutants could be caused in two ways: a constitutive up-regulation of mRNAs for GluTR and/or GSA, or from feedback regulation of either of these enzymes. If FLU exerts feedback control to down-regulate ALA synthesis in dark-grown seedlings, its inactivation should result in an enhanced rate of ALA formation. Meskauskiene et al. demonstrated that synthesis of ALA in *flu* mutants exceeds that of wild type by a factor of 3 to 4, and furthermore, that the mRNAs for the two enzymes responsible for ALA synthesis, GluTR and GSA-AT, exhibited the same dark/light fluctuations as wild

type plants. Therefore, they postulated that the FLU protein acts as a direct feedback inhibitor of tetrapyrrole biosynthesis (Meskauskiene et al. 2001).

d. Regulation of FLU and FLU-like proteins

Analyses of *FLU* mRNA and protein concentrations from light and dark grown seedlings revealed that although *FLU* mRNA levels differed radically between treatments, similar protein levels were detected. When illuminated, the *FLU* mRNA concentration in etiolated seedlings increases steadily, though the protein levels remain constant. Interestingly, the mRNA and more notably, protein levels remain elevated upon a shift to a 25-hour dark period, in contrast to the light dependent changes of the concentration of GluTR protein, which increases steadily in the light, but is hardly detectable in dark grown seedlings. The FLU protein, then, is present greatly in excess of its target GluTR in the dark, and the ratio of FLU to GluTR changes significantly upon illumination (Goslings et al. 2004).

Likewise, in *Chlamydomonas*, both FLP proteins are expressed poorly in the dark, and quickly induced upon transfer to light. However, while the l-FLP reaches a maximum after three hours then decreases, s-FLP increases steadily and remains at a high level, with an expression level five times greater than that of the long form after five hours (Falciatore et al. 2005).

Investigation of FLP light induction showed that while both are induced by blue light, this response results in equal transcript levels of the two isoforms, in contrast to the different ratios seen when induced with white light. Therefore, it appears that other photoreceptors may be involved in the control of *FLP*

expression. In plants, a similar example of light-regulated alternative splicing exists in the hydroxylpyruvate reductase gene of pumpkin (Mano et al. 1999).

Upon further investigation, Falciatore et al. found that the light-induction of *FLP* gene expression is independent of photosynthetic electron flow and not subject to control via the redox state of the plastoquinone pool, but instead, increase of the s-FLP/l-FLP ratio correlates with the decline of the tetrapyrrole intermediate Pchlide. Interestingly, among the tetrapyrrole mutants examined for FLP overexpression, the highest ratio of s-FLP/l-FLP was found in dark grown mutants affected in the Mg-chelatase H-subunit, with one CHLH mutant (*chl-1*) exhibiting a ratio similar to that of cells grown in continuous light (Falciatore et al. 2005).

1.2.2.4 Concurrent control of tetrapyrrole biosynthesis by FLU and heme

To further examine the role of FLU as an inhibitor of GluTR, Goslings et al. (2004) utilized the non-permissive light/dark growth cycle to identify suppressor mutants that antagonize the effect of the *flu* mutation on Pchlide accumulation. In a search for second site mutants no longer able to accumulate high amounts of Pchlide in the dark, the investigators employed EMS mutagenesis of the *Arabidopsis flu*1-1 line. This line has similar *flu* mRNA levels to wild type, but due to a frame shift mutation resulting in a premature stop codon, has no detectable FLU protein.

Dark grown mutagenized seedlings were examined for red fluorescence and some, termed *ulf* (for reversal of *flu*) mutants, were observed to have a decreased fluorescence from that of the original *flu* mutants. Examination of the porphyrin

concentrations verified that the original *flu* mutants had a 10 fold greater concentration than that of the *ulf* mutants. Allelism tests were then performed on the mutants, demonstrating that none of the mutants were allelic to *flu*, and that the mutants comprised four different loci. One mutant, *ulf3*, had a long hypocotyl when grown under continuous light, as with the *hy* mutants of *Arabidopsis*. This mutant was examined in further detail.

The *Arabidopsis hy1* mutant lacks a functional heme oxygenase (Muramoto et al. 1999), which serves in the heme branch of tetrapyrrole biosynthesis to degrade heme, forming biliverdin IXα (BV IXα), carbon monoxide and iron (Gisk et al. 2010; Terry et al. 1993). Although the phytochrome apoprotein is synthesized normally, this mutant does not respond to red and far red light (Chory et al. 1996), most likely due to lack of a functional phytochrome. Dark grown seedlings of *hy1* contain reduced Pchlide levels (Montgomery et al. 1999); this has been attributed to an increase in free heme and thus inhibition of GluTR (Cornah et al. 2003; Terry and Kendrick 1999). Goslings et al. (2004) showed *ulf3* allelism to *hy1*, by both mapping and an allelism test (complementation), and upon sequencing, found a lesion leading to a premature stop codon in the heme oxygenase gene of the *ulf3* mutant. These data led them to conclude that the initial steps of tetrapyrrole biosynthesis are controlled concurrently by heme and FLU, which interact with the N-terminal 30 amino acids and a C-terminal region of GluTR, respectively.

a. The Heme Oxygenases of *Arabidopsis thaliana*

In *Arabidopsis*, in addition to HY1, there are three putative heme oxygenases (HOs). All four proteins are nuclear encoded and chloroplast localized. They can

be divided into two subfamilies: HO1 and HO2. Three of the four (HY1, HO3 and HO4), belong to the HO1 subfamily, and can convert heme to biliverdin IX using ferredoxin as an electron donor. HO2, on the other hand, is unable to bind or degrade heme and therefore is not a heme oxygenase (Gisk et al. 2010).

The HO2 subfamily differs from the HO1 subfamily by an inverted 34-55 spacer sequence rich in glutamate, aspartate, and glycine residues. Additionally, phenotypic analysis of the *Arabidopsis ho2-1* mutant suggests that HO2 contributes to proper photomorphogenesis (Davis et al. 2001). Interestingly, although all four HOs are able to bind Proto IX, HO2 forms a stable complex with Proto IX, and thus, it may be involved in the regulation of tetrapyrrole metabolism (Gisk et al. 2010).

b. ALA formation in *ulf3* and *hy1*

The *ulf3* and *hy1* were put into a wild type background and analyzed for ALA formation, which in *ulf3* was slightly lower than in *hy1* and wild type, while the *flu* mutant showed an ALA formation rate more than four times higher than that of wild type. Either mutant (*ulf3* or *hy1*) put into the *flu* line reduced the rate of ALA synthesis to about one third of the rate in *flu* alone. In addition, the double mutants (*ulf3/flu* and *hy1/flu*) both show a decreased accumulation of Pchlide (Goslings et al. 2004).

c. FLU and heme control tetrapyrrole biosynthesis concomitantly in light adapted plants

Light induction of *FLU* mRNA, as well as presence of the FLU protein in light adapted plants suggest that FLU is still involved in controlling tetrapyrrole

biosynthesis after greening. To assess the function of FLU in light adapted plants, Goslings et al. first measured the rate of ALA synthesis in wild type plants. They found the rate to be three to four times greater in light adapted plants than in seedlings transferred to the dark.

In light grown *flu* mutants, the ALA synthesis rate increased fourfold, compared to wild type, while *hy1* mutants produced ALA at half the rate of wild type. Plants with a double mutation (*flu*/*hy1*) exhibited a suppressed ALA synthesis rate compared to those with the single *flu* mutation (Goslings et al. 2004).

The plants were then transferred to the dark to block the conversion of Pchlide to chlorophyllide (Chlide), and the accumulation of Pchlide in each was shown to be consistent with the rate of ALA synthesis in the corresponding light grown plant. Hence, heme and FLU seem to control the rate of ALA synthesis and the accumulation of Pchlide in both etiolated and green seedlings (Goslings et al. 2004).

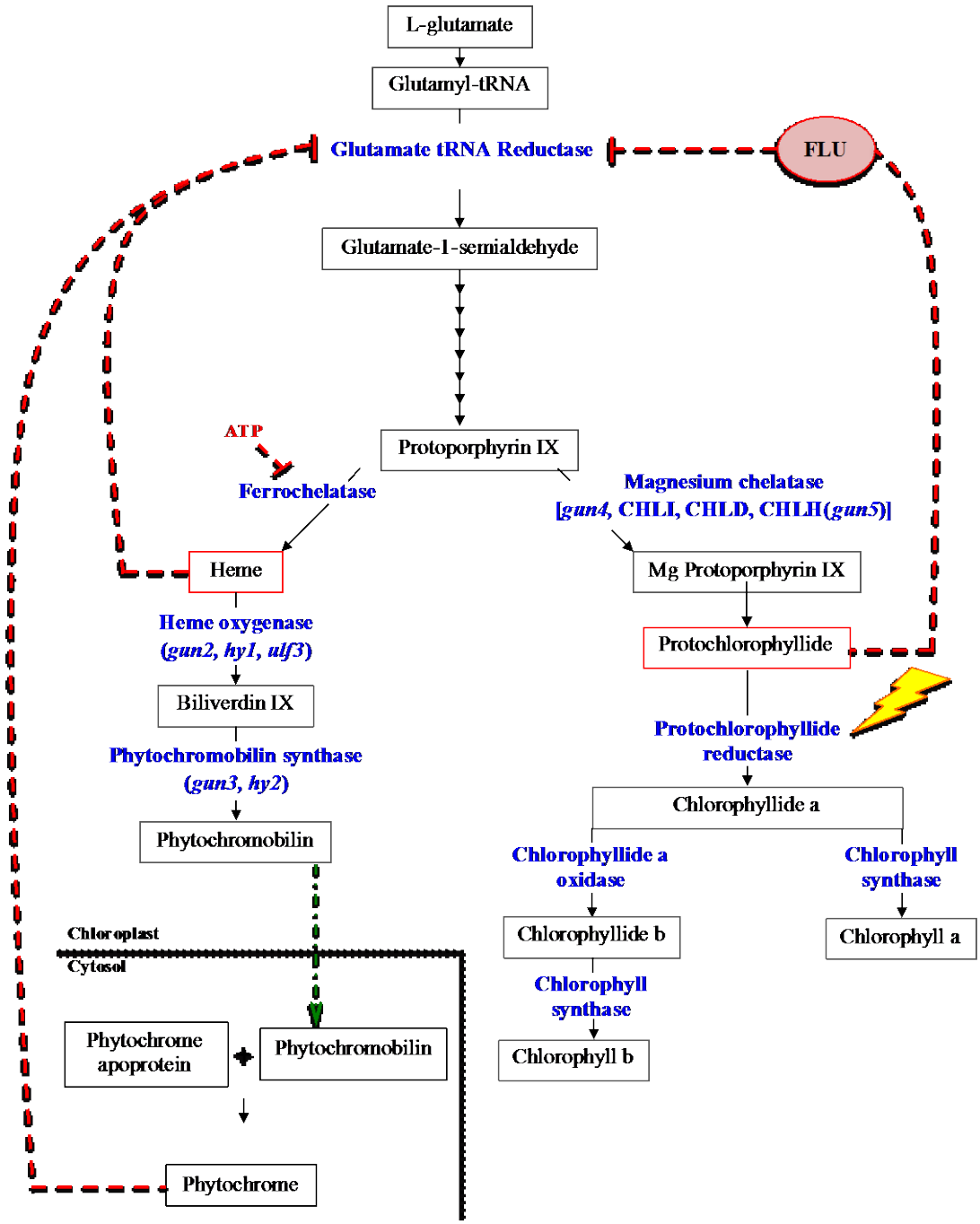


Figure 3. Regulation of Tetrapyrrole Biosynthesis

This figure shows an overview of tetrapyrrole biosynthesis and its regulation in higher plants. Enzymes are shown in blue, black arrows indicate enzymatic reactions, black names represent pathway intermediates, and dotted red lines represent regulatory interactions. The black dotted line represents a change in cellular compartment.

Protochlorophyllide reductase requires light to catalyze the conversion of Protochlorophyllide to Chlorophyllide a., as is represented by a lightning bolt. The FLU protein and the accumulation of the intermediate Heme interacts with GluTR to regulate the synthesis of ALA. In addition to phytochrome, depicted because of its relationship with this pathway, GluTR is light regulated by cryptochrome, a blue light receptor. The Arabidopsis mutants GUN2, GUN3, GUN4, and GUN5 all affect chlorophyll accumulation.

1.3 The Magnesium Chelatase Enzyme – Characterization

Mg-chelatase is the multi-subunit enzyme complex that catalyzes the first committed step in Chlorophyll biosynthesis, by the insertion of a divalent magnesium ion into the tetrapyrrole intermediate Proto IX. Comprised of three subunits, Mg-chelatase I (CHLI), Mg-chelatase D (CHLD), and Mg-chelatase H (CHLH) (Walker and Willows 1997), this enzyme catalyzes porphyrin metallation in an ATP dependent fashion (Fodje et al. 2001). In addition, the appreciable accumulation of chlorophyll in photosynthetic plant tissue requires an additional protein called GUN4 (Larkin et al. 2003). The optimal stoichiometric ratio of each component in plants has yet to be demonstrated.

Early work done on Mg-chelatase comprised studies done in photosynthetic bacteria, such as the purple bacteria *Rhodobacter sphaeroides* as well as in higher plants, including cucumber (*Cucumis sativus*) (Castelfranco et al. 1979; Fuesler et al. 1984a; Fuesler et al. 1981), pea (*Pisum sativum*) (Walker and Weinstein 1994), and barley (*Hordeum vulgare)*(Axelsson et al. 2006b).

In the 1950's, three different *Hordeum vulgare* (barley) mutants were identified as accumulating Proto (Henningsen et al. 1993), and were grouped into 3 loci, denoted Xantha-f, -g, and –h (von Wettstein et al. 1995). Castelfranco (1979) and colleagues demonstrated that Mg-chelatase activity required magnesium, ATP, and Proto as substrates (Castelfranco et al. 1979; Fuesler et al. 1981; Fuesler et al. 1984b; Pardo et al. 1980), and Walker and Weinstein (1991) showed that Mg-chelatase activity could only be measured in intact chloroplasts. They demonstrated that disruption of the plastids led to a complete loss of activity, and furthermore, that hydrolysis of ATP is necessary for enzymatic activity, as

opposed to the use of ATP as an allosteric activator, or simply as a structural component of the enzyme, for example, to facilitate the correct placement of magnesium for insertion into Proto IX (Walker and Weinstein 1991).

In developing cucumber chloroplasts, enzyme inactivation studies employing the mercurial reagents *p*-chloromercuribenzoate (PCMB, a membrane permeable organomercury compound which reacts with thiol groups) and the membrane impermeable *p*-chloromercuribenzene sulfonate (PCMBS, which does not penetrate the chloroplast envelope), led to the conclusion that Mg-chelatase was localized in the chloroplast envelope (Fuesler et al. 1984b). However, re-examination of this study by Walker and Weinstein showed no difference between the effects of the two mercurials, PCMB and PCMBS. They demonstrated, rather, that Mg-chelatase is not in the inter-envelope space, using a plastid-free assay. In developing this assay, they showed that Mg-chelatase was a multi-component system, with at least one membrane associated fraction (the others being soluble). They established that the membrane-bound fraction is not firmly attached to the membrane, and in the presence of low $MgCl_2$ concentrations, the membrane-bound fraction becomes completely solubilized. Nevertheless, they were unable to definitively localize the enzyme (Walker and Weinstein 1995).

Characterization of the subunits of Mg-chelatase is as follows.

1.3.1. Magnesium Chelatase subunit D

CHLD is a subunit comprised of four distinct regions: an N-terminal half that has sequence similarity with the AAA^+ (ATPases Associated with diverse cellular Activities) domain of CHLI (Fodje et al. 2001), a poly-proline linker region, a

glutamate/aspartate rich region, and a C-terminal half containing an Integrin I domain (Fodje et al. 2001). This subunit has a molecular mass of 60 – 90 kDa, and is thought to exist in the holo-enzyme as a hexameric ring.

Poly-proline regions have been implicated in mediating protein-protein interactions (MacArthur and Thornton 1991) and in the stabilization of oligomeric complexes (Bergdoll et al. 1997). Integrin-I domains are well-characterized domains that function within αβ-hetero-dimeric transmembrane receptors involved in cell-cell and cell-matrix interactions (Hynes 2002). Integrin-I domains characteristically include a metal ion-dependent adhesion site (MIDAS) motif consisting of two conserved sequence regions (DXSXS…TD) located approximately 90 amino acid residues apart (Lee et al. 1995). The MIDAS motif typically allows the coordination of a Mg^{2+} or a Mn^{2+} ion. Mutation of conserved residues in the MIDAS motif of the *R. capsulatus* D-subunit greatly reduces Mg-chelatase activity (Axelsson et al. 2006a).

Multiple lines of evidence have shown that the D and I subunits interact with each other. Preincubation of these subunits in the presence of ATP and Mg^{2+} hastens the onset of the Mg-chelatase reaction *in vitro* (Guo et al. 1998; Jensen et al. 1998; Willows et al. 1996; Willows et al. 1999). Yeast two-hybrid analysis has supported a direct interaction between CHLD and CHLI (Papenbrock et al. 1997). Furthermore, the D-subunit was not detectable in *chlI* mutants in barley, whereas wild type levels of the D-subunit were observed in barley *chlH* mutants, suggesting that a functional I-subunit is required for maintaining the D-subunit in the cells (Lake et al. 2004).

Single-particle cryo-electron microscopy analyses have probed the structure of the complex between the Mg-chelatase D and I subunits of *Rhodobacter capsulatus*. Reconstructions of the complex have revealed that the AAA^+ modules of subunits D and I form a hexameric complex consisting of 3 dimers related by a three-fold axis. The different conformations of the complex in the presence of ATP and ADP suggested that the C-terminal integrin-I domains of the D subunits are involved in transmitting conformational changes between the D- and I- subunits (Lundqvist et al. 2010).

The current model of Mg-chelatase assembly involves the D subunit serving as a platform or scaffold for assembly of the complex (Axelsson et al. 2006a). The subsequent steps involve the AAA^+ domains of the D and I subunits forming a hexameric complex in the presence of ATP and Mg^{2+}, while the large H subunit binds Proto IX. The H subunit complexed with Proto IX serves as a substrate for the complex between the I- and D-subunits, resulting in the insertion of Mg^{2+} into Proto IX.

1.3.2. Magnesium Chelatase subunit I

CHLI is the smallest of the subunits of Mg-chelatase subunits with a molecular weight of 37-46 kDa (Walker and Willows 1997). The I-subunit is classified as an AAA^+ family protein, and contains the distinctive Walker A and Walker B motifs of an ATPase (Fodje et al. 2001). AAA proteins have diverse range of functions across all domains of life, often involved in processes such as folding, assembly, and disassembly of protein complexes (Hanson and Whiteheart 2005; Vale 2000).

The 3-dimensional crystal structure of the I-subunit from *Rhodobacter capsulatus* has been determined to a resolution of 2.1 Angstroms (Fodje et al., 2001). This represents the only high-resolution structure of a subunit of Mg-chelatase. The structure of BchI showed a novel arrangement of domains with the C-terminal helical domain located behind the nucleotide-binding site compared to its more typical position on top of the nucleotide-binding site in other members of this family.

AAA^+ proteins typically associate as hexameric or heptameric rings. The I-subunit from *R. capsulatus* has been demonstrated to form hexameric rings (Fodje et al. 2001; Willows et al. 2004)), whereas the *Synechocystis* PCC6803 CHLI has been reported to form a heptameric ring (Reid et al. 2003).

Comparison of the Walker A and Walker B Mg^{2+}-ATPase motifs of CHLD and CHLI shows important differences, with the CHLI motifs highly conserved, whereas the CHLD motifs are more divergent. This is consistent with the activities of these proteins, as CHLI is known to be able to hydrolyse adenosine triphosphate (ATP) (Hansson et al. 2002; Kobayashi et al. 2008; Lake et al. 2004), whereas no ATPase activity has been detected from the D subunit (Hansson and Kannangara 1997; Jensen et al. 1999; Petersen et al. 1999). Upon ATP hydrolysis, CHLI undergoes a significant conformational change and essentially acts as a motor for the Mg-chelatase complex (Lundqvist et al. 2010).

1.3.3. Magnesium Chelatase subunit H

CHLH is the largest of the subunits of Mg-chelatase with a molecular weight of approximately 140 kDa. The H-subunit, but not the D- or I-subunits, has been shown to bind Proto IX, implicating it as the catalytic portion of the Mg-chelatase enzyme (Willows and Beale 1998; Willows et al. 1996). Also known as GUN5 (see section 1.4.2.1), this subunit has been shown to bind Proto IX in a ratio of 1 H-subunit: 1 ProtoIX (Walker and Willows 1997). CHLH appears to interact only transiently with the CHLD and CHLI subunits, acting as a substrate for the I/D complex (Sawicki and Willows 2008).

Low resolution structures (25-30 Angstrom) of the H subunit from have been determined by electron microscopy and single-particle three-dimensional reconstruction (Sirijovski et al. 2008) and by small-angle X-ray scattering and single particle reconstruction (Qian et al. 2012)). Based on these studies, the H-subunit appears to consist of two domains joined by a short linker: a large C-terminal domain that forms a cage-like structure, and a small globular N-terminal domain. Porphyrin-binding induces distinct conformational changes (Sirijovski et al. 2008). The cage-like structure has suggested the possibility that CHLH encloses the labile product of Mg-chelatase, Mg-Proto, and chaperones it to the active site of the next enzyme in the pathway, magnesium protoporphyrin methyltransferase, protecting it from photooxidation (Qian et al. 2012).

In addition to its catalytic role, CHLH appears to be involved in an eclectic range of other functions or interactions, this may be one reason for the large size of this subunit. The GUN4 protein in both plants and the cyanobacterium *Synechocystis*

have been shown to form a complex with CHLH (Larkin et al. 2003; Sobotka et al. 2008). GUN4 (see section 1.4.2.1) appears to be an enhancer of the Mg-chelatase reaction, and it has been proposed to act as a molecular switch by reducing the Mg^{2+} concentration required for activity at low porphyrin concentrations (Davison et al. 2005). CHLH (GUN5) itself has also been implicated in retrograde signaling (see section 1.4.2). In *Synechocystis*, it has been reported that CHLH is a negative regulator of the RNA polymerase sigma factor SigE (Osanai et al. 2009), although this may not apply to higher plants.

Interestingly, the H-subunit has also been reported to be an ABA receptor (Shen et al. 2006a; Wu et al. 2009). More recently, the role of CHLH as an ABA receptor has been questioned. Recombinant Xan-F, the barley ortholog of CHLH, was found incapable of binding ABA, and *xan-f* mutants showed normal ABA responsiveness (Muller and Hansson 2009). In *Arabidopsis*, there is evidence suggesting that CHLH affects ABA signaling in stomatal guard cells, but recombinant CHLH was not able to bind ABA (Tsuzuki et al. 2011). CHLH expression is suppressed by the circadian clock protein TOC1, which directly interacts with the CHLH promoter, expression of TOC1 itself is highly induced by ABA (Legnaioli et al. 2009). The link between CHLH and ABA remains to be fully clarified.

The multiple additional regulatory functions/interactions of CLH probably reflects the fact that Mg-chelatase is located at the key branch point between heme and chlorophyll biosynthesis (fig. 3). This is a critical regulatory point for a plant cell as insufficient levels of chlorophyll pigments would impair the photosynthetic

apparatus, but overproduction of these pigments could lead to light-induced photodamage.

1.3.4. Enzyme Assembly

The Mg-chelatase enzyme is thought to assemble in a two-step process. The first step, called the activation step, requires ATP and Mg and involves only subunits I and D. This step does not involve ATP hydrolysis; rather, ATP is used in assembly of the enzyme. The second step, chelation, requires subunit H, Proto, and in oxygenic organisms, GUN4 (Axelsson et al. 2006a).

1.3.5. Regulation of Magnesium Chelatase

A study by Castells et al. has shown that the TOC1 component regulates the diurnal expression of Mg-chelatase H (CHLH, *A. thaliana* ABAR/GUN5) by direct binding to its promoter. In this study, ABA treatment resulted in TOC1 induction at midday; this induction was demonstrated to control the phase of TOC1 binding and the expression of CHLH. Conversely, plants deficient in CHLH (via RNAi) were not subject to ABA mediated induction of TOC1 (Castells et al. 2010).

GUN4 regulation of Mg-chelatase has been shown to occur post-translationally (Peter and Grimm 2009), and transcriptional studies show that I and H are induced by light in developing *Arabidopsis* and soybean, although the induction of I is not so pronounced in Barley.

1.4 Plant Regulation and Signaling

Light is an environmental control factor of nuclear gene expression in higher plants, primarily at the transcriptional level through the phytochrome system. This has been demonstrated for genes encoding plastid proteins such as the light harvesting chlorophyll *a/b* protein (LHCII), NADP-protochlorophyllide oxidoreductase, and Rubisco (ribulose 1,5-bisphosphate carboxylase) small subunit (RBCS) (Mosinger et al. 1985).

1.4.1. Anterograde signaling

Most of the proteins localized in the chloroplast are nuclear encoded (Nott et al. 2006). Anterograde signaling is the process by which the nucleus exerts control over chloroplast-encoded genes, such as those involved in organellar gene expression. In this process, the nucleus encodes regulators that act to coordinate expression of nuclear and chloroplastic genes so that proper stoichiometry of plastidic protein complexes is achieved (Jung and Chory 2010).

1.4.2. Retrograde Signaling

According to the endosymbiotic theory, organelles are derived from prokaryotes. Over time, most of the chloroplast's genome has been relocated to the nucleus, with the exception of about 50 protein coding genes and 30 tRNAs (Zurawski and Clegg 1987). The coding genes include many genes for proteins involved in organellar gene expression, as well as some of the genes encoding photosynthetic proteins (Voigt et al. 2009). Approximately 2100 nuclear genes are predicted to encode chloroplast genes in *Arabidopsis thaliana* and as many as 4800 in *Oryza sativa* (Richly and Leister 2004). Coordination of nuclear and plastidic genome

expression is thought to be essential to chloroplast biogenesis (Goldschmidt-Clermont 1998), since a good deal of the photosynthetic machinery is comprised of complexes containing both nuclear and chloroplast encoded proteins (Nott et al. 2006).

Retrograde signaling involves a cue from a location outside the nucleus, such as the chloroplast, endoplasmic reticulum, or mitochondrion, that affects the expression of nuclear encoded genes (Liu and Butow 2006; Ron and Walter 2007). Thus, the nucleus is able to appropriately respond to intracellular cues that keep it apprised of, for example, the developmental and functional state of the chloroplast. The nature of the chloroplastic signal has yet to be revealed, although there have been extensive studies providing evidence for its existence.

The first study indicating plastid control over the nuclear genome was done by Bradbeer et al. (1979) on *Hordeum vulgare* (barley) mutants deficient in carotenoid production. Seedlings of this mutant have either white striped or completely white leaves, with the white leaves lacking plastidic ribosomes and thus protein synthesis in the plastids, though cytosolic ribosomal abundance is normal. The large and small subunits of RuBisCo, all subunits for coupling factor CF_1 (two of which are cytoplasmically synthesized and three are plastid synthesized), and cytoplasmically synthesized ferredoxin-NADP reductase are undetectable in these mutants. The white sections contain undifferentiated photosynthetically inactive plastids, and the authors found the abundance of two cytoplasmically synthesized, plastid localized Calvin cycle enzymes, Phosphoribulokinase and D-glyceraldehyde-3-phosphate:$NADP^+$ oxidoreductase (phosphorylating) (a.k.a. Glyceraldehyde-3-phosphate dehydrogenase), to be

significantly reduced (phosphoribulokinase uses one ATP to phosphorylate ribulose-5-phosphate into ribulose 1,5-bisphosphate in the last step of the Calvin cycle (Hurwitz et al. 1956), while D-glyceraldehyde-3-phosphate: $NADP^+$ oxido-reductase reduces 1,3-bisphosphoglycerate to glyceraldehyde-3-phosphate with the release of inorganic phosphate, yielding glyceraldehyde-3-phosphate (Shih et al. 1991)). They concluded that cytoplasmic synthesis of plastid localized nuclear gene products may be controlled by the plastid (Bradbeer et al. 1979).

A number of studies on retrograde signaling have employed the use of inhibitors in an effort to elucidate the mechanism by which signaling occurs, as well as the signal itself (Sagar et al. 1988; Strand et al. 2003; Sullivan and Gray 1999; Susek et al. 1993). These studies make use of Norflurazon (NF), chloramphenicol and lincomycin, and aim to identify genes affected when cross talk between the chloroplast and the nucleus is disrupted. In the absence of normal chloroplast development, nuclear encoded chloroplast localized genes are down regulated. The focus of these studies was to identify genes that are still expressed when chloroplast development is impaired. NF inhibits the enzyme phytoene desaturase in the carotenoid biosynthesis pathway, which results in photo bleaching and arrested plastid development (Breitenbach et al. 2001). Chloramphenicol and lincomycin also result in arrested plastid development, as they inhibit intraplastidic protein translation (Gray et al. 1995; Smith-Johannsen and Gibbs 1972). To date, three mechanisms have been identified in retrograde signaling. These are: inhibition of plastid gene expression (PGE), accumulation of tetrapyrrole biosynthetic intermediates, and changes in the redox state of the

plastid. These pathways are engaged via two genes, *GUN1* and *ABI4* (Koussevitzky et al. 2007a).

1.4.2.1 The Arabidopsis gun *mutants*

In a study done by Susek, Asubel, and Chory in 1993, a series of *Arabidopsis* mutants were identified in a screen designed to identify genes responsible for the coupling of the nuclear and chloroplast genomes.

The screen employed a construct comprised of the promoter of the nuclear encoded *LHCB* (*CAB3*) gene fused with a hygromycin resistance gene, as well as an *LHCB* promoter driven β-glucuronidase (GUS) construct, which functioned as a reporter by converting commercially available substrates into colored or fluorescent products under conditions where *CAB3* was normally expressed (Susek et al. 1993). The plants were EMS mutagenized, then screened on the herbicide NF. This herbicide inhibits carotenoid biosynthesis early in the pathway (at phytoene desaturase), thus resulting in photo-bleached plants due to the excitation of chlorophyll to its triplet state and subsequent reaction with molecular oxygen, to form singlet oxygen. Photo-bleaching results in a lack of chloroplastic development; under these conditions the plants were screened for lines expressing both the *CAB3*-promoter driven transgenes. In a normally responding photo-bleached plant, the *CAB3*-promoter driven transgenes would not be expressed, as a result of a retrograde chloroplastic signal alerting the nucleus of the compromised state of the plastid.

a. The *gun1* mutant

The *gun1* mutant lacks a chloroplast localized pentatricopeptide repeat protein; this family of proteins is involved in the maturation, translation and turnover of mRNAs (Koussevitzky et al. 2007a). It is defective in the switch from dark to light growth, implicating this gene in modulation of the coordination of genes involved in the early stages of photosynthetic growth (Susek et al. 1993). The *gun1* mutant has a phenotype on chloramphenicol and lincomycin, suggesting that the protein product of this gene plays a role in the plastid gene-expression dependent pathway (Nott et al. 2006). Furthermore, the GUN1 protein has a MUTS-related domain, and may allow GUN1 to bind nucleic acids (Koussevitzky et al. 2007a).

b. The *gun2* mutant

The *gun2* mutant of *Arabidopsis* is deficient in heme oxygenase (hy1) (Mochizuki et al. 2001). Heme oxygenase 1 converts heme to biliverdin IX (fig. 3) (Davis et al. 1999).

c. The *gun3* mutant

The gene responsible for the *gun3* mutation was identified as phytochromobilin synthase (hy2/elm1) (Mochizuki et al. 2001). Phytochromobilin synthase converts biliverdin to phytochromobilin (fig. 3) (Parks and Quail 1991).

d. The *gun4* mutant

GUN4 is a novel chloroplast protein that binds Proto IX and Mg-Proto. The function has not been elucidated, however, while it promotes the activity of Mg-chelatase, the enzyme does not require GUN4 for catalysis (Larkin et al. 2003).

Although *in vitro* analysis of Mg-chelatase has shown that it is possible to achieve catalysis with only the three subunits (CHLI, CHLD, CHLH) and cofactors (ATP, Mg^{2+}) it has been demonstrated that higher plants are incapable of accumulating chlorophyll effectively in the absence of the protein GUN4 (Larkin et al. 2003). GUN4 related proteins have only been found in species that carry out oxygenic photosynthesis, and the transcript abundance is restricted to green tissues (Peter and Grimm 2009). It has been suggested that the GUN4 protein might be tethered to chloroplast membranes by protein-protein interactions (Larkin et al. 2003) and has been shown to complex with CHLH and a fragment of CHLH to form a 500kDa thylakoid-localized complex. In-vitro experiments show that GUN4 from *Synechocystis* increases the Mg-chelatase reaction rate threefold, and that this rate increase is accomplished through the ability of the protein to bind both the substrate (Proto IX) and the product (Mg-Proto) of Mg-chelatase.

A recent study (Peter and Grimm 2009) using GUN4-deficient and overexpressing plants suggests that GUN4 has multiple activities, including a role in posttranslational regulation, where it contributes to protein stability of enzymes in magnesium porphyrin biosynthesis. GUN4 may be essential for ALA biosynthesis, and at high light intensities, GUN4 abundance is increased, in turn stimulating ALA synthesis and Mg-chelatase activity. GUN4 may have an additional potential role as a relief shunt to protect the cell from accumulation of excessive magnesium porphyrins.

e. The *gun5* mutant

GUN5 has been identified as the CHLH subunit of Mg-chelatase, and may have a particular role in plastid-to-nucleus signaling (Mochizuki et al. 2001). This protein is described further in section 1.3.3.

f. Additional *gun* mutants

More recently, additional *gun* mutants were identified in screens revealing four *cry1* alleles (Ruckle et al. 2007), along with *hy5* (Ruckle et al. 2007), *abi4* (*ABA*-insensitive *4*) (Koussevitzky et al. 2007b), and *glk1*/*glk2* (Waters et al. 2009).

i) the cryptochromes

The *cry1* alleles encode cryptochromes, which are blue light receptors that have a role in the regulation of germination, elongation, and circadian rhythm, amongst others (Ruckle et al. 2007). They have a flavin chromophore which, when reduced by light is transported into the nucleus where it affects the turgor pressure, causing subsequent stem elongation in the plant. Cryptochromes are derived from and closely related to photolyase, a bacterial enzyme that is activated by light and participates in DNA damage repair (Chaves et al. 2011). In eukaryotes the cryptochromes have lost their original enzymatic activity. Two chromophores: pterin (in the form of 5, 10-methenyl-6,7,8-tri-hydrofolic acid (MHF)) and flavin (in the form of flavin adenine dinucleotide (FAD)) may both absorb a photon; energy captured by pterin is transferred to flavin, which is then reduced to FADH. FADH probably mediates the phosphorylation of a cryptochrome domain, in turn triggering a signal transduction chain possibly affecting gene regulation in the nucleus) (Hsu et al. 1996; Nefissi et al. 2011).

ii) *hy5*

Acting antagonistically with COP1 (a ubiquitin protein ligase which represses photomorphogenesis in darkness by mediating ubiquitination and subsequent degradation of light induced transcription factors (Holm et al. 2002)) in regulating seedling development, HY5 is a bZIP transcription factor that binds directly to the promoters of light-inducible genes, promoting their expression and photomorphogenic development (Ulm and Nagy 2005). COP1 is a RING-finger protein with WD-40 repeats whose nuclear abundance is negatively regulated by light. COP1 directly interacts with HY5 in the nucleus to negatively regulate its activity. The abundance of HY5 is directly correlated with the extent of photomorphogenic development and the COP1-HY5 interaction may specifically target HY5 for proteasome-mediated degradation in the nucleus (Osterlund et al. 2000).

iii) *abi4*

ABI4 encodes an ABA-regulated AP2 domain transcription factor that is expressed in roots (Rook et al. 2006). Root expression of *ABI4* is enhanced by ABA and cytokinin and is repressed by auxin. ABI4 appears to have a role in controlling the formation of lateral roots via reduction of polar auxin transport (Shkolnik-Inbar and Bar-Zvi 2010).

iv) *glk1/glk2*

The *glk1/glk2* genes encode GOLDEN2-LIKE (GLK) transcription factors that are required for chloroplast development (Waters et al. 2009). An *Arabidopsis* double mutant, *glk1/glk2*, accumulates abnormal levels of chlorophyll precursors, and overexpression of the *GLK* genes leads to increased accumulation of transcripts for chlorophyll biosynthesis, light harvesting, and electron transport genes. This regulation appears to be independent of the phytochrome B signaling pathway and the two *GLK* genes are differentially responsive to plastid retrograde signals (Waters et al. 2009).

1.4.2.2 Signaling and Redox state of the Chloroplast

Because the introduction of molecular oxygen within the cell is a necessary part of photosynthesis, the cell must find ways to deal with the production of reactive oxygen species (ROS) (Schippers et al. 2012). These partially reduced or activated oxygen derivatives are highly reactive and can lead to destruction of cell components, and eventually, cell death. As a result, there are mechanisms in place to scavenge these reactive molecules, including singlet oxygen ($^1O_2$), superoxide anion (O_2^-), hydrogen peroxide (H_2O_2) and hydroxyl radical (•OH). However, recent studies suggest that plant ROS are not simply toxic byproducts of aerobic metabolism, but that they also play a role as key regulators of growth, development, and defense pathways (Mittler et al. 2004). It has been suggested that the chloroplastic signal to the nucleus employs redox signals (Apel and Hirt 2004; Fryer et al. 2003; Karpinski et al. 1997; Mateo et al. 2004; Nott et al. 2006; Sattler et al. 2006), and that compromised chloroplasts may elicit cell death

signaling pathways via the accumulation of singlet oxygen (Apel and Hirt 2004; Przybyla et al. 2008; Wagner et al. 2004).

1.4.2.3 Plastid Gene Expression dependent Signaling

NF is an herbicide that inhibits the enzyme phytoene desaturase (Breitenbach et al. 2001). This enzyme functions early in carotenoid biosynthesis, and the use of NF results in photooxidative damage of the chloroplasts in white light. In far-red light no photooxidative effects are observed, even in the virtual absence of carotenoids. Photooxidative damage of the plastids prevents phytochrome-controlled generation of nuclear-encoded mRNAs of RBCS and LHCII while synthesis or representative enzymes of cytosol, mitochondria and glyoxisomes –including phytochrome-induced proteins –was not impaired (Oelmuller and Mohr 1986).

Chloramphenicol treatment results in arrested chloroplast development as it is a potent inhibitor of protein synthesis on prokaryotic type ribosomes, including plastidic ribosomes, and thus blocks synthesis of plastidic proteins (Oelmuller et al. 1986), which in turn results in the inhibition of nuclear photosynthetic gene expression (Nott et al. 2006). Chloramphenicol, however, does not inhibit the cytoplasmic synthesis or import of nuclear encoded, Chloroplast-localized proteins. In addition, the effects of Chloramphenicol can only be seen in white mustard (*Sinapis alba* L.) when applied between 36 and 48 hours after sowing, implying the retrograde signal is not generated prior to a certain developmental stage of the plastid (Oelmuller and Mohr 1986).

1.4.2.4 Magnesium Protoporphyrin IX as a signaling molecule

Mg-Proto has been proposed as a signaling molecule to control the expression of plastid localized, nuclear encoded genes (Strand et al. 2003). However, more recent evidence demonstrates that Mg-Proto is not the signal for retrograde signaling (Mochizuki et al. 2008; Moulin et al. 2008); the molecular signal for retrograde signaling remains unknown.

1.4.2.5 Transcription Factor mediated Retrograde Signaling

Recent work has implicated a plant homeodomain transcription factor as a potential mediator of retrograde signaling (Sun et al. 2011). Plant homeodomain transcription factors affect histone modifications resulting in chromatin-mediated gene regulation. This transcription factor, PTM, has multiple transmembrane domains and localizes to the plant chloroplast envelope. PTM is proteolytically cleaved in response to retrograde signals. The N-terminal portion of cleaved PTM accumulates in the nucleus, where it activates ABI4 transcription. ABI4 is an ABA-regulated AP2 domain transcription factor (see section 1.4.2.1)

1.5 Import into the Chloroplast

Although chloroplasts retain a functional genome from their cyanobacterial ancestor, with evolution the vast majority of the chloroplastic genes have been transferred to the host nucleus until only about one hundred different proteins are retained in the chloroplastic DNA (Martin et al. 2002; Timmis et al. 2004). Over ninety percent of the 3000 different proteins needed to construct a functional

chloroplast are nuclear encoded and synthesized cytosolically (Keegstra and Cline 1999; Leister 2003).

As with the majority of chloroplast-localized proteins, subunits of Mg-chelatase are encoded in the nucleus. Efficacy of this enzyme therefore depends upon proper import into the chloroplast using an N-terminal targeting sequence, called the transit peptide (TP). This portion of the peptide allows the preprotein to enter the plastid, and is usually removed upon arrival. Some transit peptides are non-cleavable; however, these are mostly found in outer envelope proteins (OEP), such as TOC34 (translocon at the outer envelope of chloroplasts) and TOC159 (Strittmatter et al. 2010). Chloroplast targeted precursor proteins made in the cytosol enter the stroma via the TOC/TIC import machinery located in the outer and inner envelope membrane (Bedard and Jarvis 2005; Kessler and Schnell 2006; Smith 2006; Soll and Schleiff 2004). In this ATP driven process, the precursor proteins are guided to the TOC transporter located on the outer envelope by cytosolic chaperones, translocated across in an ATP dependent manner, and processed in the stroma (Ruprecht et al. 2010). After import into the chloroplast via TOC/TIC, proteins further compartmentalized to the thylakoid membrane or lumen use one of four targeting pathways (Gutensohn et al. 2006; Jarvis 2008; Jarvis and Robinson 2004; Keegstra and Cline 1999).

The Secretory (Sec) and Twin-arginine translocase (Tat) pathways are used by proteins targeted to the thylakoid lumen; these proteins have a cleavable targeting signal which, as with the transit peptide, is removed upon import. The ATP-dependent Sec pathway accepts only unfolded proteins. In contrast, the Tat pathway, using the thylakoidal proton gradient as its sole energy source, is able to

translocate fully folded proteins which likely achieve their final conformation, by oligomerization or binding of a co-factor, in the stroma (Jarvis 2008).

Proteins localized in the thylakoid membrane use either the signal recognition particle (SRP)-dependent pathway or the spontaneous insertion pathway. The GTP-driven SRP-dependent pathway is mainly concerned with the insertion of integral membrane proteins of the light-harvesting complexes. Finally, the 'spontaneous' pathway has no energy requirement or protein transport machinery (Jarvis 2008).

The translocon at the outer envelope of chloroplasts, referred to as 'TOC', is made up of five components: Toc159, Toc34, TOC75, TOC64, and TOC12 (fig. 4). TOC159 and TOC34 are related GTPases involved in preprotein recognition and are thus considered receptor proteins (Jarvis 2008). TOC159 is held in the outer envelope by C-terminal membrane anchoring domain (the M-domain), and projects a central GTP-binding domain (G-domain) and a highly acidic domain (A-domain) toward the cytosol (Andres et al. 2010; Jarvis 2008). TOC34 is anchored in the outer membrane by a short hydrophobic sequence near the C-terminus, and has a G-domain that also faces the cytosol. The two receptor proteins, together with TOC75, which oligomerizes to form the translocation pore, make up the 'core complex' (Strittmatter et al. 2010). This complex is responsible for recognition of the preprotein at the chloroplast surface and outer envelope translocation (Jarvis 2008). Finally, phosphorylation of TOC34 and TOC159 may inhibit the association of the TOC complex, or elicit its dissociation, though as the phosphorylation site is at the interface of dimerization, the former seems more likely (Oreb et al. 2008).

TOC12 is involved in tethering an intermembrane space HSP70 to the exit site of the TOC complex (Andres et al. 2010), and with this chaperone, plus TOC64 and TIC22, it forms the intermembrane space complex, which is thought to aid transfer of preproteins between the two translocon complexes (Becker et al. 2004a). TOC64 is involved in preprotein recognition, as described below.

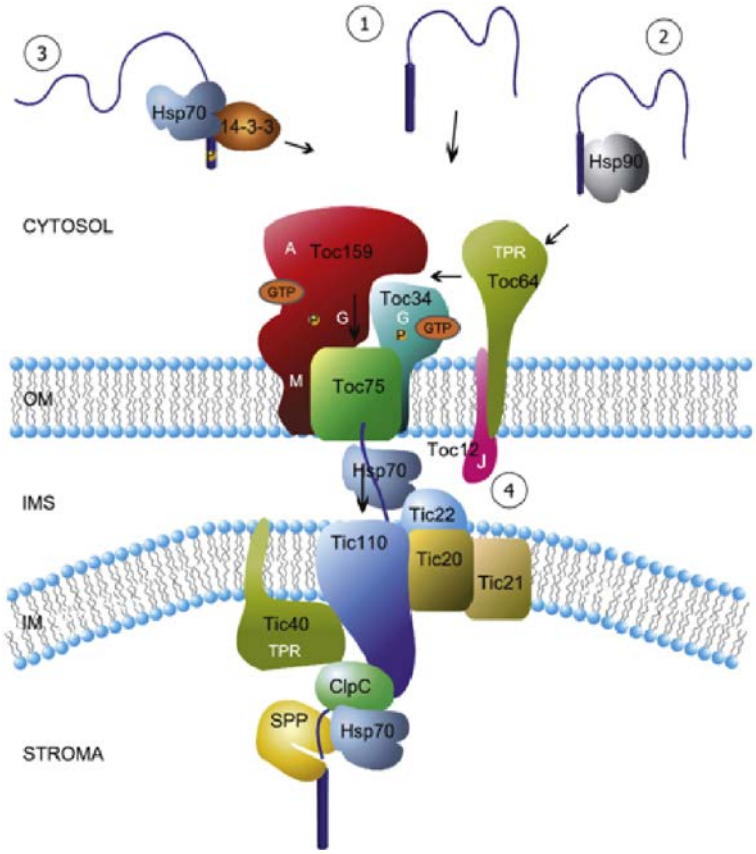


Figure 4 TICTOC import machinery

Graphical representation of the TICTOC protein import machinery showing each of the subunits and their sub-cellular localization. Preproteins destined for import to the chloroplast reach the outer envelope in three different ways: 1. direct binding to the receptors TOC34 and TOC159; 2. non-phosphorylated preproteins use the TOC64/Hsp90 pathway; and 3. pre-proteins with a phosphorylated transit peptide use the guidance complex, comprised of Hsp70 and a 14-3-3 protein

Reproduced with permission from Andres 2008

1.5.1.1 Import of photosynthetic vs. non-photosynthetic prepeptides

There is much controversy surrounding TOC159, which was initially isolated from pea chloroplasts as an 86 kDa protein, and thus assigned the name TOC86 (Hirsch et al. 1994) (Seedorf et al. 1995) (Kessler et al. 1994) (Perry and Keegstra 1994). The 86 kDa fragment was later found to be a fragment of the larger

TOC159 protein, with the acidic A-domain missing (Bolter et al. 1998; Chen and Li 2007).

Additionally, a soluble form of TOC159 has been reported to exist in abundance equal to the membrane bound form (Hiltbrunner et al. 2001; Ivanova et al. 2004), and, in the 'targeting model' of TOC receptor function, this form was proposed as having the role of recruiting prepeptides in the cytosol to the TOC complex at the chloroplastic outer envelope (Hiltbrunner et al. 2001). In this model, the TOC159-preprotein complex then docks at and forms a heterodimer with TOC34 (fig. 4), and upon GTP hydrolysis, the preprotein is transferred to TOC34. An HSP70 protein located in the intermembrane space drives subsequent translocation across the plastid membrane (Becker et al. 2004a).

Interestingly, separate transport complexes may exist for photosynthetic, as opposed to housekeeping, proteins. In *Arabidopsis*, three additional forms of TOC159 have been identified: TOC90, TOC120, and TOC132 (Fulgosi and Soll 2002; Kubis et al. 2003). Analysis of the *Arabidopsis* mutant *ppi2*, which lacks TOC159, demonstrated that while photosynthesis related proteins were transcriptionally down-regulated, housekeeping proteins accumulated normally (Bauer et al. 2000). This finding led to the supposition that TOC159 is specific as an importer of photosynthetic proteins, and that other receptors are responsible for the import of housekeeping proteins, namely TOC120 and TOC132 (Bauer et al. 2000; Jarvis 2008). More recent work supports this idea; specifically, that *Arabidopsis* TOC120 and TOC132 are present in import complexes excluding TOC159, and that import complexes containing atTOC159 (*Arabidopsis thaliana* TOC159) do not include atTOC120 or atTOC132. Moreover, atTOC33, an

isoform of TOC34, is more commonly found in atTOC159 complexes, while atTOC34 is the predominant form found in atTOC132/atTOC120 complexes (Ivanova et al. 2004). Finally, the sequence motif for a photosynthetic transit peptide interaction specifically with TOC159 has been identified: serine residues in the N-terminal 12 amino acids of the RbcS transit peptide are crucial for TOC159 dependent import of the preprotein (Lee et al. 2009).

1.5.1.2 Docking of the preprotein at the chloroplastic outer membrane

Interaction of the preprotein with cytosolic factors and subsequent docking and detection of the preprotein is relatively poorly understood. Currently, two models exist for TOC receptor function; these explain translocation of phosphorylated and un-phosphorylated prepeptides.

a. Interaction of transit peptides with cytosolic factors

Since protein import to the chloroplast is a post-translational process, it seems reasonable that soluble, cytosolic factors act to aid the passage of the preprotein from the 80s ribosome to the plastid outer membrane (Hiltbrunner et al. 2001; Jackson-Constan et al. 2001). Additionally, as proteins are imported in an unfolded state, the same cytosolic factors may function to maintain the proteins in an unfolded conformation, though the chloroplast does have the ability to unfold and import pre-folded precursor proteins (America et al. 1994; Della-Cioppa et al. 1986).

Transit peptides have little similarity at the primary sequence level, however, all contain a majority of positively charged and hydroxylated amino acids, such as serine and threonine (Cline 2000); phosphoserine and phosphothreonine often

function as the central residue in binding sites for 14-3-3 proteins (Jarvis and Robinson 2004). Serine and threonine have been demonstrated to be targets for phosphorylation by a small protein kinase family with activity limited to chloroplastic targeting signals, excluding proteins targeted to the mitochondrion (Martin et al. 2006; Waegemann and Soll 1996).

In addition, many transit peptides have putative binding sites for HSP70 proteins, and it has been shown that these proteins are able to interact with chloroplast transit peptides (Jackson-Constan et al. 2001). HSP70 proteins have been proposed as participants at numerous points along the import pathway, in addition to their involvement in the folding of newly synthesized proteins (Bukau et al. 2000; Young et al. 2004).

b. Targeting of phosphorylated preproteins

Upon phosphorylation in the cytosol, the preprotein becomes competent for interaction with the guidance complex, comprised of HSP70 and a cytosolic 14-3-3 protein. Binding of the preprotein to the guidance complex has been demonstrated to increase the import efficiency (May and Soll 2000). At the outer membrane of chloroplasts, TOC34 is an initial receptor for preproteins delivered by the guidance complex or for preproteins in monomeric form (Becker et al. 2004b), and has been shown to have a higher affinity to phosphorylated preproteins (Kovacs-Bogdan et al. 2010). It recognizes both the transit peptide and the 14-3-3 protein of the guidance complex (Schleiff et al. 2003; Schleiff et al. 2002), dimerizes with TOC159, and the precursor protein is handed from TOC34 to TOC159 (Jarvis 2008). As TOC159 binds only non-phosphorylated

preproteins (Becker et al. 2004b), the preproteins delivered by the guidance complex must be de-phosphorylated prior to association with this receptor.

c. Targeting of un-phosphorylated preproteins

Un-phosphorylated preproteins are imported to the chloroplast via an association with HSP90, which binds to the TOC core complex using TOC64 as the initial receptor (Andres et al. 2010). In Pea, TOC64 is associated with the TOC core complex intermittently (Schleiff et al. 2003b), and recognizes preproteins associated with chaperones distinct from the guidance complex (i.e. a mechanism other than the guidance complex) (May and Soll 2000; Qbadou et al. 2006). While HSP70 may also be involved, it may not be essential for targeting (Qbadou et al. 2006).

Using a clamp type C-terminal TPR domain, TOC64 interacts with the HSP90 chaperone, but not directly with the preprotein, to build a docking site receiving HSP90 complexed preprotein. Using a region of the TPR distinct from the docking site of HSP90, TOC64 interacts with core complex receptor TOC34, which recognizes the preprotein upon dissociation of the HSP90. When the preprotein is delivered from TOC64 to the core complex, TOC64 dissociates from the core complex (Qbadou et al. 2006).

1.5.1.3 Translocation of the preprotein across the chloroplast envelope

Translocation across the chloroplast envelope is an energetically costly process (Andres et al. 2010) involving at least three biochemically distinct steps (Perry and Keegstra 1994; Schnell and Blobel 1993).

The first step, called ‘energy independent binding’, involves binding of the preprotein at the chloroplast surface. This process is reversible and does not require energy (Kouranov and Schnell 1997; Perry and Keegstra 1994). In the next step, referred to as the ‘early import intermediate’ stage, the preprotein is inserted across the outer membrane (Andres et al. 2010). This stage is irreversible, and requires a concentration of ATP less than or equal to 100 μM (Schnell and Blobel 1993). In fact, in this stage, TOC34 and TOC159 likely use GTP as an energy source, as non-hydrolyzable GTP analogs inhibit the process (Schnell et al. 1994). The final stage, or the ‘late intermediate stage’, requires millimolar (>1 mM) concentrations of ATP in the stroma, and in this step the preprotein is translocated simultaneously across both envelope membranes into the chloroplast stroma, where the transit peptide is removed (Pain and Blobel 1987; Theg et al. 1989).

TOC159 is thought to complex with TOC75 and TOC34 such that four channels, each comprised of one TOC75 and one TOC34, surround the central TOC159 finger like domain (Schleiff et al. 2003b). As an alternative to the targeting model described in section 1.6.1.1, in the motor model (Jarvis 2008), TOC159, acting as a GTP-driven motor, then threads the preproteins into the TOC75 channels in a sewing machine-like fashion (Schleiff et al. 2003).

1.5.2. The TIC complex

The translocon at the inner envelope of chloroplasts, or TIC complex, is responsible for translocation of prepeptides across the inner envelope (Strittmatter et al. 2010). There is less knowledge about this translocon, but to date, eight

proteins have been proposed as being part of the translocation complex. They are: TIC110, TIC20, TIC21, TIC40; components involved in regulation TIC62, TIC55, and TIC32; and finally, TIC22 (Kovacs-Bogdan et al. 2010).

TIC22 is the only member of the complex that resides in the intermembrane space. It is only peripherally associated with the inner membrane and is thought to work in conjunction with the J-domain protein TOC12, TOC64, and an intermembrane space HSP70 to facilitate the passage of preproteins from the TOC complex to the TIC complex (Becker et al. 2004a; Qbadou et al. 2007; Schnell et al. 1994).

1.5.2.1 Pore-forming components

The basic function of the TIC complex is to form an aqueous pore for preprotein conductance (Jarvis 2008); three proteins have been proposed as pore-forming components: TIC110 (Heins et al. 2002; van den Wijngaard and Vredenberg 1999), TIC20 (Chen et al. 2002; Kouranov et al. 1998; Kouranov and Schnell 1997), and TIC21 (Teng et al. 2006). The function of TIC21 was questioned, however, when Duy et al. presented the same protein as a metal permease (Duy et al. 2007). While function of this protein as a translocon component has not been disproved, only Duy et al. (2007) were able to obtain a wild type phenotype in their complementation studies Still, it is unresolved as to whether or not this protein plays a role in preprotein import.

1.5.2.2 The TIC motor

Together with TIC110 and HSP93, TIC40 is a component central to the proposed stromal motor of the TIC complex (Bedard et al. 2007). The motor is thought to drive the translocation of the prepeptide across the inner envelope membrane into the stroma in an ATP dependent manner, during which TIC40 may act as a co-chaperone, by regulating the chaperone HSP93. TIC40 has a large stromal domain and structural and functional similarity to co-chaperones from other organisms called the HIP (HSP70-interacting protein) and HOP (HSP70/HSP90-organizing protein) chaperone family (Bedard et al. 2007; Chou et al. 2006; Stahl et al. 1999). The motor also seems to involve TIC110 and associated chaperone HSP93/CLPC (Heat shock protein 93kDa/Caseinolytic protease, subunit C).

Recently, a model for TIC function was proposed by Chou et al (fig. 5), in which TIC40 uses a TPR domain to bind TIC110, an interaction favored when the transit peptide-binding site of TIC110 is occupied (Inaba et al. 2003). When TIC40 binds to TIC110, two things are allowed to happen: first, the transit peptide is released from TIC110 allowing the stromal peptidase to cleave it from the emerging preprotein and second, the TIC40 Hip/Hop domain is exposed, which causes it to stimulate ATP hydrolysis by HSP93. The resultant energy is used to translocate the prepeptide into the stroma (Chou et al. 2006). In this manner, TIC40 may act as a timing device to coordinate the steps of the latter import stages (Jarvis 2008).

1.5.3. Import into the thylakoids

Both nuclear and chloroplast encoded thylakoid localized proteins require transport into or across the thylakoid membranes to arrive at their functional locations. This process is believed to employ four independent precursor-specific thylakoid transport pathways, characterized as being Spontaneous, Signal

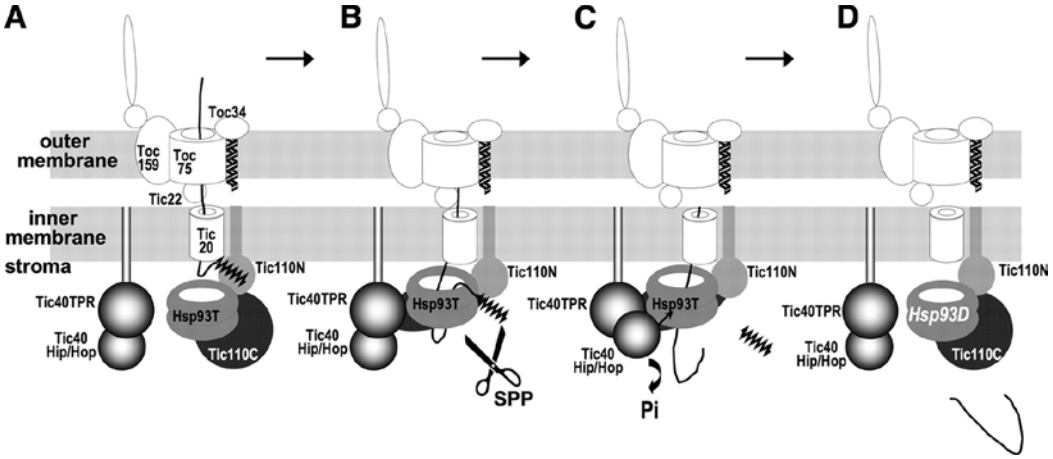


Figure 5. The model for TIC function

The lines moving across the two membrane channels represent a preprotein, with the zigzag line representing the transit peptide. SPP is the stromal processing peptidase; Hsp93T and Hsp93D, are Hsp93 with ATP or ADP bound, respectively. Other translocon components are labeled only in A, for simplicity. A. The transit peptide emerging from the inner envelope membrane channel is bound by the N-terminal TIC110 stromal domain. B. Binding the transit peptide causes a conformation change in TIC110 and recruits TIC40TPR binding. C. Binding of TIC40TPR to TIC110 causes release of the transit peptide, which is then cleaved by the stromal processing peptidase; it also un-shields the TIC40 Hip/Hop domain, which stimulates ATP hydrolysis by Hsp93. D. The energy produced by ATP hydrolysis is then used to translocate the processed mature protein into the stroma

Reproduced from:

Chou et al., 2006. Originally published in The Journal of Cell Biology. doi: 10.1083/jcb.200609172

Recognition Particle (SRP), Secretory (Sec), or Twin-arginine translocase (Tat) dependent (Schunemann 2007). Proteins residing in the Thylakoid Lumen are almost always encoded in the nucleus, and thus require transport across both chloroplast membrane systems (Jarvis and Robinson 2004). These proteins are synthesized with a bipartite presequence that contains two separate targeting

signals in tandem (Hageman et al. 1986; Smeekens et al. 1986) and transported via the ΔpH/Tat and Sec dependent pathways. Proteins localized in the Thylakoid membrane are transported primarily either by the SRP pathway or are spontaneously inserted, with some few reports of membrane proteins using the Tat or Sec dependent pathways (Aldridge et al. 2009; Jarvis 2008).

1.5.3.1 Spontaneous insertion pathway

Some outer envelope proteins can insert spontaneously into the outer envelope. Although the process of insertion is not energy dependent, a step in the insertion was shown to be stimulated by nucleotides (Salomon et al. 1990). The spontaneous insertion pathway has no requirement for nucleotides or import machinery, and is thought to be used by single (thylakoid) membrane spanning proteins, such as the CF_0II subunit of the ATP synthase and Photosystem II subunits PSBW and PSBX, as well as more complex proteins such as PSI constituents PSAK and PSAG (Aldridge et al. 2009). Both of these proteins insert into the membrane with two spans. PSAG has a stroma-exposed loop region connecting the two spanning regions, and positive charges in the loop region were found to be essential for both insertion and function (Zygadlo et al. 2006). In an even more complex process, PSBY undergoes a series of proteolytic events as it is converted into two individual membrane spans (Thompson et al. 1999).

1.5.3.2 The Signal Recognition Particle pathway

Signal recognition particle systems are co-translational and need the presence of the ribosome and a conserved RNA component, although in higher plants, this

RNA component is not always necessary (Li et al. 1995). The SRP mediated chloroplastic import of LHCII has been intensively studied, and this protein has been shown to localize correctly within the thylakoid membrane without the bipartite transit peptide used in the Tat- and Sec-dependent import of other thylakoid localized proteins. Instead, the thylakoid targeting sequence of SRP substrates is included in the mature protein (Viitanen et al. 1988).

1.5.3.3 The Secretory dependent pathway (thylakoid lumenal proteins)

The Sec-dependent pathway for protein secretion is conserved in Gram-negative bacteria and consists of a SecA ATPase and a SecYEG translocon (Natale et al. 2008). The Sec system provides one pathway for transport of lumenal proteins across the thylakoid membrane into the lumen, where they are released following cleavage of the thylakoid targeting signal. The Sec pathway transports proteins in an unfolded state, in an ATP-dependent manner (Albiniak et al. 2012).

1.5.3.4 The Twin-Arginine Translocase dependent pathway

The twin-Arginine translocase dependent pathway is localized to the cytoplasmic membranes in bacteria, and, correspondingly, in the thylakoid membranes of chloroplasts (Robinson and Bolhuis 2004). The most striking feature of this pathway is the fact that it transports pre-folded, and even oligomeric proteins across membranes. In these cases, the transported proteins are usually synthesized with or associated with folded partner proteins that possess distinctive N-terminal signal peptides bearing a common amino-acid sequence motif. This motif (called the ‘twin-arginine-motif) has a consensus sequence of SRRxFLK, where the

arginine residues are highly conserved and essential for efficient protein targeting (Robinson and Bolhuis 2004). In addition, the system seems to have a built-in 'quality control' activity that prevents transport of unfolded polypeptides. This is an innate ability to reject immature or incorrectly assembled proteins (Sargent et al. 2006).

1.6 Redox control of chloroplast protein import

The redox state of the chloroplast has been implicated in regulation of the preprotein import into the chloroplast. This redox control seems to be accomplished by altering the activity or composition of participating transport components, and that through these changes, chloroplasts communicate with other cell compartments (Balsera et al. 2010).

1.6.1. Light regulation of protein import

In a study examining the chloroplastic import of the photosynthetic and non-photosynthetic isoforms of maize ferredoxin (FdxI/FdxIII), and ferredoxin-NADP reductase (FNRI/FNRIII), it was found that light had a strong, specific effect on import of the non-photosynthetic isoforms (Hirohashi et al. 2001). In both light and dark treated chloroplasts, the photosynthetic forms were transported into the stroma, with correct processing of the preprotein and transit peptide cleavage. In the light treated plastids, however, the non-photosynthetic preproteins accumulated in the chloroplast intermembrane space. This phenotype could not be rescued by subsequent transfer into the dark, but if the preproteins were re-

isolated and imported into non-light treated plastids, they were imported and correctly processed (Hirohashi et al. 2001).

1.7 Localization of Magnesium Chelatase: where does catalysis take place?

This enzyme is thought to be localized in the chloroplast envelope (Fuesler et al. 1984b) and/or thylakoid membranes (Larkin et al. 2003). While chlorophyll content is extremely low (0.3 nmol chlorophyll a/mg protein), envelope membranes contain low amounts of chlorophyll precursors, specifically Chlide and Pchlide, in the range of 0.1 to 1.5 nmol/mg protein. When the molar ratios of Pchlide/Chlide to chlorophyll are compared, they are found to be 100 to 1000 times higher in envelope membranes than thylakoid membranes (Pineau et al. 1993)

The localization of *Arabidopsis thaliana* CHLH/ABAR was shown to be dependent upon the Mg^{2+} concentrations of the medium used in chloroplast fractionation. At a concentration of Mg^{2+} above 5 mM, CHLH localizes predominantly to the envelope, while at a lower concentration (1 mM) CHLH is found predominantly in the stroma fraction (Gibson et al. 1996a).

1.8 CHLH, Abscisic Acid, and the Role of Carotenoids

In addition to being a GUN mutant, the CHLH subunit of Mg-chelatase has been reported as an ABA receptor (Shen et al. 2006b). ABA is a phytohormone responsible for the regulation of numerous processes in plant development, including bud dormancy, seed dormancy/germination, stomatal aperture and the

expression of stress response genes (Seo and Koshiba 2002). Moreover, it is the end product in the biosynthetic pathway of carotenoids, which play numerous essential roles in plant biology, among them light harvesting and photo-protection (Milborrow 2001).

The role of photo-protection is a vital one for the health of the plant, due to the generation of singlet oxygen and free radicals generated in high light environments. When absorbed light exceeds the capacity of the photosynthetic apparatus to drive photochemistry, the lifetime of the singlet state of excited chlorophyll is extended. In this case, the excited chlorophyll can return to ground state by the emission of light (chlorophyll fluorescence), heat (thermal dissipation), or by an energy transfer to oxygen resulting in singlet oxygen, a highly reactive oxygen species (Murchie and Niyogi 2010). Singlet oxygen has been shown to be almost exclusively responsible for chloroplast lipid peroxidation in a process called photo-oxidative stress (Triantaphylides et al. 2008). The chloroplast, which is comprised of thylakoid membranes, is particularly at risk due to the increased incidence of photo-sensitizers such as heme, chlorophylls and the photo-reactive tetrapyrrole intermediates including and after Uroporphyrinogen III. To achieve this protective effect, carotenoids act by quenching excess energy absorbed by the light-harvesting complex of photosynthesis.

Carotenoid biosynthesis in chloroplasts seems to be localized to the envelope and in some cases, the thylakoid membrane (Joyard et al. 2009). All known enzymes of this pathway are nuclear-encoded and post-translationally imported into the chloroplast (Conti et al. 2004). The chloroplast signal recognition particle

heterodimer (scSRP54; FFC) and (scSRP43; CHAOS) and its receptor (cpFTSY) are involved in targeting light-harvesting proteins to the thylakoid membranes. In *fcc*/*chaos* and *cpftsy* mutants total carotenoid levels are reduced, but interestingly, there is also a greater than 80% reduction in chlorophyll (Asakura et al. 2004). The reduction in carotenoids may be attributable to the reduction of light harvesting complexes (LHCs), or it could be an effect of retrograde signaling.

Plants employ numerous mechanisms to protect from the absorption of excess light. Some examples include upright cereal leaves, which do not affect the saturation of photosynthesis but do reduce the amount of light absorbed around mid-day, as well as the alignment of chloroplasts along the periclinal walls perpendicular to the incident light.

A recent study by Castells et al. has shown that the TOC1 component of the chloroplast import machinery regulates the diurnal expression of Mg-chelatase H (*A.t* ABAR/GUN5) by direct binding to its promoter. In this study, ABA treatment resulted in TOC1 induction at midday; this induction was demonstrated to control the phase of TOC1 binding and the expression of CHLH. Conversely, plants deficient in CHLH (via RNAi) were not subject to ABA mediated induction of TOC1(Castells et al. 2010).

Characterization of the *A. thaliana* ABAR/CHLH demonstrated that it specifically binds ABA and mediates ABA signaling as a positive regulator in stomatal movement, seed germination, and postgerminative growth (Wang et al. 2011). More recently, CHLH has been identified as a positive regulator of ABA signaling associated with ripening in strawberry fruit. In this study, Jia et al. used an RNAi approach to down-regulate CHLH, which in turn led to a significant

increase in ABA levels; they postulated that this increase may be attributed to a feedback effect on the number of available ABA signal molecules when the ABA signaling pathway is repressed (Jia et al. 2011).

The mechanism for this phenomenon was further elucidated upon study of the Mg-chelatase H subunit of *Arabidopsis*, which was shown to play a role in ABA binding and signaling via its C-terminal half (Wu et al. 2009), and was found to be a key player connecting ABA-mediated plant responses to drought with the circadian clock (Legnaioli et al. 2009). In addition, CHLH was found to antagonize a group of WRKY transcription repressors involved in ABA signaling, interacting with these factors via the cytosolically exposed C-terminal half of the envelope associated form (Shang et al. 2010). WRKY transcription factors serve as negative regulators of ABA signaling in seed germination and post germination growth. In particular, WRKY40 inhibits expression of ABA responsive genes such as *ABI5*. In response to high ABA levels, the WRKY40 is recruited from the nucleus and its association with CHLH is enhanced. In turn, CHLH relieves inhibition of *ABI5* by repressing WRKY40 expression (Shang et al. 2010).

1.9 The study of transiently expressed Magnesium Chelatase subunits *in Planta*

The study of Mg-chelatase subunits expressed *in Planta* may help define the processes and effects specifically of the Mg-chelatase enzyme. Previous studies have used etiolated/developing seedlings, the drawback being that it is difficult to determine that the effects seen are not influenced by changes in the developmental state of the plant, due to the effects of etiolation, or of seedling development.

For example, the function of PORA is confined to the very early stages of transition from etiolated to light growth (Apel 1981; Batschauer and Apel 1984; Forreiter et al. 1990; Mosinger et al. 1985). The amounts of both PORA protein and *PORA* mRNA decrease drastically soon after the beginning of illumination, due in part to rapid proteolytic turnover of the enzyme protein (Reinbothe et al. 1996). Changes in the expression of a protein acting similarly to PORA could be mistakenly attributed to the experimental treatment, if the protein were not well-characterized.

Transient expression of the Mg-chelatase subunits in mature leaf tissue allows the elimination of the developmental effects seen with other systems. In the utilization of this system, we hope to elucidate the effects brought about directly by over-expression of the individual proteins *per se*.

Chapter 2. Cloning of *Zea mays* Magnesium Chelatase subunits

Mg-chelatase is expressed in both greening and mature photosynthetic tissues, but the majority of the plant studies performed to date have been done almost exclusively on seedlings. This is likely due to the fact that maturing tissue has a greatly increased flux in tetrapyrrole synthesis, facilitating the detection of changes between a fully functioning system and one that is impaired. However, the use of seedlings may pose a problem in that the plant is undergoing vast physiological changes to bring about a new developmental state. In these cases, the transition into photosynthetic growth demands a major shift in gene expression and presumably employs a plethora of signaling.

We were interested in studying the effects of over-expression of Mg-chelatase subunits in mature leaf tissue, in order to elucidate the direct effects of these subunits on stable, as opposed to developing, tissue. The use of mature tissue greatly reduces ambiguity regarding changes to the transcriptome or proteome, leading to conclusions that can be directly attributed to overexpression of a particular protein. To this effect, the subunits of Mg- chelatase and GUN4 were cloned and inserted into a binary vector (pBin61) for *in Planta* expression.

The use of a transient expression system in this case is crucial. In contrast to the use of stable transformants, transient expression allows rapid expression of the protein of interest, and eliminates the effects of a change in the developmental state of the plant.

Prior to my PhD candidature, I worked in the laboratory of Thomas Brutnell under the supervision of Dr. Ruraidh Sawers to clone the *Zea mays CHLI* gene and transiently express the protein in *Nicotiana benthamiana* (Sawers et al. 2006a). In this work, we demonstrated the efficacy of *Agrobacterium* mediated expression of the *CHLI* gene, showing that the *Zea mays* I-subunit of Mg-chelatase expressed in *Nicotiana benthamiana* is capable of interacting with endogenous subunits D and H. This interaction is best demonstrated by expression of the semi-dominant mutant *Oil yellow 1* (Oy1/CHLI) construct, which shows a clear chlorotic phenotype (Sawers et al. 2006b). The chlorotic phenotype indicates two things. First, that the transiently expressed protein, with a *Zea mays* transit peptide, is correctly imported into the *N. benthamiana* chloroplast, and second, that the transiently expressed protein is capable of interacting with endogenous proteins. If either of these were untrue, expression of the mutant construct would result in a wild type phenotype, since in that case the plant would rely solely on endogenous proteins. Furthermore, infiltration with the wild type construct of CHLI leads to a green phenotype, indicating that the transiently expressed protein does not impair enzymatic function.

This chapter describes the cloning and expression of the remaining subunits of Mg-chelatase, which builds on the preliminary work done with *Zea mays* CHLI.

2.1 Materials and Methods

2.1.1. RNA isolation and cDNA synthesis

RNA was prepared from *Zea mays* cultivar T43 as described previously (Sawers et al. 2006a). Approximately 80 – 90 mg of tissue was ground in liquid nitrogen prior to RNA extraction using the RNeasy Plant Mini kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. DNAse treatment of the RNA was accomplished on the column during the RNA extraction using the RNase-Free DNase Set (Qiagen). Reverse transcription of RNA into cDNA was accomplished using the SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen, Carlsbad, CA). Oligo dT primers were used to obtain total cDNA, while gene specific cDNA was obtained using primers specific for CHLD or CHLH.

2.1.2. Primer design

Primer design was accomplished using the Primer3 program (Rozen and Skaletsky 2000), with a target melting temperature (Tm) of between 68 and 73°C. The high melting temperature was desirable to increase target specificity as well as in some cases employing the primer in a two-step amplification program, which excludes a separate annealing step. In addition, the primers were designed with an optimal G-C content of 50%, and an optimal length of 24 bases. As discussed in individual results sections, some primers included sequence to incorporate a restriction enzyme cutsite or non-contiguous DNA in the amplicon. A complete list of primers can be found in Appendix A.

2.1.3. Polymerase Chain Reaction

Because we were interested in expressing the proteins, it was necessary to perform PCR amplifications resulting in as few errors as was possible, to avoid any changes in the resultant amino acid sequence. We therefore used primarily high fidelity polymerases, most of which produce a blunt ended amplicon. Since blunt ended cloning is more difficult than sticky-ended cloning and the vectors are typically more expensive, we chose to add A's at the end of the reaction and clone the amplicons into pGemT Easy vector to facilitate sequence analysis. The A's were added by simply pausing the reaction after cycling and adding 0.3 µl GoTaq polymerase (Promega; Madison, WI, USA), before proceeding to a final extension step of approximately 10 minutes at 72°C.

Typical reaction programs included a modified touchdown stage, where the target DNA template was enriched by annealing at the estimated melting temperature, then dropping the annealing temperature one degree each cycle. During this stage, annealing of the primer to the target DNA would have to be exact, and thus the likelihood of non-specific annealing is reduced. After the initial touchdown-enrichment, the reaction was programmed to proceed as in a normal PCR, with an annealing temperature approximately 5°C below the estimated melting temperature. This style of cycling was primarily used with hot-start polymerase mixes, as they tend to be more heat-stable than standard polymerases.

2.1.3.1 Polymerases

Amplifications were accomplished using a number of different polymerases, but mainly Novagen KOD DNA polymerase (formerly KOD HiFi DNA Polymerase)

from EMD4Biosciences (Damstadt, Germany), TripleMaster PCR system (Eppendorf; Hamburg, Germany), FastStart High Fidelity PCR System and Expand High Fidelity PCR System, both from Roche Applied Science (Penzberg, Germany). All RACE reactions employed the Invitrogen (Carlsbad, CA, USA) GeneRacer kit.

2.1.3.2 PCR Reaction Additives

In reactions where amplification was exceptionally recalcitrant, we used 5% v/v DMSO as an additive in the PCR reaction. DMSO has been shown to reduce non-specific priming by changing the Tm of primer-template hybridization. Furthermore, DMSO reduces secondary structure, and facilitates strand separation (Filichkin and Gelvin 1992; Masoud et al. 1992; Pomp and Medrano 1991).

2.1.4. Agarose Gel purification

To reduce cloning of non-target amplicons, all PCR reactions were run out on agarose gels (0.8-1.0% w/v) containing either ethidium bromide (0.5 µg/ml) or Gel Red (Biotium, Haward, CA, USA) according to the manufacturer's recommendation in TAE buffer (40 mM Tris-acetate (pH 8.0), 2 mM EDTA). The correctly sized band was excised from the gel and processed using the QIAquick Gel extraction kit (Qiagen; Hilden, Germany), according to the manufacturer's recommendations.

2.1.5. Cloning into pGem-T Easy vector and *E. coli* transformation

Cloning of amplified products into the pGem-T Easy vector (fig. 6) was accomplished using the pGem-T Easy vector system. The amplicon A-overhangs were allowed to anneal to the vector's T-overhangs, and the kit T4-DNA ligase was used to join the ends in a reaction set up according to the manufacturer's recommendations.

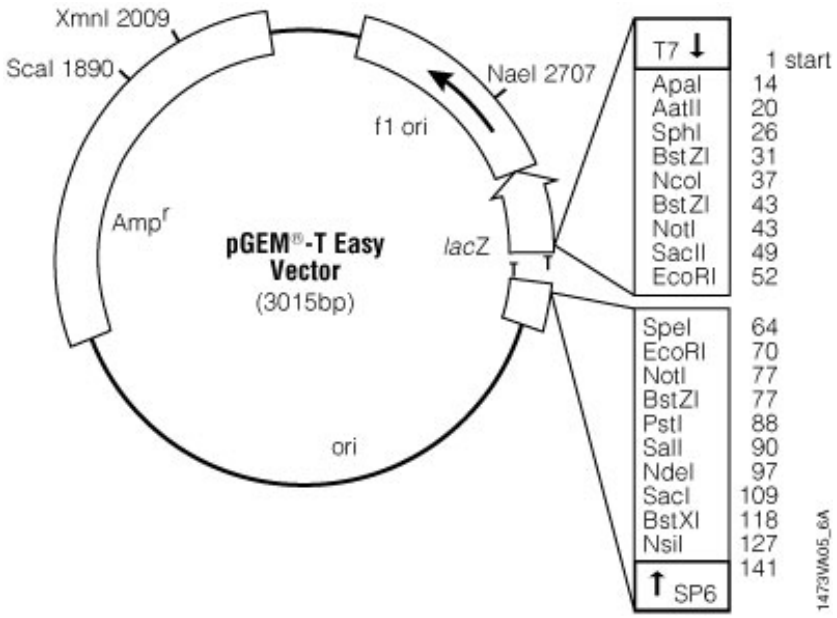


Figure 6. pGEM-T Easy vector map

The figure shows a map of the cloning vector obtained from:

http://www.promega.com/resources/product-guides-and-selectors/protocols-and-applications-guide/cloning/

Transformations of *Escherichia coli* were accomplished chemically. In this process, frozen competent *E. coli* cell aliquots were carefully thawed on ice, and the ligation reaction was kept in the same ice bucket to allow the temperature of the ligation to equilibrate with that of the cells. Next, 2 µl of ligation reaction was added to a 50 µl competent cell aliquot, and the transformation mix was incubated

on ice for 30 minutes prior to a heat shock (30-45-second water bath incubation at 42°C). To each transformation, 125 µl SOC medium (2% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 8.56 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20mM D-glucose) was then added, and the cells were cultured at 37°C with shaking for one hour. Finally, the cells were plated on LB-Agar (1.0% Tryptone (w/v), 0.5% Yeast extract (w/v), 0.5% sodium chloride (w/v), and 1.5% BactoAgar (w/v)) with ampicillin added for selection of cells carrying the pGEM-T easy vector and 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-gal) added to facilitate blue/white selection. The gene for ampicillin resistance is incorporated on the pGEM-T Easy vector; therefore cells successfully transformed with the vector will have resistance to ampicillin. The blue/white selection is accomplished using the β−galactosidase gene, which spans the vector's multiple cloning site (MCS), and functions to cleave the β-glycosidic bond in X-gal, resulting in blue bacterial colonies. An insertion into the MCS causes a disruption in this gene; thus, bacterial colonies transformed with the cloned insert are white and can be grown on ampicillin, while colonies with empty vector are blue.

2.1.6. Sequence Analysis

Nucleotide sequence files (AB1 files) were analyzed using the Sequencher program (Gene Codes; Ann Arbor, MI, USA). Using this program, individual sequence files were assembled into contiguous sequences, and the amino acid translations were compared to consensus protein sequences of CHLD and CHLH. To obtain the consensus protein sequences, multiple sequence alignments of

orthologous CHLD and CHLH proteins from other organisms were generated using ClustalW (Larkin et al. 2007).

2.2 Results

2.2.1. Cloning of Magnesium chelatase D

The Mg-chelatase D subunit was cloned in a two-part process, with the 5' and 3' parts of the gene cloned separately and subsequently fused in a separate PCR amplification reaction. This strategy was used due to the difficulty in finding a full length *Zea mays* mRNA sequence with homology to a known CHLD query sequence.

2.2.1.1 RT-PCR of CHLD 3' Fragment

To identify mRNA encoding an amino acid sequence orthologous to the Mg-chelatase D query sequence from *Hordeum vulgare* (Xantha G, accession AJ011926), a TBLASTX search of the NCBI *Zea mays* expressed sequence tags (EST) database was performed. This type of Blast search is useful when searching for homologous/orthologous protein sequences, without necessarily maintaining high nucleotide sequence similarity. The program translates both the query sequence and database sequences in all six frames, and compares the translated sequences to identify homologous regions.

Several accessions were identified with high similarity to the query sequence, however, most were between 550 and 1000 bp, considerably shorter than the expected size (~2.4 kb). To identify longer *Zea mays* accessions, the query sequence was limited to the portion of the gene with high sequence identity. Upon

identification of a 2kb putative *Zea mays* partial *CHLD* gene (accession #AY109815.1) the Primer3 program (Rozen and Skaletsky 2000) was used to design primers for use in RT-PCR.

The initial 2016 bp gene fragment (depicted in fig. 7 as 2 kb) was PCR amplified using KOD polymerase (Damstadt, Germany), with A-overhangs added at the end of the PCR amplification to facilitate cloning into the pGem T Easy vector (fig. 7). A-overhangs were incorporated on the amplicon ends by adding Promega brand Taq polymerase (Madison, USA) after the completion of cycling, and incubating the reaction at 72ºC for 30 minutes without a melting step.

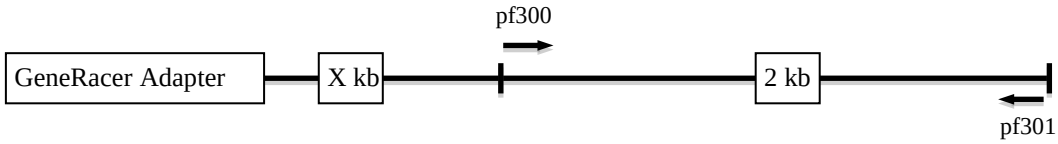


Figure 7. Cloning of magnesium chelatase D 2kb gene fragment

The D subunit was cloned in two stages. The initial stage was comprised of amplification of a 2016 bp 3' fragment obtained using primers pf300 and pf301. The 5' end of the gene was obtained using 5'RACE, with an amplicon of unknown size (X kb).

The final product was gel purified, cloned into Promega vector p-GemT Easy utilizing the vector's T-overhangs, and the resultant plasmid DNA was used to chemically transform Invitrogen Top10 *E.coli* cells (Top10 cells). Sequence verification was performed using Sanger sequencing and gene specific primers (see Appendix 1 for *CHLD* sequencing primer details). Because the *Zea mays* sequence for the putative *CHLD* did not encompass the full length of the gene, the 5' end of the mRNA sequence was obtained using 5'Rapid Amplification of cDNA ends (5' RACE).

2.2.1.2 5'RACE of magnesium chelatase D

The 5' end of the *CHLD* gene was obtained using the 5' RACE GeneRacer kit from Invitrogen (Grand Island, USA). In the first step of this three stage process, the gene specific cDNA was reverse transcribed from *Zea mays* cultivar T43 RNA using Superscript III reverse transcriptase (Invitrogen) and the gene specific pf301 primer along with the GeneRacer 5' primer (fig. 8A). Next, the first round of PCR amplification was done using the gene specific cDNA as template and the primers

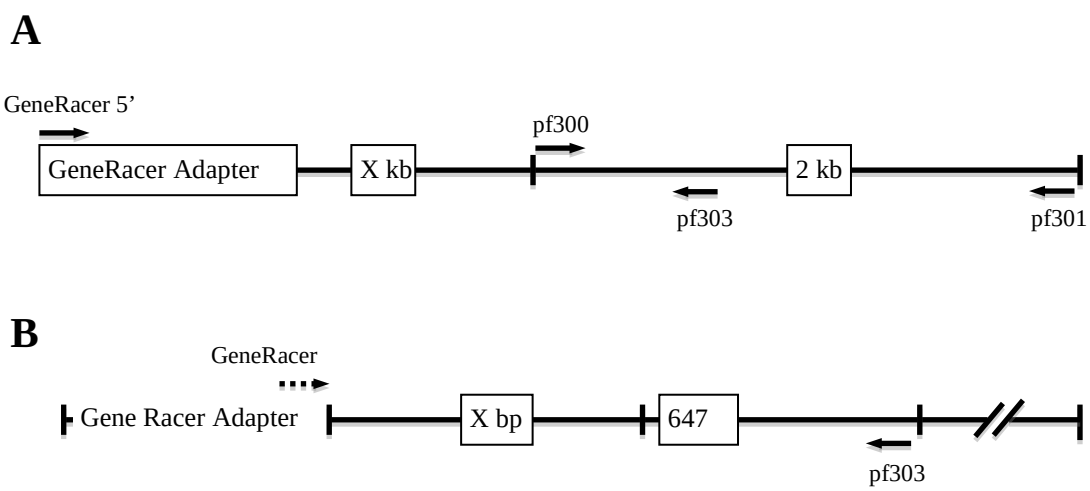


Figure 8. Cloning of magnesium chelatase D 5' end using RACE

The 5' end of Mg-chelatase D was cloned using the GeneRacer kit from Invitrogen. In this process, the fragment was cloned using a two stage fully nested PCR scheme. ***A.*** *In the first stage, the unknown region of 'X bp' was cloned using the GeneRacer 5' primer coupled with the gene specific pf301 primer. The template used in this stage was gene specific cDNA from Zea mays T43.****B.*** *The second stage, done in a separate PCR reaction, used the GeneRacer 5'nested primer with gene specific pf303. In this figure, the GeneRacer 5' and 5' nested primers are depicted above the GeneRacer adapter in solid and dashed arrows, respectively.*

pf301 and GeneRacer 5' primer. As these primers were used to create the cDNA, the amplicon length did not change in this step. In the second round of PCR, the amplification was done in a fully nested manner, using the product from round 1

as template and primers pf303 along with GeneRacer 5'-nested. The round 2 product included the 647 bp of known sequence as well as approximately 400 bp of unknown 5' sequence (fig. 8B), with 238 bases being the 5' end of the gene (from the end of the known 647 bp sequence to the ATG start codon) and the remaining sequence 5' UTR (un-translated region).

2.2.1.3 Fusion of magnesium chelatase D 5' and 3' fragments

To obtain the full-length *CHLD* clone, PCR amplification was performed using the two fragments as template: the 5' RACE product and the 3' 2 kb gene fragment from RT-PCR of *Zea mays* mRNA. The two templates had an overlapping region that would anneal after the melting stage of the program, allowing the polymerase to use the resultant double stranded regions in place of primers (fig. 9). The end primers used were the 5' GeneRacer nested primer and pf301, and the amplification was performed using Roche Expand High-Fidelity

Figure 9. Fusion of magnesium chelatase D 5' and 3' fragments

The 5' RACE product and 2 kb fragments of Mg-chelatase D were fused in a PCR reaction employing only the GeneRacer 5'nested primer and pf301. Primers pf300 and pf303 were not included in the reaction, but show the overlapping region between the two templates. This overlap eliminated the need for primers in this region, as annealing of the overlapping strands formed a double strand of DNA, allowing extension by the polymerase.

polymerase. A-overhangs were added using the same method used for the 2 kb *CHLD* fragment (see above). The 2274 bp product was cloned into the pGem T Easy vector and sequence verified using Sanger sequencing.

Of the twelve clones analyzed, eight had a mutation(s) resulting in a change to the amino acid sequence, or that resulted in a frame-shift. Four clones were identified as having a single point mutation each, resulting in each case in a silent mutation. Given that the protein products would be identical in sequence to the wild type protein, one of these was chosen randomly for use in *in Planta* expression experiments (clone CHLD 2-2) (fig. 10).

```
1        ATGGCGACGCCCACC GCGCTCTCCACCTCA CTCCCCTACCTCCCG CCCCGCCGCGTCATC TCATTCCCATCCGCC
1         M  A  T  P  T   A  L  S  T  S   L  P  Y  L  P   P  R  R  V  I   S  F  P  S  A
76       GCCGCCGTCTCCCTC CCCGTCACCTCCCGC CCCGCCCGGCTGCGG GATTCCCGCCTCGCG GCCGCGGCAACCTCG
26        A  A  V  S  L   P  V  T  S  R   P  A  R  L  R   D  S  R  L  A   A  A  A  T  S
151      GCCTCCGAGGTCCTC GAGTCCACCAACGGC GCCGTCCCCACTGCG GCCAAGGGCGGCGCG TGGCGCGGGTATGGG
51        A  S  E  V  L   E  S  T  N  G   A  V  P  T  A   A  K  G  G  A   W  R  G  Y  G
226      AGGGAGTACTTCCCC CTGGCTGCCGTCGTT GGGCAGGATGCTATC AAAACTGCTCTGCTG CTTGGGGCGATTGAT
76        R  E  Y  F  P   L  A  A  V  V   G  Q  D  A  I   K  T  A  L  L   L  G  A  I  D
301      CGTGAGATCGGAGGC ATTGCCATCTCAGGG AAGCGTGGGACGGCA AAGACAGTGATGGCT CGTGGTTTGCATGCT
101       R  E  I  G  G   I  A  I  S  G   K  R  G  T  A   K  T  V  M  A   R  G  L  H  A
376      ATGCTTCCACCCATT GAAGTGGTGGTTGGT TCCATTGCAAATGCT GACCCTAACTCCCCT GACGAATGGGAGGAT
126       M  L  P  P  I   E  V  V  V  G   S  I  A  N  A   D  P  N  S  P   D  E  W  E  D
451      GGTTTAGCTGATCAA ATACAGTATGACTCT GATGGTAATGTCAAA TCCGAGATCGTCAAA ACACCTTTTGTGCAG
151       G  L  A  D  Q   I  Q  Y  D  S   D  G  N  V  K   S  E  I  V  K   T  P  F  V  Q
526      ATTCCACTTGGTGTG ACGGAGGATAGGCTC ATTGGATCAGTTGAT GTTGAAGCATCTGTG AGATCAGGGACTACT
176       I  P  L  G  V   T  E  D  R  L   I  G  S  V  D   V  E  A  S  V   R  S  G  T  T
601      GTATTTCAACCTGGT CTTCTTGCTGAAGCA CATAGAGGTGTTCTT TATGTTGATGAAATA AATCTATTGGATGAT
201       V  F  Q  P  G   L  L  A  E  A   H  R  G  V  L   Y  V  D  E  I   N  L  L  D  D
676      GGCATAAGCAATCTA CTTCTGAATGTCTTG ACGGAGGGAGTTAAC ATTGTGGAAAGAGAG GGCATTAGCTTTCGC
226       G  I  S  N  L   L  L  N  V  L   T  E  G  V  N   I  V  E  R  E   G  I  S  F  R
751      CATCCCTGCAAACCA CTTCTAATTGCTACT TACAATCCAGAGGAA GGATCTGTACGTGAA CACTTGCTTGATCGT
251       H  P  C  K  P   L  L  I  A  T   Y  N  P  E  E   G  S  V  R  E   H  L  L  D  R
826      ATTGCAATTAATTTA AGTGCTGATCTTCCA ATGAGTTTTGATGAC CGCGTTGAAGCAGTG GATATTGCAACACGG
276       I  A  I  N  L   S  A  D  L  P   M  S  F  D  D   R  V  E  A  V   D  I  A  T  R
901      TTTCAGGAGTCTAGC AAAGAAGTTTTCAAA TTGGTGGAAGAAAAA ACTGAAACTGCAAAA ACTCAGATAATTTTT
301       F  Q  E  S  S   K  E  V  F  K   L  V  E  E  K   T  E  T  A  K   T  Q  I  I  F
976      GCAAGAGAGTATCTG AAGGATGTTACTATT AGCACAGAGCAGCTC AAATATCTTGTCATG GAAGCTATACGAGGT
326       A  R  E  Y  L   K  D  V  T  I   S  T  E  Q  L   K  Y  L  V  M   E  A  I  R  G
1051     GGCTGTCAGGGGCAT CGTGCTGAGTTGTAT GCTGCCCGAGTTGCA AAATGTCTAGCTGCT ATGGAAGGACGTGAA
351       G  C  Q  G  H   R  A  E  L  Y   A  A  R  V  A   K  C  L  A  A   M  E  G  R  E
1126     AAAGTATTTGTGGAT GACCTCAAGAAAGCT GTAGAGCTGGTCATT CTACCTCGCTCCATC CTATCTGATAATCCA
376       K  V  F  V  D   D  L  K  K  A   V  E  L  V  I   L  P  R  S  I   L  S  D  N  P
1201     CAGGATCAGCAGCAA GAGCAACCACCCCCA CCCCCGCCGCCACCA CCTCCAGAAAATCAA GATTCTTCAGAAGAC
401       Q  D  Q  Q  Q   E  Q  P  P  P   P  P  P  P  P   P  P  E  N  Q   D  S  S  E  D
1276     CAAGATGAGGAAGAC GAAGACCAAGAGGAT GATGAAGAAGAAAAT GAACAACAAGACCAA CAGATACCTGAGGAG
426       Q  D  E  E  D   E  D  Q  E  D   D  E  E  E  N   E  Q  Q  D  Q   Q  I  P  E  E
1351     TTCATTTTTGATGCT GAAGGTGGTTTAGTA GATGACAAACTTCTT TTCTTTGCCCAGCAA GCACAGAGACGACGT
451       F  I  F  D  A   E  G  G  L  V   D  D  K  L  L   F  F  A  Q  Q   A  Q  R  R  R
1426     GGAAAAGCTGGGCGA GCAAAGAATGTCATC TTCTCAGAAGATAGG GGCCGTTACATAAAG CCTATGCTTCCTAAG
476       G  K  A  G  R   A  K  N  V  I   F  S  E  D  R   G  R  Y  I  K   P  M  L  P  K
1501     GGTCCAGTAAGGAGG TTAGCTGTTGATGCC ACGCTTAGAGCAGCT GCACCATACCAAAAA CTGCGCAGAGAGAAA
501       G  P  V  R  R   L  A  V  D  A   T  L  R  A  A   A  P  Y  Q  K   L  R  R  E  K
1576     GAACGTGACAAAACA AGAAAGGTTTTTGTT GAAAAGACTGACATG AGAGCCAAAAGAATG GCTCGAAAAGCAGGT
526       E  R  D  K  T   R  K  V  F  V   E  K  T  D  M   R  A  K  R  M   A  R  K  A  G
1651     GCTCTAGTCATATTT GTTGTGGACGCTAGT GGTAGCATGGCTCTG AATCGTATGCAGAAT GCTAAAGGTGCGGCG
551       A  L  V  I  F   V  V  D  A  S   G  S  M  A  L   N  R  M  Q  N   A  K  G  A  A
1726     TTGAAGTTGCTTGCA GAAAGCTACACCAGC AGAGATCAGGTTTCA ATTATTCCTTTTCGT GGAGATTATGCTGAG
576       L  K  L  L  A   E  S  Y  T  S   R  D  Q  V  S   I  I  P  F  R   G  D  Y  A  E
1801     GTTTTGCTTCCACCA TCAAGATCTATAGCA ATGGCCCGGAAACGT CTTGAGAAGCTACCA TGTGGTGGTGGTTCT
601       V  L  L  P  P   S  R  S  I  A   M  A  R  K  R   L  E  K  L  P   C  G  G  G  S
1876     CCTTTAGCTCATGGC CTAAGTACAGCTGTC AGAGTGGGTCTGAAT GCTGAAAAGAGTGGC GATGTTGGGCGTATC
626       P  L  A  H  G   L  S  T  A  V   R  V  G  L  N   A  E  K  S  G   D  V  G  R  I
1951     ATGATTGTTGCAATC ACCGATGGAAGAGCT AATGTATCACTGAAG AGATCCACTGACCCA GAAGCTGCTGCTGCT
651       M  I  V  A  I   T  D  G  R  A   N  V  S  L  K   R  S  T  D  P   E  A  A  A  A
2026     TCAGATGCACCAAGA CCTTCTTCTCAAGAA CTGAAGGACGAGATA CTTGAGGTGGCTGGC AAAATATACAAGGCA
676       S  D  A  P  R   P  S  S  Q  E   L  K  D  E  I   L  E  V  A  G   K  I  Y  K  A
2101     GGAATGTCCCTTCTT GTCATCGACACTGAG AACAAGTTTGTATCC ACGGGATTTGCCAAG GAAATTGCAAGGGTT
701       G  M  S  L  L   V  I  D  T  E   N  K  F  V  S   T  G  F  A  K   E  I  A  R  V
2176     GCCCAGGGGAAATAT TATTACCTCCCTAAT GCTTCAGATGCTGTA ATTTCTGCTGCCACC AAGACCGCCCTGACA
726       A  Q  G  K  Y   Y  Y  L  P  N   A  S  D  A  V   I  S  A  A  T   K  T  A  L  T
2251     GACTTGAAGAGCTCA AATCACTAG
751       D  L  K  S  S   N  H  *
```

Figure 10. *Zea mays* magnesium chelatase D sequence

Sequence for clone CHLD 2-2. Each row of nucleotide sequence is immediately followed by a row of predicted protein sequence based on the standard genetic code. Nucleotide/amino acid numbering is shown to the left of each row, and represents the position of the first character in that row. The silent mutation for this clone, the change of a guanine to adenine, occurs at position 810 (shown above highlighted in black), both resulting in a glutamic acid (E) in the amino acid sequence.

2.2.1.4 Sub-cloning into the pBin-61 binary vector

In the final step of this cloning strategy, cut-sites for the restriction enzymes *Xba*1 and *Bam*H1 were engineered onto the 5' and 3' ends of the gene, respectively, in order to clone the construct into the pBin-61 binary vector. Simultaneously, the 3' primer (pf331) was engineered to ablate the stop codon at the end of the gene, in order to make use of the tags included in pBin-61. The 5' primer pf320 was then used along with pf331 and Roche Fast Start High Fidelity polymerase to amplify the full-length sequence. Sequence analysis was performed to verify the integrity of the translated sequence (fig. 10). Protein prediction of the nucleotide sequence is depicted in figure 11, which shows the predicted motifs.

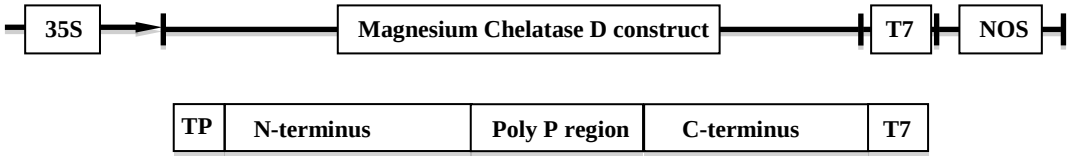


Figure 11. The Magnesium Chelatase D construct in pBin-61 vector

Expression of the magnesium chelatase D protein was driven by the 35S cauliflower mosaic virus promoter (35S), and terminated using the Nopaline Synthase terminator (NOS). The construct was comprised of a transit peptide (TP), the 5' end of the gene (N-terminus), the Poly-P region (Poly P), the 3'end of the gene (C-terminus), and a T7 tag

2.2.2. Cloning of Magnesium chelatase H

The Mg-chelatase H-subunit was cloned using a strategy similar to that used to clone the first two subunits, with some adjustments. Initially PCR of this sequence was attempted to obtain the clone in two fragments: amplification of the majority of the 3887 bp gene fragment identified as a putative *CHLH* (target sequence length is 3802 bp), and the 5' RACE product that comprised the 5' end of the gene, as was done with Mg-chelatase D. Amplification of the target sequence proved to be recalcitrant, however, and after repeated attempts at amplification of the 3802 bp fragment (using nested primers pf314, pf316, pf315 and pf317), the cloning strategy was modified to generate four separate fragments. The known target sequence was PCR amplified in 3 overlapping parts, and the 5' RACE fragment obtained separately. Subsequently, the fragments were joined in separate PCR amplification steps.

As done previously, the fragments would then be cloned into the pGemT easy vector and sequenced using Sanger sequencing prior to assembly in a subsequent PCR reaction. In order to facilitate ligation of the product to the T-overhangs of the vector, A's were added to the fragment ends by adding Promega *Taq* polymerase to the reaction prior to a final extension incubation of 15 minutes, without a melting step.

The previously identified *CHLH* sequence from *Hordeum vulgare* (accession number HVU26916) was used as a query sequence in a TBLASTX search, resulting in identification of *Zea mays* accession AY109455.1. This sequence was translated and verified as CHLH sequence upon comparison to known Mg-

chelatase H protein sequences from *Rhodobacter sphaeroides* (*bchH* gene Accession number Q9RFD5), *Nicotiana tabacum* (Mg protoporphyrin IX chelatase Accession number AAB97152), *Synechocystis elongates* (Mg-chelatase subunit H Accession number YP_401154), and *Hordeum vulgare (Xantha F* Accession number AY039003). The 3887 bp gene fragment identified in this search represents the majority of the gene, but includes a portion of the 3' untranslated region (UTR), and omits approximately 1 kb of the 5' end of the gene; the translated portion of the gene included in the target sequence is 3802 bp.

2.2.2.1 Overview of magnesium chelatase H cloning strategy

The *CHLH* gene was cloned in four separate fragments, each with 200-300 bp of overlap onto the next fragment. These overhangs were used to fuse the fragments together in a subsequent PCR reaction. In order from 5' to 3', the first fragment of unknown size would be obtained using RACE. The remaining fragments were obtained using RT-PCR and subsequent amplification of the target region, with the second fragment having an expected size of 1075 bp, the third 1982 bp and the last 1262 bp. (fig. 12). These fragments will be referred to as 5' (RACE), and

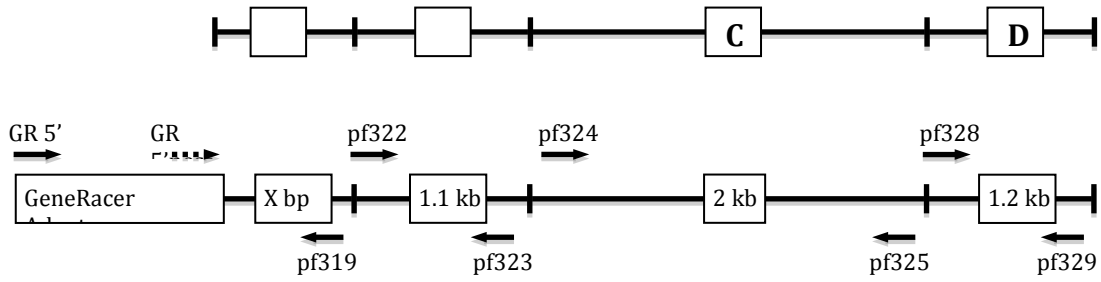


Figure 12. Overview of magnesium chelatase H cloning strategy

The four fragments of CHLH were cloned as A. the 5'RACE fragment of unknown size, utilizing primers GeneRacer 5'nested (GR 5' nest) and pf319; B. fragment 1, 1075 bp, using primers pf322 and pf323; C. fragment 3, 1982 bp in length, used primers pf324 and pf325, and D. the 1262 bp 3' fragment, referred to as fragment 3, used pf328 and pf329.

fragments 1, 2 and 3, respectively.

2.2.2.2 Cloning of the magnesium chelatase H 5'RACE fragment

5' RACE of the H-subunit was performed using the Invitrogen GeneRacer kit. With the GeneRacer adapter ligated to the T43 cDNA template, the first round of PCR employed the GeneRacer 5' primer, coupled with primer pf319, to amplify the 5' end of the gene (fig. 12). Although the size of the 5' end was unknown, based on comparison to orthologous copies of CHLH, the product was expected to be a minimum of 512 bp. The second round of PCR, performed in a half-nested manner (using GeneRacer 5' nested primer and pf319; fig. 12), resulted in a product of approximately 750 bp (fig. 13) (see Appendix A for primer sequences). Upon gel extraction, cloning, and sequencing of this fragment, the length from the ATG start site to the pf319 primer was found to be 538 bp; the remainder of the cloned fragment was 5' UTR, and was omitted from the final construct (fig. 13). Both reactions employed Roche Fast Start High Fidelity polymerase mix

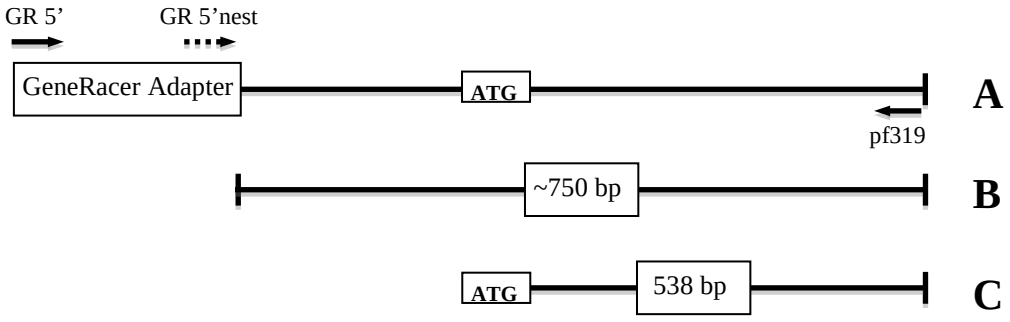


Figure 13. 5' RACE of magnesium chelatase subunit H

A. A portion of the cDNA template is shown, with the location of primers pf319 and GeneRacer primers, the ATG start codon for magnesium chelatase subunit H, and the Generacer Adapter B. The 5'RACE product of ~750 bp, includes 5'UTR upstream of the ATG start codon C. The 538 bp region included in the final construct is shown, including the ATG start codon.

according to the manufacturer's recommendations, with the addition of dimethyl sulfoxide (DMSO).

2.2.2.3 Cloning of magnesium chelatase H fragment 1

Mg-chelatase H fragment 1 proved extremely difficult to amplify, and after multiple attempts using a variety of polymerase mixes, was ultimately accomplished using Roche Fast Start High Fidelity Polymerase with the addition of DMSO, and by employing a gel extraction between the two rounds of RT-PCR, in addition to a long initial melting step (prior to amplification cycling) in each reaction.

The first round of PCR was performed with T43 cDNA ligated to the GeneRacer adapter as template and primers pf322 and pf323 (fig. 12). These primers were designed with very high melting temperatures (74.53°C and 73.61°C respectively) in order to increase specificity to the target cDNA template. In addition, DMSO was included, as this reagent has been shown to enhance the yield and specificity of PCR amplification by facilitating strand separation and disrupting base pairing (Frackman et al. 1998). The polymerase mix manufacturer (Roche) recommended a two to ten-minute incubation at high temperature to activate the enzyme, and because previous attempts at amplification of this section of the cDNA were unsuccessful, the incubation time was increased to the upper limit to reduce any possible secondary structure within this fragment.

While the first round product band was the correct size (1075 bp), when the entire 50 µl reaction was run on an agarose gel, the product band was extremely weak. Preliminary attempts at round-two RT-PCR yielded no product. In order to

eliminate any non-specific products that might be interfering with the reaction by depleting the primer pool, the weak band was gel extracted prior to further amplification. Again, the round two amplification failed; the reaction was only successful when the polymerase was activated with a ten-minute initial heat soak at 97°C, and the gel extracted round 1 product was used as template.

2.2.2.4 Cloning of magnesium chelatase H fragment 2

The cDNA template used to amplify this region was derived from oligo dT primers, and obtained from the *Zea mays* line T43. Primers pf324 and pf325 were employed along with Roche Expand High fidelity polymerase to amplify this region of 1982 bp in a two-step RT-PCR.

2.2.2.5 Cloning of magnesium chelatase H fragment 3

The RT-PCR product for CHLH was used as template to obtain this region of the gene. Using primers pf328 and pf329 (see Appendix A for primer sequences), the 1262 bp product was amplified using Roche Expand High-fidelity polymerase (Penzberg, Germany).

2.2.2.6 Fusion of CHLH and sub-cloning into the pBin61 binary vector

The assembly of full-length Mg-chelatase H construct was attempted several times using the four fragments as template (fig. 12), however the reaction failed to produce a product. The fusion therefore was accomplished by first generating the 5’ half (1425 bp) and 3’ half (2921 bp) of the gene separately, followed by joining of the halves. Once again, fragment one seemed to be the limiting factor, as

amplification of the 5’ half of the gene was far more recalcitrant than that of the 3’ half.

The 5’ half of *CHLH* was amplified using the 5’ RACE product (~750 bp) and fragment one (1075 bp) sequence-verified plasmid DNAs as template. The RACE product at this point still included some 5’ UTR, and the two fragments were designed with an overlapping region to enable annealing of the melted strands, therefore the expected product size was 1425 bp. Eppendorf Triple Master PCR system was used as polymerase, DMSO was included in the reaction, and A’s were added at the end of cycling using Promega Taq polymerase, as described previously.

Preliminary attempts to amplify this region failed, until the ten-minute 97°C incubation prior to amplification was added to the cycling program. Because the Eppendorf polymerase was not a hot-start polymerase, to decrease damage to the enzyme, the reaction was prepared without polymerase, and the enzyme was added after the initial high temperature incubation.

Amplification of the 3’ half of the Mg-chelatase H gene was accomplished in the same manner as the 5’ half, including the long incubation. Although fragment one was the problematic region in the past, because of the greater length of the 3’ templates, the long initial incubation (but not DMSO) was added to this reaction as well, in order to circumvent any issues with the possible formation of secondary structure.

The templates used for this reaction were fragment 2 (1982 kb) and fragment 3 (1262 kb), with an overlapping region of about 160 bp for an expected product size of 2921 bp. Promega Taq polymerase was added after cycling to incorporate

A's on the product ends, and both 3' and 5' halves were cloned into pGem T Easy vector (Promega) and sequence verified.

The reaction to fuse the two halves of *CHLH* also served to incorporate *Xba*1 and *Bam*H1 on the ends of the construct, in order to facilitate cloning into the pBin61 binary vector. This was accomplished using primers pf332 and pf333, along with the Eppendorf Triplemaster PCR system, to generate a product of 4143 base pairs. Promega Taq polymerase was once again employed to add A's to the ends of the amplicon, which was gel extracted, cloned into pGemT Easy vector, and sequence verified using Sanger sequencing prior to sub-cloning into the pBin61 expression vector.

Sequence verification of the full-length construct was accomplished by first performing restriction enzyme digests to identify plasmids with the correct insert size. Twelve clones were identified and partially sequenced using vector primers T7 (pf403) and SP6 (pf404) (see Appendix A for primer sequences). Of these, three clones were verified as having no errors in the 5' and 3' ends, and were sequenced further using gene specific primers pf326, pf345, pf347 and pf349 (see Appendix A for primer sequences).

Two clones were verified as error free when compared to the consensus sequence from an amino acid alignment (see section 2.1.2 for alignment details; figs. 14,15). These clones were double digested with restriction enzymes *Xba*1 and *Bam*H1, and ligated to pBin61-Myc vector digested with the same enzymes (fig. 14).

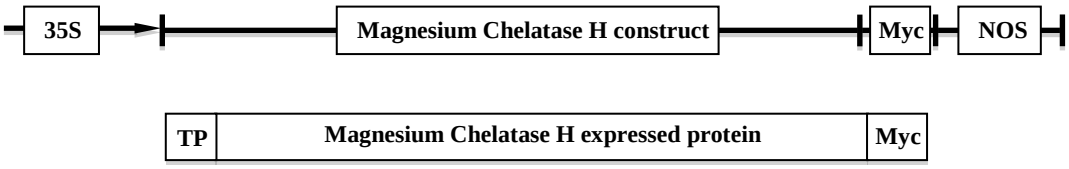


Figure 14. The Magnesium Chelatase H construct

Magnesium chelatase H expression was driven by the 35S promoter and terminated with NOS. The expressed protein included a transit peptide and a Myc tag.

```
1     ATGTCGTCGTCCCTA GTGTCCACCCCATTT GCCGCCGCGGCGCAG AAGCGGCTCCTGGCG GCGCCCGTGCCGCTG
1      M  S  S  S  L   V  S  T  P  F   A  A  A  A  Q   K  R  L  L  A   A  P  V  P  L
76    CACTCGTTCCCCCTG AGCGGCCGGCGCCAG CCTCCGCGCCGCGCC GGCACCATCCGGTGC GCGGTCGCCGGCGGC
26     H  S  F  P  L   S  G  R  R  Q   P  P  R  R  A   G  T  I  R  C   A  V  A  G  G
151   AACGGCCTCTTCACG CAGACCAAGCCCGAG GTGCGGCGCGTGGTG CCCTCCGACCCGCGG GGCCTGCCGCGGGTC
51     N  G  L  F  T   Q  T  K  P  E   V  R  R  V  V   P  S  D  P  R   G  L  P  R  V
226   AAGGTCGTCTACGTC GTGCTGGAGGCGCAG TACCAGTCGTCCGTA ACCGCCGCCGTGCAG CAGCTCAACGCCGAC
76     K  V  V  Y  V   V  L  E  A  Q   Y  Q  S  S  V   T  A  A  V  Q   Q  L  N  A  D
301   CCGCGCCGCGCCGCC GCGTTCGAGGTCGTG GGCTACCTCGTCGAG GAGCTCCGCGACGAG GACACCTACGCCACC
101    P  R  R  A  A   A  F  E  V  V   G  Y  L  V  E   E  L  R  D  E   D  T  Y  A  T
376   TTCTGCGCCGACCTC GCCGACGCCAACGTC TTCATCGGCTCCCTC ATCTTCGTCGAGGAG CTGGCCCTCAAGGTC
126    F  C  A  D  L   A  D  A  N  V   F  I  G  S  L   I  F  V  E  E   L  A  L  K  V
451   AAGGCCGCCGTCGAG AAGGAGCGCGACCGC ATGGACGCCGTCCTC GTCTTCCCCTCAATG CCCGAGGTCATGCGC
151    K  A  A  V  E   K  E  R  D  R   M  D  A  V  L   V  F  P  S  M   P  E  V  M  R
526   CTCAACAAGCTCGGC TCCTTCAGCATGTCG CAGCTGGGGCAGTCC AAGAGCCCCTTCTTC CAGCTCTTCAAGCGC
176    L  N  K  L  G   S  F  S  M  S   Q  L  G  Q  S   K  S  P  F  F   Q  L  F  K  R
601   AACAAGGCCAACTCC AGCAACTTCGCCGAC AGCATGCTCAAGCTC GTGCGCACGCTGCCC AAGGTGCTCAAGTAC
201    N  K  A  N  S   S  N  F  A  D   S  M  L  K  L   V  R  T  L  P   K  V  L  K  Y
676   CTGCCCTCCGACAAG GCGCAGGACGCCCGG CTCTACATCCTCAGC CTCCAGTTCTGGCTC GGCGGCTCGCCGGAC
226    L  P  S  D  K   A  Q  D  A  R   L  Y  I  L  S   L  Q  F  W  L   G  G  S  P  D
751   AACCTCCAGAACTTC CTTAAGATGATCGCC GGCTCCTACGTGCCT GCCCTCAAGGGCGCC GGCATCAAGTACGAC
251    N  L  Q  N  F   L  K  M  I  A   G  S  Y  V  P   A  L  K  G  A   G  I  K  Y  D
826   GACCCCGTGCTCTAC CTCGACTCCGGCATC TGGCACCCGCTGGCG CCCACCATGTACGAG GACGTCAAGGAGTAC
276    D  P  V  L  Y   L  D  S  G  I   W  H  P  L  A   P  T  M  Y  E   D  V  K  E  Y
901   CTCAACTGGTACGAC ACGCGCCGGGACGCC AACGACAAGCTCAAA GACCCCAAGGCGCCC GTCATCGGCCTCGTC
301    L  N  W  Y  D   T  R  R  D  A   N  D  K  L  K   D  P  K  A  P   V  I  G  L  V
976   CTGCAGAGGAGCCAC ATTGTCACCGGCGAC GACGGGCACTACGTC GCCGTCATCATGGAG CTCGAGGCCAAGGGC
326    L  Q  R  S  H   I  V  T  G  D   D  G  H  Y  V   A  V  I  M  E   L  E  A  K  G
1051  GCCAAGGTCATACCC ATCTTCGCCGGCGGC CTCGACTTCTCCGGG CCCACACAGCGCTAC CTGGTCGACCCGATT
351    A  K  V  I  P   I  F  A  G  G   L  D  F  S  G   P  T  Q  R  Y   L  V  D  P  I
1126  ACCGGCAAGACGTTC GTGAACGCCGTGGTG TCTCTCACCGGGTTC GCGCTCGTCGGCGGG CCGGCGAGGCAGGAC
376    T  G  K  T  F   V  N  A  V  V   S  L  T  G  F   A  L  V  G  G   P  A  R  Q  D
1201  CATCCCAAGGCCATT GCCGCGCTGCAGAAG CTCGACGTGCCGTAC ATTGTCGCGCTCCCG CTCGTGTTCCAGACA
401    H  P  K  A  I   A  A  L  Q  K   L  D  V  P  Y   I  V  A  L  P   L  V  F  Q  T
1276  ACGGAGGAGTGGCTC AACAGCACCTTGGGG CTTCACCCAATTCAG GTGGCGCTGCAGGTC GCGCTGCCGGAGCTC
426    T  E  E  W  L   N  S  T  L  G   L  H  P  I  Q   V  A  L  Q  V   A  L  P  E  L
1351  GACGGTGGGATGGAG CCCATCGTGTTCGCC GGCCGGGACCCCAGG ACAGGGAAGTCACAT GCATTGCACAAGAGA
451    D  G  G  M  E   P  I  V  F  A   G  R  D  P  R   T  G  K  S  H   A  L  H  K  R
1426  GTGGAGCAGCTCTGC ACTAGAGCCATCAGA TGGGCAGAACTGAAG AAGAAAACTAAGGAG GAGAAGAGACTGGCG
476    V  E  Q  L  C   T  R  A  I  R   W  A  E  L  K   K  K  T  K  E   E  K  R  L  A
1501  ATCACTGTTTTCAGT TTCCCACCTGACAAG GGCAACGTCGGGACT GCAGCATATCTGAAC GTGTTCAGCTCCATC
501    I  T  V  F  S   F  P  P  D  K   G  N  V  G  T   A  A  Y  L  N   V  F  S  S  I
1576  TACTCTGTGCTCTCA GACCTCAAGAAGGAC GGCTACAACGTGGAG GGTCTTCCGGACACA CCTGAAGCCCTCATC
526    Y  S  V  L  S   D  L  K  K  D   G  Y  N  V  E   G  L  P  D  T   P  E  A  L  I
1651  GAGGAGGTGATCCAT GACAAGGAAGCTCAG TTCAACAGCCCCAAC CTAAATGTTGCTTAC CGCATGAATGTGAGG
551    E  E  V  I  H   D  K  E  A  Q   F  N  S  P  N   L  N  V  A  Y   R  M  N  V  R
1726  GAGTACCAGTCGCTG ACCTCCTACGCCTCC TTGCTGGAGGAGAAC TGGGGGAAGCCACCT GGGCACCTCAACTCT
576    E  Y  Q  S  L   T  S  Y  A  S   L  L  E  E  N   W  G  K  P  P   G  H  L  N  S
1801  GATGGCGAGAACCTC CTCGTCTACGGGAAG CAGTACGGCAATGTC TTCATCGGGGTGCAG CCCACCTTTGGGTAC
601    D  G  E  N  L   L  V  Y  G  K   Q  Y  G  N  V   F  I  G  V  Q   P  T  F  G  Y
1876  GAAGGTGATCCCATG CGGCTTCTCTTCTCA AAGTCTGCCAGCCCT CACCATGGATTTGCA GCATACTACACCTTT
626    E  G  D  P  M   R  L  L  F  S   K  S  A  S  P   H  H  G  F  A   A  Y  Y  T  F
1951  GTCGAGAAGATCTTC CAGGCCGATGCTGTT CTGCACTTTGGAACA CACGGGTCCCTCGAG TTCATGCCTGGCAAG
651    V  E  K  I  F   Q  A  D  A  V   L  H  F  G  T   H  G  S  L  E   F  M  P  G  K
2026  CAGGTTGGGATGAGT GACGCCTGCTTCCCT GACAGCCTCATTGGC AACATCCCCAACATC TACTACTATGCTGCA
676    Q  V  G  M  S   D  A  C  F  P   D  S  L  I  G   N  I  P  N  I   Y  Y  Y  A  A
2101  AACAACCCATCAGAG GCCACGGTGGCCAAG CGCCGGAGCTACGCG AACACCATCAGCTAC CTGACCCCACCGGCC
701    N  N  P  S  E   A  T  V  A  K   R  R  S  Y  A   N  T  I  S  Y   L  T  P  P  A
2176  GAGAACGCCGGCCTC TACAAGGGGCTCAAG CAGCTGTCAGAGCTC ATCTCTTCCTACCAG TCTCTCAAGGACACC
726    E  N  A  G  L   Y  K  G  L  K   Q  L  S  E  L   I  S  S  Y  Q   S  L  K  D  T
```

```
2251  GGGCGTGGTCCTCAG ATTGTGAGCTCCATC GTCAGCACTGCAAAG CAGTGCAACCTCGAC AAGGATGTCCCGCTG
751     G  R  G  P  Q   I  V  S  S  I   V  S  T  A  K   Q  C  N  L  D   K  D  V  P  L
2326  CCCGAGGAAGGGGAG GAGCTCCCACCAAAG GAGCGTGACCTTGTC GTTGGGAAGGTGTAC GCCAAGATCATGGAG
776     P  E  E  G  E   E  L  P  P  K   E  R  D  L  V   V  G  K  V  Y   A  K  I  M  E
2401  ATAGAGTCACGGCTC CTTCCCTGCGGTCTG CATGTCATCGGCGAG CCGCCGAGTGCCATC GAGGCGGTGGCCACG
801     I  E  S  R  L   L  P  C  G  L   H  V  I  G  E   P  P  S  A  I   E  A  V  A  T
2476  CTGGTGAACATAGCT GCCCTCGACCGCCCC GAGGACGGCATAACC TCGCTGCCCGGCATA CTTGCCGCCACAGTG
826     L  V  N  I  A   A  L  D  R  P   E  D  G  I  T   S  L  P  G  I   L  A  A  T  V
2551  GGCAGGGATATTGAA GATGTGTACAGGGGA AGTGACAAGGGCATA CTGGCTGACGTCGAG CTTCTGAGGCAGATC
851     G  R  D  I  E   D  V  Y  R  G   S  D  K  G  I   L  A  D  V  E   L  L  R  Q  I
2626  ACTGAGGCTTCGCGC GGCGCCATCACCGCC TTCGTTGAGAAGACC ACAAACAGCAAAGGG CAAGTCGTCAATGTT
876     T  E  A  S  R   G  A  I  T  A   F  V  E  K  T   T  N  S  K  G   Q  V  V  N  V
2701  ACCAACAACCTCAGC AAGATACTTGGTTTC GGTCTGTCGGAACCA TGGGTGCAGTACCTG TCCACGACCAAGTTC
901     T  N  N  L  S   K  I  L  G  F   G  L  S  E  P   W  V  Q  Y  L   S  T  T  K  F
2776  GTCAGAGCGGACAGA GAGAAGATGAGGGTT CTGTTTGGGTTCTTG GGGGAGTGCCTGAGG CTCGTCGTGCAAGAC
926     V  R  A  D  R   E  K  M  R  V   L  F  G  F  L   G  E  C  L  R   L  V  V  Q  D
2851  AACGAGCTGGGAAGC TTGAAGCTTGCCCTC GAGGGAAGCTACGTC GAGCCTGGCCCTGGC GGCGACCCGATCCGT
951     N  E  L  G  S   L  K  L  A  L   E  G  S  Y  V   E  P  G  P  G   G  D  P  I  R
2926  AACCCGAAGGTGCTC CCGACAGGGAAGAAC ATCCACGCTCTCGAT CCGCAGGCCATCCCA ACCACGGCTGCCTTG
976     N  P  K  V  L   P  T  G  K  N   I  H  A  L  D   P  Q  A  I  P   T  T  A  A  L
3001  AAGAGCGCCAAGATC GTCGTGGACCGTCTC CTGGAGAGGCAGAAG GCTGACAATGGCGGC AAGTACCCTGAGACG
1001    K  S  A  K  I   V  V  D  R  L   L  E  R  Q  K   A  D  N  G  G   K  Y  P  E  T
3076  GTCGCACTTGTCCTG TGGGGCACCGACAAC ATCAAGACCTATGGT GAGTCACTAGCCCAG GTGCTGTGGATGATT
1026    V  A  L  V  L   W  G  T  D  N   I  K  T  Y  G   E  S  L  A  Q   V  L  W  M  I
3151  GGAGTTCGGCCAGTT GCCGACACCTTCGGC CGTGTCAACCGTGTG GAGCCTGTCAGCCTT GAGGAGCTTGGACGC
1051    G  V  R  P  V   A  D  T  F  G   R  V  N  R  V   E  P  V  S  L   E  E  L  G  R
3226  CCAAGGATCGATGTC GTCGTCAATTGCTCG GGTGTTTTCAGAGAT CTTTTCATCAACCAG ATGAACCTGCTGGAC
1076    P  R  I  D  V   V  V  N  C  S   G  V  F  R  D   L  F  I  N  Q   M  N  L  L  D
3301  AGGGCAGTGAAGATG GTCGCCGAACTGGAC GAGCCAGCAGAGATG AACTACGTGCGCAAG CACGCCCAGGAGCAG
1101    R  A  V  K  M   V  A  E  L  D   E  P  A  E  M   N  Y  V  R  K   H  A  Q  E  Q
3376  GCGGAGGAGCTCGGC GTGTCGCTAAGGGAG GCGGCGACAAGGGTG TTCTCGAACGCATCA GGCTCCTACTCGTCC
1126    A  E  E  L  G   V  S  L  R  E   A  A  T  R  V   F  S  N  A  S   G  S  Y  S  S
3451  AACGTGAACCTGGCC GTGGAGAACGCGTCA TGGACCGACGAGAAG CAGCTCCAGGACATG TACCTGAGCCGCAAG
1151    N  V  N  L  A   V  E  N  A  S   W  T  D  E  K   Q  L  Q  D  M   Y  L  S  R  K
3526  TCCTTCGCGTTCGAC AGCGACGCCCCAGGG GCAGGCATGAAGGAG AAGCGCAAGGCGTTC GAGCTCGCCCTGGCG
1176    S  F  A  F  D   S  D  A  P  G   A  G  M  K  E   K  R  K  A  F   E  L  A  L  A
3601  ACGGCGGACGCCACG TTCCAGAACCTCGAC TCGTCGGAGATCTCG CTGACGGACGTGAGC CACTACTTCGACTCG
1201    T  A  D  A  T   F  Q  N  L  D   S  S  E  I  S   L  T  D  V  S   H  Y  F  D  S
3676  GACCCGACCAAGCTC GTGCAGGGGCTGCGC AAGGACGGGCGGGCG CCGTCCTCGTACATA GCCGACACCACCACG
1226    D  P  T  K  L   V  Q  G  L  R   K  D  G  R  A   P  S  S  Y  I   A  D  T  T  T
3751  GCGAACGCCCAGGTG AGGACGCTGTCGGAG ACGGTGCGCCTCGAC GCGAGGACCAAGCTG CTGAACCCCAAGTGG
1251    A  N  A  Q  V   R  T  L  S  E   T  V  R  L  D   A  R  T  K  L   L  N  P  K  W
3826  TACGAGGGGATGATG AAGAGCGGGTACGAG GGGGTCAGGGAGATC GAGAAGCGGCTCACC AACACCGTCGGGTGG
1276    Y  E  G  M  M   K  S  G  Y  E   G  V  R  E  I   E  K  R  L  T   N  T  V  G  W
3901  AGCGCCACGTCTGGG CAGGTCGACAACTGG GTCTACGAGGAGGCC AACTCCACGTTCATC GAGGACGAGGCGATG
1301    S  A  T  S  G   Q  V  D  N  W   V  Y  E  E  A   N  S  T  F  I   E  D  E  A  M
3976  AGGAAGAGGCTCATG GACACCAACCCCAAT TCGTTCAGGAAGTTG GTGCAGACCTTCCTG GAAGCCAGTGGCAGA
1326    R  K  R  L  M   D  T  N  P  N   S  F  R  K  L   V  Q  T  F  L   E  A  S  G  R
4051  GGCTACTGGGAGACA ACGGAGGAGAACCTG GACAGGCTCAGGGAG CTCTATTCGGAGGTT GAAGACAAGATTGAG
1351    G  Y  W  E  T   T  E  E  N  L   D  R  L  R  E   L  Y  S  E  V   E  D  K  I  E
4126  GGGATTGACAGGGGA TCCTAA
1376    G  I  D  R  G   S  *
```

Figure 15. Magnesium chelatase H sequence

Sequence for cloned magnesium chelatase subunit H. Each row of nucleotide sequence is immediately followed by a row of predicted protein sequence based on the standard genetic code. Nucleotide/amino acid numbering is shown to the left of each row, and represents the position of the first character in that row.

2.2.3. Magnesium Chelatase D truncation constructs

Preliminary *in Planta* overexpression of the Mg-chelatase D wild type protein resulted in an unexpected chlorotic phenotype (to be discussed in Chapter 3). We were interested in investigating the cause for this phenotype, and specifically if the observed chlorosis could be attributed to a specific region of the Mg-chelatase D protein. We therefore separated the CHLD protein into two truncations, the first comprised of the N-terminus and poly-Proline linker region, and the second consisting of the C-terminal region, with an engineered transit peptide fused to its N-terminus (fig. 16).

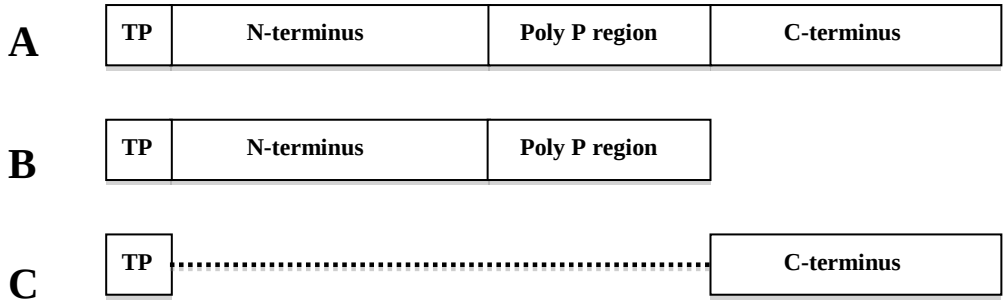


Figure 16. Magnesium chelatase D truncation constructs

Shown in this figure are the truncation constructs generated from CHLD full-length sequence. A. Magnesium chelatase D full length construct encoding wild type protein B. Magnesium chelatase D N-terminal construct encoding a protein comprised of the transit peptide, the N-terminal region and the poly-Proline linker region C. Magnesium chelatase D C-terminal construct encoding a protein comprised of the transit peptide engineered onto the C-terminal region.

2.2.3.1 Magnesium chelatase D 5' truncation construct

This section of the DNA, comprised of the N-terminal region plus the poly-Proline linker region (fig. 16B), was obtained in a PCR reaction using primers pf320 and pf392. These primers were designed to incorporate the *Xba*1 and

*Bam*H1 cut-sites necessary to sub-clone the product into the pBin61 expression vector, and produced an amplicon 1253 bp long. The amplification program included an initial 95°C incubation to activate the Hot Star polymerase prior to thirty rounds of cycling.

2.2.3.2 Magnesium chelatase D 3' truncation construct

To investigate the 3' end of the *CHLD* gene, two constructs were generated. The first was an un-modified version, starting just after the poly-Proline region and extending to the stop codon, which was ablated in order to allow us to express a tag, if we desired. The second 3' truncation construct modified the 3' region (described above) by the addition of a 5' transit peptide. The second construct was generated to enable import of the transiently expressed protein into the plastid. To

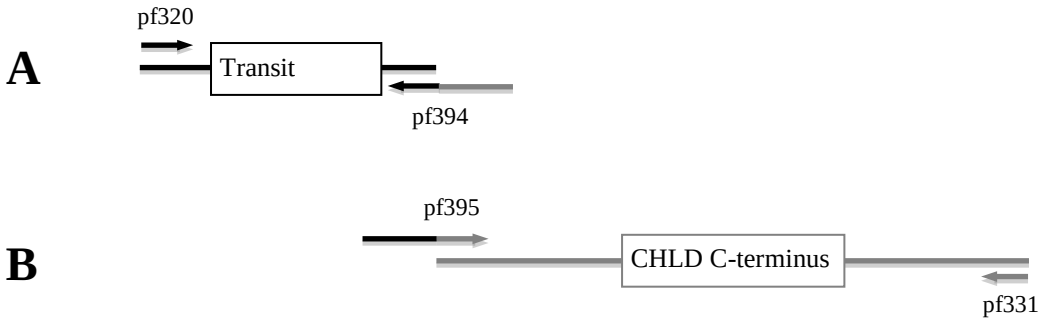


Figure 17. Synthesis of CHLD C-terminal truncation construct

Synthesis of the CHLD 3' construct, encoding the C-terminal region of the protein, occurred in two stages. In the first stage, fragments A and B were PCR synthesized, with stage two accomplishing fusion of the two fragments. The 3' reverse primer of fragment A (pf394) was engineered to include sequence from the 5' end of fragment B, so that in the second stage of this process, melting of fragment A would generate sequence complementary to fragment B, and thus allow fusion of the two. Primer pf395 was engineered in a similar way, to further enable fusion of the two fragments.

do this, we utilized the ChloroP prediction software (Emanuelsson et al. 1999) to determine the region of the *Zea mays CHLD* gene most likely responsible for

chloroplastic import, and amplified this region separately from the 3’ end of the *CHLD* gene (fig. 17). The primers used on the 3’ end of the transit peptide region and the 5’ end of the genic region encoding the C-terminus were engineered so as to produce complementary ends after melting (fig. 16).

2.3 Discussion

The subunits of Mg-chelatase were cloned to facilitate study of over-expression of this enzyme in mature leaf tissue. Use of a transient expression system allowed us to study the effect of over-expression of a specific Mg-chelatase subunit, without the confusion resulting from developing tissue (seedlings) or from pleiotropic effects of a breakdown in carotenoid biosynthesis, as is seen with experiments employing NF.

2.3.1. PCR amplification of magnesium chelatase D- and -H subunits

PCR amplification of *CHLD* was a relatively straightforward process, with the majority of the nucleotide sequence obtained using basic polymerase chain reaction. The 5' end of the gene, obtained using 5' RACE, was performed using a kit (see materials and methods section) and proceeded without issue.

Amplification of the *CHLH* gene, however, proved more difficult, taking over a year to complete. Troubleshooting the amplification of this gene encompassed many strategies, including re-design of primers, adjustment of the magnesium concentration, testing of multiple polymerase mixes (including high fidelity, long template, and hot start mixes from various companies), and adjustment of the annealing and extension temperatures. After repeated failures to amplify the full-length sequence, the cloning strategy was modified to amplify the gene in 4 separate reactions, and to fuse the amplicons in a final PCR amplification. The primers for three of the fragments (called fragments 1, 2, and 3) were designed

using a *Z. mays* partial mRNA sequence identified by blast (see materials and methods section), while the 5' end of the gene was obtained using 5' RACE.

Of the four amplicons, the most difficult to amplify was CHLH fragment 1. Re-design of the primers proved to be ineffectual, and successful amplification of the target was ultimately accomplished only by the addition of DMSO, including an initial high temperature soak, and performing two rounds of PCR amplification. To obtain the fragment, the product of the first round of amplification was gel purified and used as template in the second round of amplification (which, like round one, included the adjustments listed above). In this way, the round two reaction was enriched for the target sequence.

Examination of the GC content shows that this region of the gene was extremely GC rich, with an average GC content of 65% (fig. 18). This may explain the

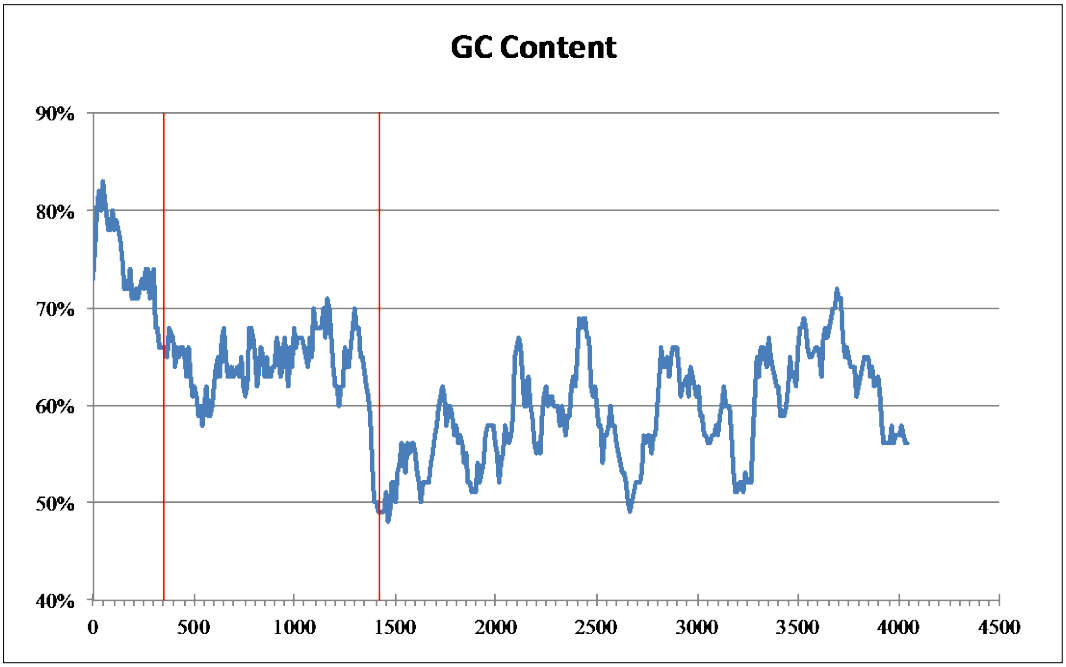


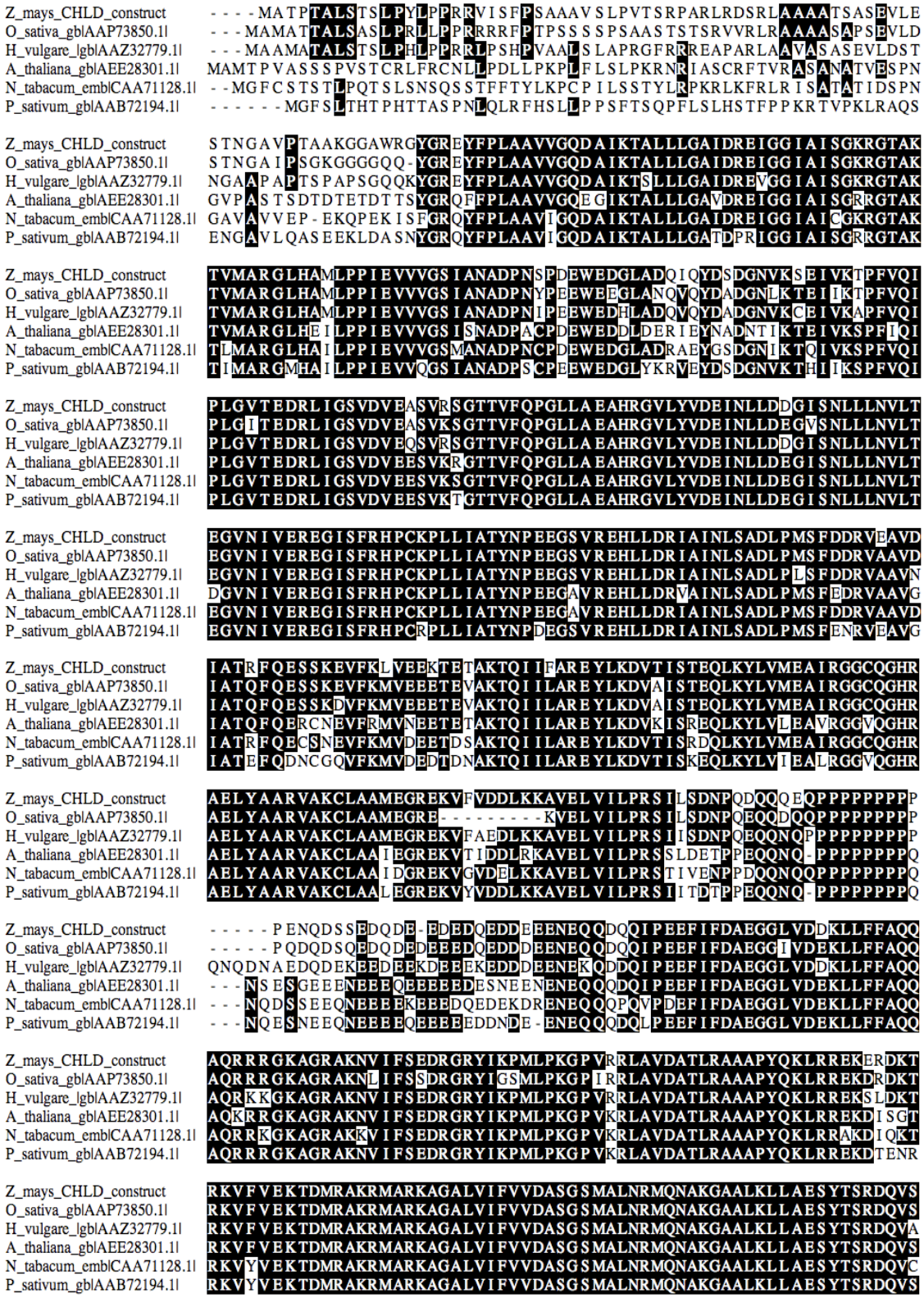
Figure 18. Magnesium chelatase H GC content

This figure shows the GC content of full-length magnesium chelatase H. The x-axis of the graph reflects the nucleotide position on the ChlH gene, while the y-axis shows the respective GC content for each position. Fragment one is depicted as the region bound by red lines, where the red lines represent the amplification primers pf322 and pf323. This fragment, with an amplicon length of 1075 bp, has an average GC content of 65%.

difficulty in the amplification of the region. However, if GC content were the only factor influencing the reaction, amplification of the 5'RACE fragment should have proved more difficult. This was not, in fact, the case. While GC content for the 5'RACE fragment averaged at 74.6%, and in theory, should make its amplification more difficult than that of fragment 1, amplification of this fragment was relatively straightforward. In comparison to the other CHLH fragments, amplification of the 5'RACE fragment was only more involved than that of fragments 2 and 3 due to the additional steps involved in addition of the RACE adapter. In light of this fact, I would speculate that there was some other factor contributing to the difficulty of fragment 1 amplification, and would speculate that perhaps the cDNA in some way reflects secondary structure existing in the mRNA.

2.3.2. Construct verification using multiple sequence alignment

Due to the fact that un-annotated mRNA sequence was used to design primers for amplification of the *CHLD* and *CHLH* sequences, a final verification step was performed using ClustalW multiple sequence alignment of the translated cloned sequence compared to known protein sequences (figs. 19 and 20, respectively).
Both sequences show high levels of sequence similarity throughout their sequence lengths. Interestingly, some amino acid residues appear to be conserved only within the monocot/dicot groupings. Based on the high level of sequence similarity, it seems highly probable that the cloned sequences are in fact *Zea mays* Mg-chelatase D-subunit and H-subunit genes.


Z_mays_CHLD_construct
O_sativa_gb|AAP73850.1|
H_vulgare_|gb|AAZ32779.1|
A_thaliana_gb|AEE28301.1|
N_tabacum_emb|CAA71128.1|
P_sativum_gb|AAB72194.1|

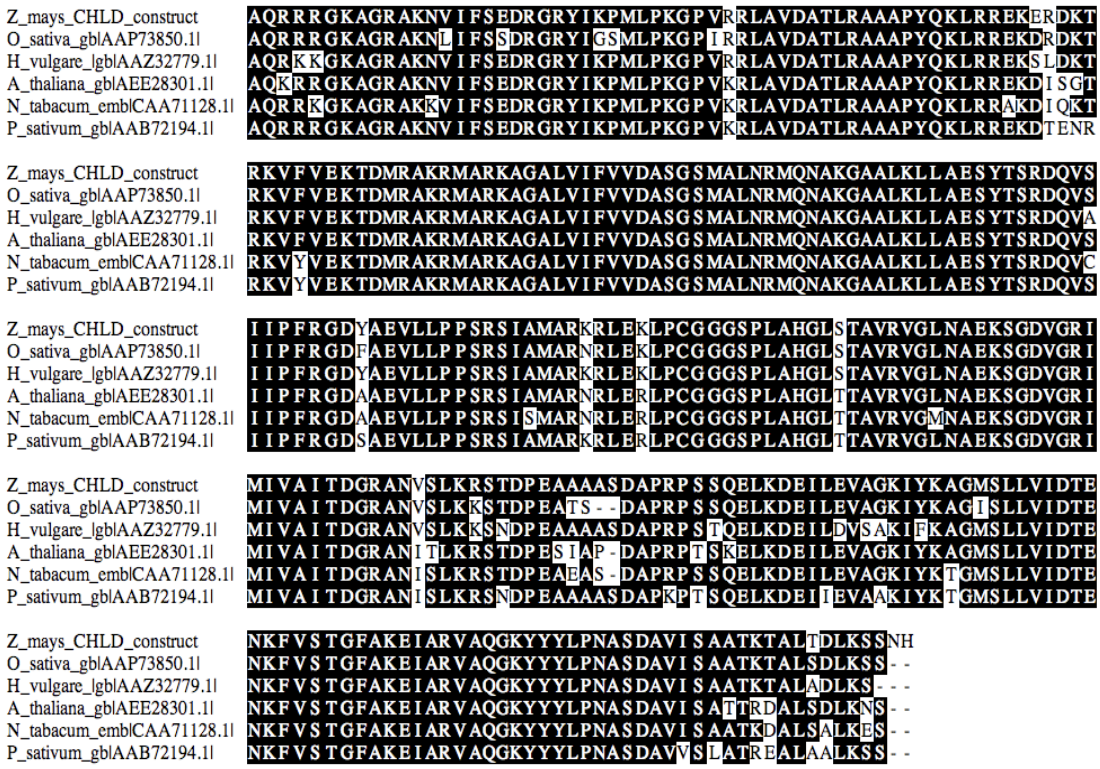
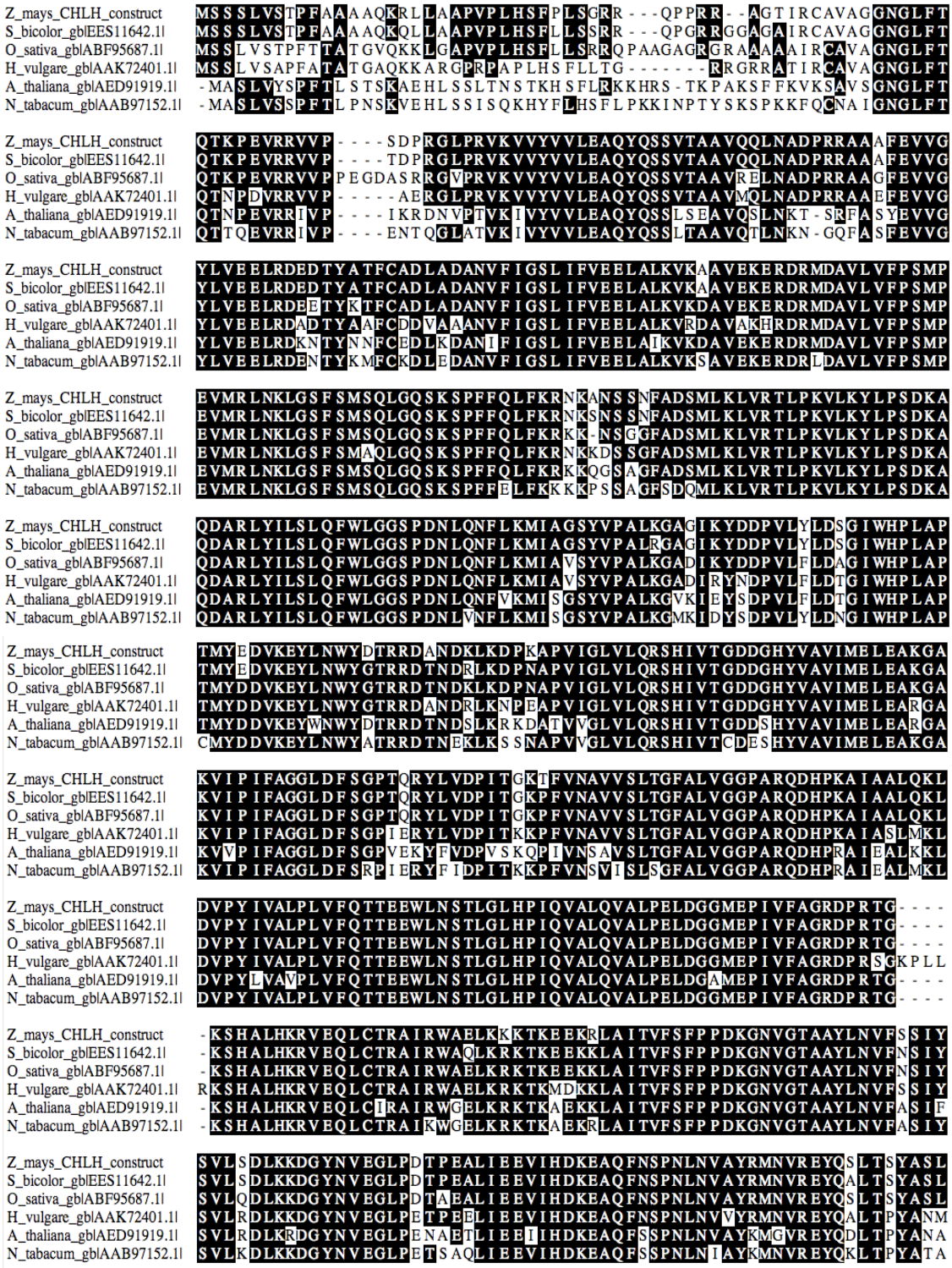


Figure 19. Multiple sequence alignment of CHLD proteins

This figure shows the protein alignment comparing the cloned putative CHLD sequence (Z. mays CHLD construct) compared to magnesium chelatase protein sequences from rice (O. sativa), barley (H. vulgare), thale cress (A. thaliana), tobacco (N. tabacum) and pea (P. sativum).


Z_mays_CHLH_construct
S_bicolor_gb|EES11642.1|
O_sativa_gb|ABF95687.1|
H_vulgare_gb|AAK72401.1|
A_thaliana_gb|AED91919.1|
N_tabacum_gb|AAB97152.1|

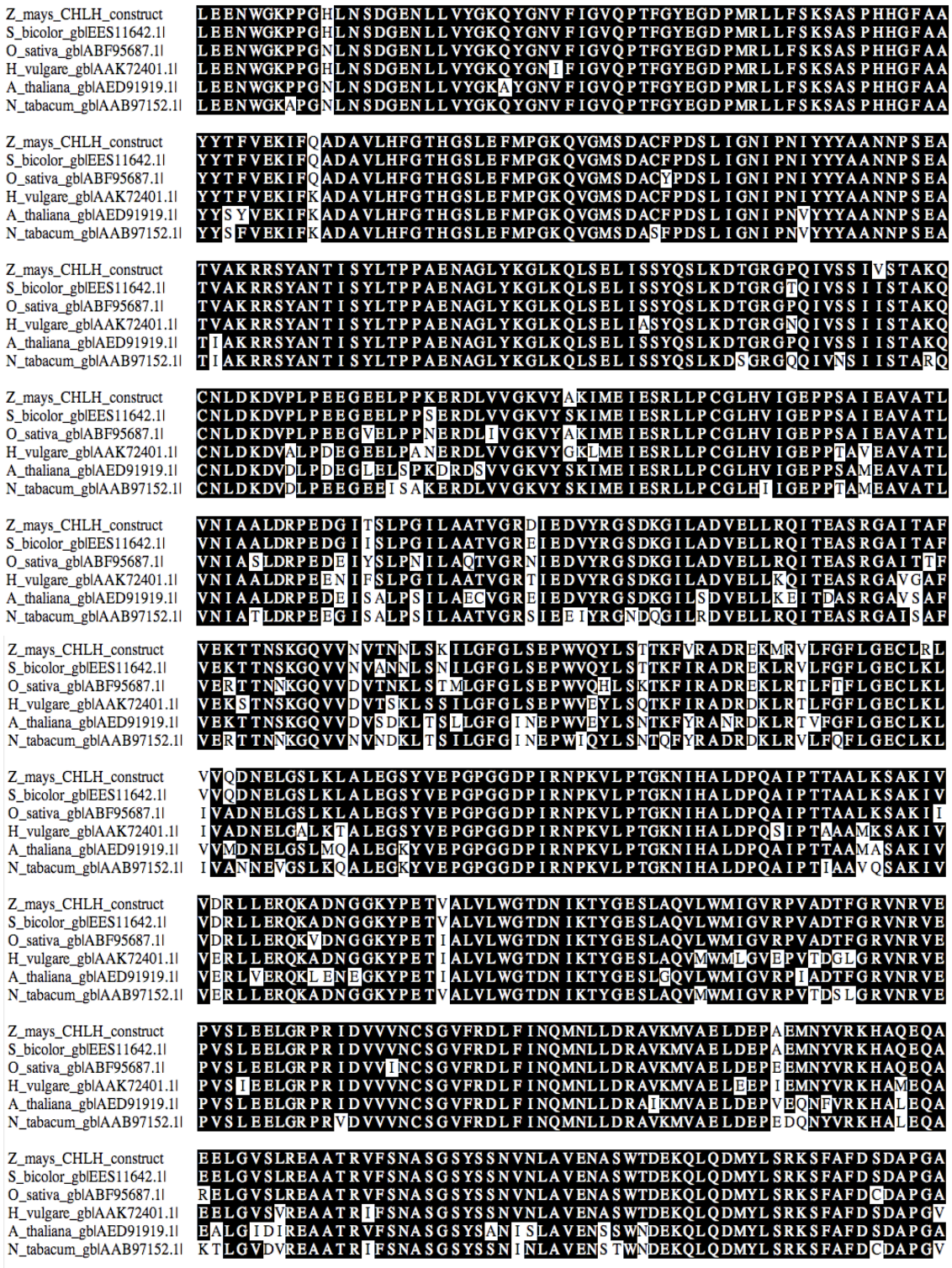

Z_mays_CHLH_construct
S_bicolor_gblEES11642.1|
O_sativa_gblABF95687.1|
H_vulgare_gblAAK72401.1|
A_thaliana_gblAED91919.1|
N_tabacum_gblAAB97152.1|

Figure 20. Magnesium chelatase H multiple sequence alignment

This figure shows the protein alignment comparing the cloned putative CHLH sequence (Z. mays CHLH construct) compared to magnesium chelatase protein sequences from Sorghum, (S. bicolor) rice (O. sativa), barley (H. vulgare), thale cress (A. thaliana), tobacco (N. tabacum).

Chapter 3. Phenotypic analysis of Magnesium Chelatase subunit over-expression in *Nicotiana benthamiana* leaf tissue

Traditionally, studies of Mg-chelatase have employed the use of mutant plant lines and/or seedlings treated with the herbicide NF (Gadjieva et al. 2005; Mochizuki et al. 2001; Susek et al. 1993). While these types of studies were instrumental in the illustration of a chloroplastic retrograde signal to the nucleus (Susek et al. 1993), the use of NF may result in pleiotropic or secondary effects, as correct development of the etioplast into the chloroplast is compromised. The use of transient expression allows correct development of the chloroplast prior to rapid over-expression of a specific protein (in this case one or multiple subunits of Mg-chelatase) and functional analysis of the expressed protein in native, or near native environments (Janssen and Gardner 1990; Sawers et al. 2006a). Thus, changes in the physiology of the plant differing from that of the empty-vector control can be attributed directly to the over-expressed protein. Furthermore, disruption of Mg-chelatase function is easily scorable as it results in chlorosis of the infiltrated section of the leaf (Sawers et al. 2006a; Sawers et al. 2006b).

Previous work has shown the efficacy of *Zea mays* CHLI interaction with *Nicotiana benthamiana* endogenous proteins (Sawers et al. 2006a). Furthermore, this system allows the delivery of several transgenes into the same cell (Kapila et al. 1997) so that multimeric proteins can be expressed and assembled (Vaquero et al. 1999). In this study, we examine the phenotypic effects of Mg-chelatase subunit over-expression in mature leaf tissue, using *Agrobacterium* mediated transient expression. We also investigate the interaction of CHLD and CHLI,

using truncation constructs of Mg-chelatase D and a CHLD mutant construct with an amino acid substitution in the MIDAS motif (see section 3.2.4).

3.1 Materials and Methods

3.1.1. Plant production

Plants were initially grown in pure potting soil, however, the composition of the soil proved too dense to raise healthy plants, so I formulated a soil mixture (39.9% coir, 39.9% vermiculite, 19.9% potsoil, 0.2% osmocote fertilizer) loosely based on Cornell mix (Boodley and Sheldrake 1982). This mix was used for both starting the seedlings as well as transplanting.

Flats were seeded lightly and germinated in the greenhouse at 25°C without supplemental lighting for approximately 3 weeks before transplant into clean 2 liter pots, and grown under the same conditions up to and during experimentation. Transplanted plants were used for infiltration experiments 10 days after transplant.

3.1.2. *Agrobacterium* infiltration

3.1.2.1 Cell growth

Agrobacterium tumefaciens cells were grown on LB Agar plates with tetracycline (Tet) (15µg/ml) to select for the p19 helper plasmid and Kanamycin (Kan) (50µg/ml) to select for the pBin61 plasmid at 28° C for two days. Discrete colonies were then sterilely transferred into 7-10 milliliters of L-medium (0.1% w/v bacto-tryptone, 0.5% w/v bacto-yeast extract, 0.5% NaCl, 0.1% D-glucose) with tetracycline and kanamycin (15 µg/ml and 50 µg/ml, respectively) selection and grown at 28°C overnight with shaking. Cells were then pelleted and re-

suspended in infiltration buffer (10 mM MES, 10 mM $MgCl_2$, 100-150 µM Acetosyringone). The infiltration solution was made by measuring the cell titer using an Eppendorf BioPhotometer and diluting the culture with infiltration buffer to an optical density of 200.

3.1.2.2 Empty vector as a control

Because *Agrobacterium* is a plant pathogenic bacterium, we anticipated that infiltrations done into the leaf tissue may have a baseline effect on the plant. For that reason, to accurately assess the phenotype of over-expressed proteins, empty pBIN61 vector was used to transform *Agrobacterium*, and this was used as a negative control.

3.1.2.3 Syringe infiltration

Preliminary experimentation showed that selection of a healthy young plant with five to six true leaves is crucial for efficient phenotype development. The leaf used for the infiltration experiment was not allowed to exceed a diameter of 5 cm, with a 3 cm diameter being ideal. Using a scalpel, a small incision was made in the lower epidermal layer of the fourth or fifth mature leaf, midway between the mid-vein and the leaf margin. The infiltrate-loaded syringe was then pressed firmly but gently against the lower leaf surface and enough infiltrate injected into the leaf to make an infiltration area with a diameter of approximately 1 – 1.5 cm. Infiltration experiments were repeated for biological replication (table 1).

3.1.3. Site-directed mutagenesis of the CHLD MIDAS motif

Using the primer pair pf320 (CHLD full length forward with engineered Xba1 site) and pf368 (reverse primer engineering the MIDAS mutation DASGS to DAAGS), a 5’ fragment of 1685 bp was generated with a mutation in the MIDAS of Mg-chelatase D. The 3’ end of the gene was amplified using pf331 and pf369, to form a product of 611 bp with the same mutation incorporated. Full-length Mg-chelatase D with the MIDAS mutation (DASGS to DAAGS) was obtained by using the two fragments as template in a PCR reaction employing KOD high fidelity polymerase. In a final step after cycling, A’s were incorporated onto the amplicon ends by adding Promega Taq polymerase, and incubating at 72°C for 10 minutes without a melting step. The final product was cloned into pGEM-T Easy vector using the vector T-overhangs and sequence verified using Sanger sequencing before sub-cloning into the pBIN-61 binary vector as described in section 2.2.4.

3.2 Results

Infiltration experiments were repeated a minimum of 3 times, with a greater number of infiltrations for single infiltrations of the basic constructs (table 1).

Table 1. Phenotypes of leaf infiltrations

Construct infiltrated	*Biological replicates*	*Phenotype*
Empty vector	*14*	*wt*
CHLI	*21*	*wt*
CHLD	*29*	*chlorosis*
CHLD with mutated MIDAS (DM9)	*9*	*chlorosis*
CHLH	*15*	*wt*
GUN4	*6*	*wt*
CHLD + Empty vector	*8*	*chlorosis*
CHLD + CHLI	*12*	*wt*
CHLD + GUN4	*4*	*chlorosis*
CHLD N-terminal construct	*18*	*chlorosis*
N-term + CHLI	*14*	*chlorosis*
CHLD C-terminal construct	*3*	*wt*
CHLD C-terminal construct with transit peptide (CTP)	*9*	*wt*
CTP + CHLI	*4*	*wt*
CHLD + CHLI + GUN4	*6*	*wt*
CHLD + CHLH	*4*	*chlorosis*
N term + CTP	*12*	*chlorosis*
N term + CTP + Empty vector	*4*	*chlorosis*
N term + CTP + CHLI	*14*	*wt*
DM9 + GUN4	*4*	*chlorosis*
DM9 + CHLI	*8*	*chlorosis*
DM9 + CTP	*10*	*chlorosis*
DM9 + CTP + Empty vector	*4*	*chlorosis*
DM9 + CTP + CHLI	*11*	*wt*
DM9 + N term	*10*	*chlorosis*
DM9 + N term + Empty	*6*	*chlorosis*
DM9 + N term + CHLI	*10*	*chlorosis*
CHLI + GUN4	*5*	*wt*
CHLI + CHLD + CHLH	*3*	*wt*
CHLI + CHLD + CHLH + GUN4	*5*	*wt*
POY4	*4*	*chlorosis*
POY4 + CHLD	*4*	*chlorosis*
POY5	*4*	*chlorosis*
POY5 + CHLD	*4*	*chlorosis*

*Wild type is defined as having a phenotype that is a green indistinguishable from un-infiltrated leaf tissue or as being visually identical to the phenotype seen with an infiltration of Empty vector, which gives a very slightly (just discernable) chlorsis.

**CTP: C-terminal region of CHLD with transit peptide added

***N term: N-terminal region of CHLD, including the poly-proline region

****DM9: CHLD with mutated MIDAS motif

*****POY4, POY5: CHLI semi-dominant mutants Oy1-N1989 and Oy1-N700, respectively

3.2.1. Over-expression of magnesium chelatase subunits in *Nicotiana benthamiana*

Transient expression of individual subunits (CHLI, CHLD and CHLH) of Mg-chelatase in *Nicotiana benthamiana* was accomplished using *Agrobacterium tumefaciens* carrying a subunit construct inserted into the binary vector pBin61 (see Chapter 2 for generation of the constructs). Plants were then grown for four days, prior to examination for any visible phenotype. Construct expression was verified using iTRAQ analysis (see chapter 5). As all constructs encoded wild type proteins, the expected result was that each would generate a wild type phenotype.

The HA-tagged wild type CHLI encoding construct was infiltrated into *N. benthamiana* and no visually detectable difference was seen between this construct and the empty vector (fig. 21A). Transient over-expression of Myc-

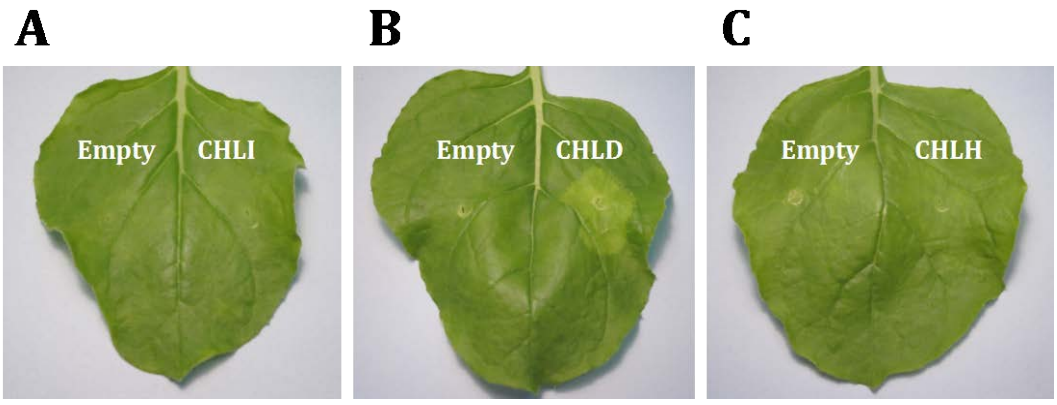


Figure 21. Over-expression of magnesium chelatase I, D and H subunits

Phenotypic analysis of magnesium chelatase subunits over-expressed in Nicotiana benthamiana leaf tissue *shows that while A. CHLI and C. CHLH result in wild type phenotypes, over-expression of the CHLD subunit results in chlorosis of the infiltrated region (B).*

tagged Mg-chelatase H, likewise, had a wild type phenotype (fig. 21C), as expected. In contrast, infiltration of the construct encoding T7-tagged wild type Mg-chelatase D subunit resulted in a strongly chlorotic (yellow) phenotype (fig. 21B). The CHLD construct was re-sequenced, re-analyzed and verified as being wild type, and the infiltration experiment was repeated three more times to verify the phenotype.

To investigate the possibility that this phenotype was a result of the tag, the CHLD construct was sub-cloned into pBIN-61 vector with a Flag-6xHis tag, and into pBIN-61 with the HA tag. Infiltration experiments using the CHLD fused to either Flag or HA had the same chlorotic phenotype.

3.2.2. Rescue of CHLD chlorosis by co-infiltration with CHLI

Because the current model for Mg-chelatase enzyme assembly uses the CHLD hexamer as a platform to build a double-hexameric ring with CHLI (Axelsson et al. 2006a; Willows et al. 2004), we decided to try a double infiltration comprised of both the CHLI construct and the CHLD construct to examine the effect of dual over-expression of these subunits.

This procedure has been demonstrated previously, with the expression of a chimeric antibody in tobacco (Vaquero et al. 1999), and we employed the technique to make a phenotypic assessment of the double infiltration. As was done previously, the transgenes to be co-expressed were present in different *Agrobacterium* cultures that were mixed prior to infiltration (Vaquero et al. 1999; Voinnet et al. 2003).

As a positive control for this experiment, a semi-dominant mutant construct of the Mg-chelatase I subunit (*Oy1-N1989*) was used to elicit chlorosis (generously provided by Thomas Brutnell) (Sawers et al. 2006a). This construct carries the mutation L176F, which results in strong chlorosis of the infiltrated region. The infiltrations were performed with a 1:1 ratio of CHLD to CHLI, each with an initial titer optical density (OD) of 0.2, as measured on the BioPhotometer (Eppendorf). Combining the infiltrates in this manner resulted in each having a final OD of 0.1, since each is effectively diluted by the co-infiltrate.

Because this reduction in CHLD infiltrate concentration could lead to a decrease in phenotype severity, the co-infiltration of CHLD and CHLI was compared to a

co-infiltration of CHLD and Empty vector (fig. 22C) prepared in the same manner.

As seen in figure 22B, the severity of wild type CHLD chlorosis is comparable to that of the semi-dominant CHLI mutant. However, co-infiltration of wild type CHLD with CHLI results in an almost complete rescue of the phenotype (fig. 22C).

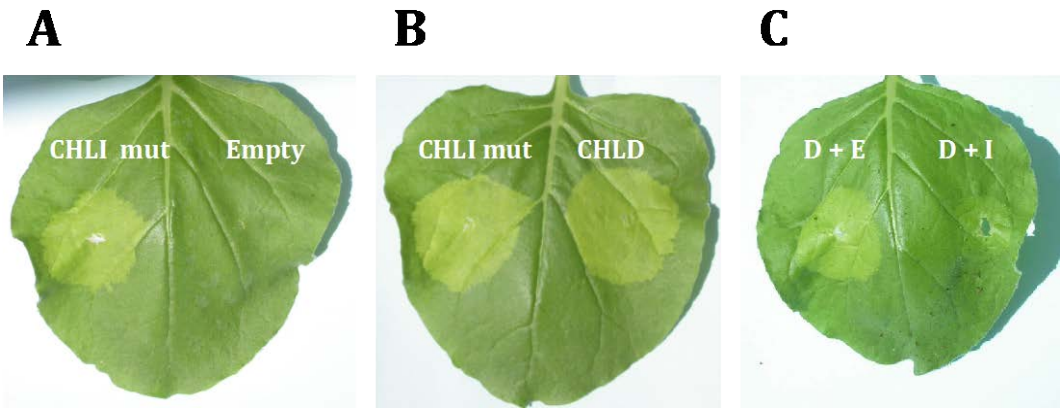


Figure 22. Phenotypes of CHLI- and CHLD-infiltrated leaf tissue

A. *Over-expression of the CHLI semi-dominant mutant results in severe chlorosis.* ***B.*** *Over-expression of wild type CHLD results in chlorosis similar to that of the CHLI mutant.* ***C.*** *The chlorosis seen with over-expression of CHLD is not appreciably reduced by co-infiltration with Empty vector, however, co-expression of CHLD with the CHLI subunit results in almost complete rescue of the chlorotic phenotype.*

3.2.3. Magnesium chelatase D truncation constructs

To elucidate the region of the CHLD responsible for the chlorotic phenotype, truncation constructs were generated as described in sections 2.2.3.1 and 2.2.3.1 (pp. 81-83). These included a construct comprised of the 5' region of the gene encoding the N-terminal half of the protein and ending with the poly-Proline region, an un-modified construct encompassing the 3' region beginning just after the poly-Proline linker region and extending to the end of the gene, and a third

construct encoding the C-terminal half of the protein, but with a transit peptide engineered onto the N-terminus (fig. 23D).

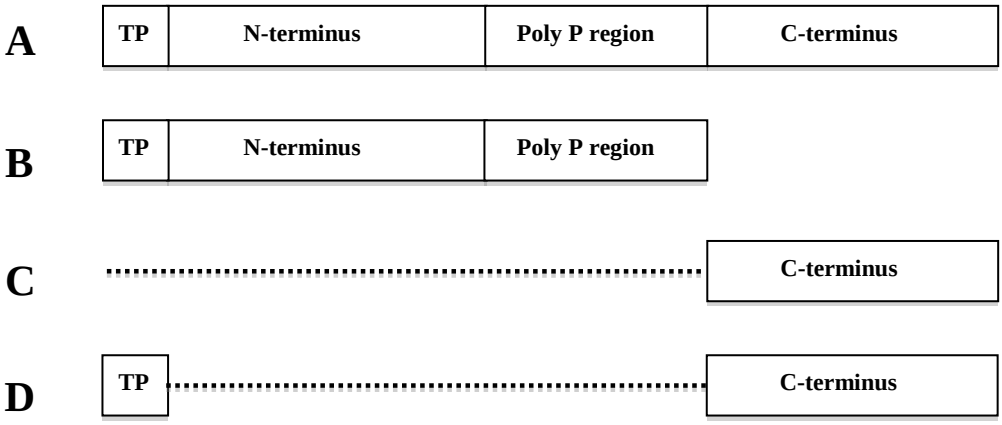


Figure 23. Mg-chelatase D truncation proteins

The truncation constructs used resulted in the production of protein fragments, as compared to the full-length protein depicted in A. The N-terminal region of the protein is shown in B, and includes the N-terminal half as well as the poly-proline region. The un-modified C-terminal region is shown in C, with the C-terminal region including a transit peptide shown in D.

Over-expression experiments were performed using the truncated regions of CHLD, alone or in conjunction with the CHLI subunit. Singular over-expression of the un-modified 3' construct resulted in a wild type phenotype. Infiltration of this construct was done mainly to ensure that over-expression of the un-modified C-terminus would not result in chlorosis, and after verification of this phenotype, the construct encoding the C-terminus with the engineered transit peptide (CTP) was used exclusively. Infiltration of the 5' construct (encoding the N-terminal half of CHLD) alone results in chlorosis (fig. 24A), while the CTP construct by itself generates a wild type phenotype (fig. 24A).

To investigate the interaction of the truncated proteins with the CHLI subunit, the CHLD N-terminal half was over-expressed along with CHLI, and the phenotype was compared to an infiltration comprised of the N-terminal, C-terminal and

CHLI constructs. While the chlorotic phenotype of the N-terminus was not rescued by co-infiltration with CHLI (fig. 24A), upon inclusion of the C-terminal construct, rescue was accomplished, as seen with the full-length CHLD (fig. 24B).

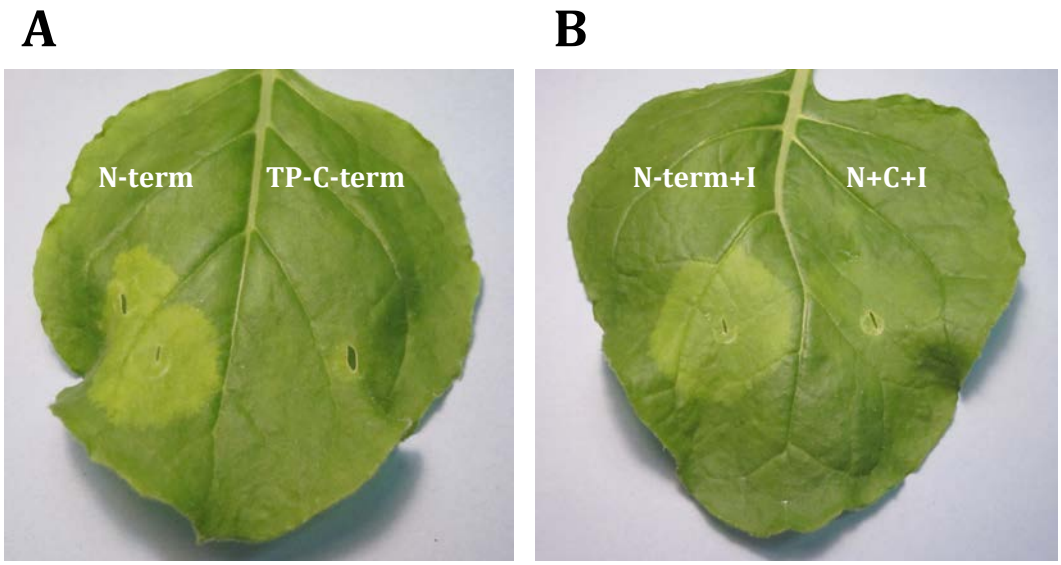


Figure 24. Phenotypic analysis of CHLD truncation constructs

A. *Over-expression of CHLD truncation proteins: the N-terminal half of CHLD (N-term) and the C-terminal half fused to the CHLD transit peptide (TP-C-term) result in chlorotic and wild type phenotypes, respectively.* ***B.*** *While co-expression of full-length CHLD and CHLI results in rescue of chlorosis (see figure 18C), co-expression of CHLI with the N-terminal half of CHLD does not result in rescue (N-term+I). Co-expression of both truncated halves of CHLD with CHLI, however, results in rescue comparable to that seen with full-length wild type CHLD (N+C+I).*

3.2.4. CHLD MIDAS mutant

Rescue of the chlorotic phenotype of Mg-chelatase D by co-expression of CHLI was an interesting result, and we wanted to investigate the CHLD:CHLI interaction further. To do so, we attempted to ascertain the region of CHLD responsible for the association with CHLI using a mutant CHLD construct.

The C-terminal region of CHLD contains a von Willebrand factor type A (VWA) domain (Whittaker and Hynes 2002). These types of domains are often protein-protein interaction sites in cell adhesion proteins, such as integrins and extracellular matrix proteins (Whittaker and Hynes 2002). In these cases, the interaction is mediated by the MIDAS motif (Michishita et al. 1993), which is defined by the amino acid sequence DXSXS… T… D, where X is any amino acid (Lee et al. 1995; Whittaker and Hynes 2002). To investigate the putative role of this domain, a mutant construct was generated from the wild type construct using PCR-based site directed mutagenesis to modify the MIDAS where the wild type DASGS amino acid sequence was mutated to DAAGS (fig. 25). The mutation changed the central Serine (an uncharged polar residue) to an Alanine (an uncharged hydrophobic residue).

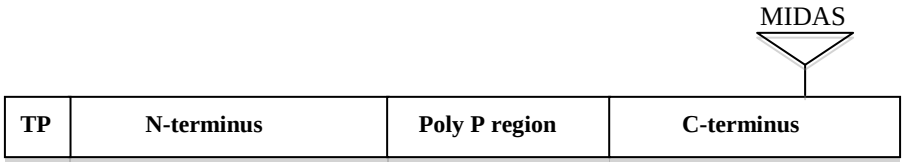


Figure 25. Magnesium chelatase D MIDAS mutant construct

The figure shows the location of the CHLD midas motif, with residue sequence DASGS. The constuct shown here, produced using PCR based directed mutagenesis, was engineered to produce a protein product with a central Alanine in lieu of a Serine residue.

3.2.4.1 Phenotypic analysis of the CHLD MIDAS mutant

Whereas the chlorotic phenotype of wild type Mg-chelatase D is rescued by co-infiltration with CHLI, rescue is inhibited when the central serine residue of the MIDAS motif in the C-terminal region is mutated (fig. 26).

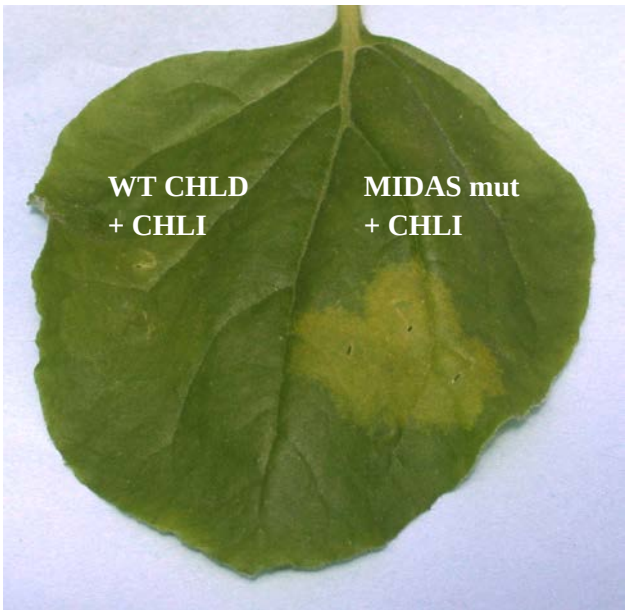


Figure 26. Co-infiltration of CHLD MIDAS mutant with CHLI

While the chlorotic phenotype elicited by over-expression of wild type magnesium chelatase D can be rescued upon co-expression with magnesium chelatase I, no such rescue is observed when CHLI is co-expressed with the MIDAS mutant of CHLD. This mutant has had a site directed mutagenesis performed to change the amino acid sequence of the MIDAS motif from DASGS to DAAGS, causing the polar central Serine to be changed to a hydrophobic Alanine residue.

3.3 Discussion

Over-expression of the I- and H-subunits of Mg-chelatase results in a barely discernible phenotype lighter than that of the surrounding leaf tissue, but visually identical to that seen with an Empty vector infiltration (fig. 21A and 21C). For that reason, we attributed this slight phenotype to the effect of an *Agrobacterium* infiltration, in itself, and scored it as wild type.

Unexpectedly, infiltration of the wild type CHLD construct resulted in severe chlorosis, such as that seen with the semi-dominant mutant of Mg-chelatase I (fig. 22B). Initially, we thought this chlorosis could be due to a sequence error generated in the cloning process, or perhaps a result of addition of the C-terminal T7 tag. The clone was therefore re-sequenced and verified as having a wild type amino acid sequence when compared to the original *Z. mays* mRNA sequence (GI# AY109815.1; see section 2.2.1). Furthermore, when independently generated CHLD clones were expressed with either a Flag or an HA tag, same severely chlorotic phenotype was observed. Since these tags are relatively small (table 2), we tentatively attributed the observed chlorosis directly to Mg-chelatase D over-expression.

Table 2. Protein epitope tags

Epitope Tag	Tag amino acid sequence
HA	YPYDVPDYA
T7	MASMTGGQQMG
Myc	EQKLISEEDL
Flag-6xHis	DYKDDDDKHHHHHH

Assembly of Mg-chelatase is thought to be accomplished utilizing a hexameric ring of Mg-chelatase D as a platform upon which the rest of the enzyme is built (Axelsson et al. 2006a; Hansson et al. 2002; Willows et al. 2004). Construction of the enzyme begins with the activation form, which is comprised of a double hexamer of six D- and six I-subunits. The transient expression system used in this study, which employs the p19 helper plasmid, has been demonstrated to express the transient proteins in excess of 50 fold higher than transient expression without

p19 (Voinnet et al. 2003). We postulated, therefore, that the CHLD over-expression could be leading to a surfeit of this subunit in the infiltrated region, resulting in an inability to assemble functional enzyme.

An extremely high level of CHLD would allow the formation of a plethora of D-hexamer platforms. However, because assembly of the activation form of the enzyme would rely, in part, upon a lower level of endogenously expressed Mg-chelatase I subunits, formation of functional proteins would be very low. In essence, the CHLD platforms may "mop up" the endogenous CHLI proteins, leading to the absence of functional Mg-chelatase, a breakdown in chlorophyll production, and ultimately, chlorosis.

Alternatively, the chlorosis seen with CHLD could have been due to either 1. steric interference of *any* tag fused to the C-terminus of the protein or 2. over-expression of CHLD leading to a retrograde signal to down-regulate chlorophyll production. To investigate these possibilities, co-expression experiments were performed to implement simultaneous expression of both CHLD and CHLI. The aim of these experiments was to rescue the chlorotic phenotype.

Because co-infiltration of separate *Agrobacterium* cultures carrying the constructs for both CHLD and CHLI would essentially dilute the final infiltrate, co-infiltration of CHLD with empty vector was performed as a control. As seen in figure 22C, while co-infiltration of CHLD/Empty still results in severe chlorosis, co-expression of CHLD/CHLI results in almost complete rescue. We therefore concluded that the chlorosis observed with over-expression of the CHLD subunit was not due to steric interference of enzyme assembly by the C-terminal tag, but rather as a result of the CHLD over-expression *per se*.

To further investigate the cause of CHLD elicited chlorosis, we generated truncation constructs of this protein to separate the N-terminal half from the C-terminal half, as described in sections 2.4.1 and 2.4.2. As seen in figure 16, the N-terminal half included the transit peptide, the poly-Proline region, as well as the sequence between these two regions. The construct for the C-terminal half of the protein was based on the region directly following the poly-Proline linker region, continuing to the stop codon. To enable correct compartmentalization of this truncated protein, the CHLD transit peptide was engineered onto the 5' end of the construct. In addition, the stop codon was ablated to allow expression of a C-terminal tag, which was encoded on the pBin61 vector. This was done to remain consistent with expression of the full-length protein, rather than for use of the tag in these experiments.

As depicted in figure 23A, over-expression of the N-terminal half of the protein resulted in a chlorotic phenotype, while over-expression of the C-terminal half of CHLD resulted in a wild type phenotype. To determine whether or not this phenotype could be rescued by co-expression of the CHLI protein, we performed experiments to over-express both the CHLD N-terminal half and CHLI. As seen in figure 23B, co-expression of the CHLD N-terminal truncation with CHLI does not result in rescue of the phenotype. However, when both truncation halves of CHLD are expressed with CHLI, rescue is re-established, and is visually comparable to the rescue seen with full-length CHLD (figs. 22C and 23B). These data imply that 1. while the chlorotic phenotype is generated by the N-terminal half of the CHLD protein, the rescue by CHLI is accomplished via the C-terminal

half of CHLD, and 2. the two halves of CHLD are either able to re-assemble into a functional subunit or they are acting independently of one another.

Rescue of chlorosis by the N-terminal half of CHLD by co-expression of the CHLD C-terminal half and CHLI led to an examination of the C-terminal region of CHLD. We were looking for a motif or moiety that could be responsible for the CHLD:CHLI interaction. The integrin I domain, with the MIDAS motif was a likely candidate, so we generated a mutant construct which changed the central serine of the MIDAS motif from a polar to a hydrophobic residue (DASGS to DAAGS). Chlorosis resulting from over-expression of this protein cannot be rescued by co-expression with CHLI, leading us to the conclusion that the integrin I domain, and specifically the MIDAS motif, is a key component to the interaction between CHLI and CHLD.

Interestingly, chlorosis resulting from over-expression of the CHLD MIDAS mutant *can* be rescued if a wild type form of the CHLD protein is included, implying either that the chlorosis is generated via a retrograde signal, or that the two halves of CHLD do, in fact, function independently of one another, as suggested earlier in this discussion.

Note that in this case the expressed proteins would include CHLI, CHLD with the mutated MIDAS, *and* an additional wild type CHLD C-terminal half. Phenotype rescue in the plant infiltrated with these constructs raises the question: how does the mutant half, which is attached to the rest of CHLD and wild type half (expressed independently) of CHLD work together without causing some kind of steric hindrance? Furthermore, in light of these data, what is the role of the N-terminal half of CHLD in formation of the activation complex?

3.4 Conclusions

In this chapter, we have demonstrated that over-expression of the Mg-chelatase D subunit results in a chlorotic phenotype. We have shown that this phenotype is not a result of steric hindrance of the C-terminal tag, but rather results from CHLD over-expression *per se*. This work does not resolve, however, the question of the basis for this chlorosis, which might result from either 1. The plant's inability to form functional Mg-chelatase enzymes, due to excess CHLD 'mopping up' endogenously produced CHLI, or 2. over-expression of CHLD results in generation of a retrograde signal which, in turn, elicits the phenotype.

We have shown that the chlorotic phenotype seen with CHLD over-expression can be rescued when the protein is co-expressed with the Mg-chelatase I subunit. Furthermore, the chlorotic phenotype appears to be derived from the N-terminal half (and/or the poly-Proline linker region) of the protein. Rescue of the chlorosis elicited from over-expression of the CHLD N-terminal half cannot be rescued when co-expressed with CHLI, however, rescue is achieved when the two halves of CHLD are co-expressed with the I-subunit. This implies that interaction of the CHLD and CHLI proteins involves the C-terminal half of CHLD (rescue of N-terminal half-derived chlorosis only occurs in the presence of the C-terminal half protein truncation). Specifically, the MIDAS seems to be integral to the interaction between CHLD and CHLI.

The mechanism of rescue is unclear. It could be that co-expression of CHLI provides the subunits necessary to establish complete activation forms of the enzyme, which then go on to interact with endogenously generated CHLH

proteins, thereby re-establishing chlorophyll biosynthesis. Alternatively, if over-expression of the CHLD subunit results in generation of a retrograde signal to elicit chlorosis, co-expression of the CHLI subunit could serve to somehow ablate this signal.

Chapter 4. Analysis of pigments and Magnesium Chelatase activity in infiltrated *Nicotiana benthamiana* leaf tissue

In this chapter, the relationship between plant phenotype and Mg-chelatase activity in infiltrated leaf tissue is investigated by performing pigment analysis using HPLC. These experiments were done, in part, to investigate the activity of Mg-chelatase in leaf tissues over-expressing either an individual subunit of Mg-chelatase, or a combination of subunits. Specifically, we were interested in comparing the relative activity of Mg-chelatase in leaf tissue over-expressing the CHLD subunit, which has a chlorotic phenotype, to leaf tissue co-expressing CHLD and CHLI, which is rescued back to a wild type phenotype.

In addition, we were interested in studying the effect of CHLH over-expression on enzymatic activity. It has been reported that Mg-chelatase H behaves co-operatively *in-vitro* (Jensen et al. 1998; Sawicki and Willows 2008). In other words, this subunit essentially behaves as a substrate for the enzymatic I:D double hexamer. Thus, the addition of CHLH subunits should, theoretically, enhance Mg-chelatase activity. To investigate this hypothesis, we analyzed the Mg-chelatase activity level of CHLH over-expressing leaf tissue.

Finally, we were interested in investigating the possible role of CHLH as an abscisic acid receptor, and its impact on chlorophyll biosynthesis. To this effect, we looked for a link between CHLH over-expression and carotenoid production.

4.1 Materials and Methods

4.1.1. Tissue collection

Plants were grown and infiltrated as described in sections 3.1.1 and 3.1.2. Timing of tissue collection was designed to precede the onset of chlorosis and thus circumvent any secondary effects of bleaching of the plastid. Transient green fluorescent protein (GFP) expression can be detected at day 3 with the aid of a UV lamp, and is detectable until approximately day 10. Likewise, in leaf tissue infiltrated with a construct generating chlorosis, the infiltrated region begins to pale at approximately 3 days after infiltration. To minimize secondary effects resulting indirectly from chlorosis derived from over-expression of Mg-chelatase subunits, tissue was harvested approximately 48 hours after infiltration.

4.1.2. Pigment isolation and HPLC analysis

Using a glass Tenbroeck tissue grinder, one square centimeter of fresh leaf tissue (0.022 – 0.028 g.) was ground in extraction buffer (0.6% Ammonia, Methanol) on ice in darkness, before transfer to a pre-labeled 1.5 mL microfuge tube. Samples were then centrifuged at 14k rpm at room temperature for 30 seconds. Supernatant was transferred to an amber sample vial, and 50 µl was injected into C18 reverse phase high performance liquid chromatography (HPLC) column.

A Shimadzu HPLC system was used at 1 ml/min with an Alltech C18 column (50-4.6mm), diode array detector and a Shimadzu RF-535 fluorescence detector. The HPLC protocol was modified from Sawicki and Willows (Sawicki and Willows 2007), and employed a gradient as follows: 0-10 min (0.15 M NH_4, pH

7.5) to (80% v/v Acetonitrile, 20% 0.15 M NH_4Ac, pH 7.5); 10-20 min linear gradient from 80% v/v Acetonitrile, 20% 0.15 M NH_4Ac, pH 7.5 to 50% Methanol, 40% v/v Acetonitrile, 10% 0.15 M NH_4Ac, pH 7.5, 20-40 min linear gradient from 50% Methanol, 40% v/v Acetonitrile, 10% 0.15 M NH_4Ac, pH 7.5 to 100% Methanol. Fluorescence excitation (Ex) was set at 410 nm, emission, (Em) was 630 nm, to detect Proto, and 595 nm to detect Mg-Proto. Standards for Proto IX and Mg-Proto eluted from the column at 17 and 16 minutes, respectively. HPLC analyses of *N. benthamiana* leaf tissue over-expressing single or combinations of Mg-chelatase subunits were performed with six biological replicates.

4.1.3. Porphyrin Extraction

Using 1 square centimeter of leaf tissue, samples were ground in a glass grinder with 1 ml porphyrin extraction solution (80% aqueous acetone; 22 mM Ammonia), prior to the addition of 200 µl Hexane. Next, 150 µl of the supernatant was diluted with 250 µl Hexane and used for spectrophotometric analysis to obtain readings at relevant wavelengths (443 nm chlorophyll a soret band; 453 nm chlorophyll b soret band; 490-500 nm carotenoid soret band; 637 nm Proto fluorescence; 650 chlorophyll a Qy band; 663 chlorophyll b Qy band). Spectrophotometry was performed using a Beckman Du64 Spectrophotometer.

4.1.4. Chloroplast isolation

Using 0.25-0.50 grams of infiltrated leaf tissue samples were gently ground in 5.0 ml isolation buffer (0.1 M Tricine-NaOH, 1 mM EDTA, 0.1% w/v BSA, 4mM

$MgCl_2$, 0.5M sorbitol, 1 mM DTT) with a mortar and pestle on ice. A glass tissue grinder was not used in order to preserve the integrity of the chloroplasts. The ground sample was then passed through a 100 micron filter before centrifugation at 4°C, 200xG for 30-60 seconds to collect particulate matter. The supernatant was transferred to a pre-labeled, weighed microfuge tube, and centrifuged for seven minutes at 2500g, 4°C. After carefully removing all supernatant, chloroplasts were weighed and re-suspended in an equal weight isolation buffer.

4.1.5. Flow cytometry

Chloroplast counting was performed using a Becton Dickinson FACS Caliber flow cytometer. To prepare samples for flow cytometry, serial dilutions were done to obtain a sample with barely detectable green color, typically approximately 1:20,000 v/v of chloroplast preparation to chloroplast isolation buffer. The dilution was performed because the flow cytometer is unable to measure a concentrated sample. Diluted sample (50 µl) was used for flow cytometry and the plastid count was used to determine the plastid concentration (in plastids/ml) of the original sample.

4.1.6. Magnesium chelatase assay

Samples were assayed for Mg-chelatase activity by first adding 20 µl of carefully measured chloroplasts to 180 µl reaction buffer (0.079 M Tricine NaOH, 0.794 mM EDTA, 0.397M Sorbitol, 0.794 mM DTT, 0.014 M $MgCl_2$, 6.67 mM Mg-ATP, 2 units creatine phosphokinase, 22.22 µM creatine phosphate, 6 µM Proto IX, 1.079% w/v BSA, 1.11 mM dipyridyl), and incubating reactions for 30

minutes at 30°C. To stop the assay, 800 µl 80%v/v Acetone 20% v/v 0.14M Ammonia was added to each sample. After mixing, 150 µl Hexane was added, and samples were mixed well before brief centrifugation and Fluorimetric measurement of the Acetone layer.

Immediately after stopping the assay, each sample (500 µl of the middle acetone layer) was measured using a Perkin Elmer L50 fluorimeter. Fluorescence was measured from 550 to 700 nm, with excitation set at 418nm. Intensity values were noted at curve minimum (approximately 572-575 nm), Mg-Proto maximum (595 nm), and Proto maximum (632 nm). Experimental sample peak values for Mg-Proto and Proto were determined by subtracting the minimum value from the values at 595 nm and 632 nm, respectively.

Assay control reactions were set up in tandem with the experimental reactions, but omitted the addition of Proto IX, and were measured immediately after set-up, without an incubation step. Peak values were obtained as described above, and gross activity was determined by subtracting each sample's control peak value from experimental sample peak value for Mg-Proto. Net activity for each sample was determined by dividing the gross activity by the number of plastids used in the assay (measurement obtained by flow cytometry, see section 4.1.5).

4.2 Results

4.2.1. HPLC Analysis of *Nicotiana benthamiana* leaf tissue overexpressing magnesium chelatase subunits

Leaf tissue over-expressing semi-dominant mutant Mg-chelatase I subunits (*Oy1-N1989, Oy1-N700)* (Sawers et al. 2006b) referred to here as pOY4 and pOY5, can

be expected to be deficient in Mg-chelatase activity, as the I-subunit acts as the enzyme motor (Lundqvist et al. 2010). Import of mutant CHLI subunits into the chloroplast and interaction with endogenous proteins is evidenced by the chlorotic tissue phenotype. In addition, tissues over-expressing the Mg-chelatase D-subunit may be impaired in Mg-chelatase activity due to an inability to form functional holo-enzymes. In either case, the deficiency would likely result in an accumulation of Proto, the substrate for Mg-chelatase. As an alternative to the possible detrimental effect postulated, expression of *Z. mays* Mg-chelatase D could result in enhanced Mg-chelatase activity, and perhaps in an accumulation of the enzyme's product, Mg-Proto, eliciting chlorosis via retrograde signaling.

HPLC-chromatogram data analysis was performed with a focus on the chlorophylls and carotenoids. In the analyses of the HPLC spectra, these pigments were expected to elute from the HPLC column in a predictable manner (fig. 27).

Trace data from HPLC were analyzed for the accumulation of Proto, Mg-Proto, chlorophylls a and b and carotenoids.

Identification of chromatograph peaks was based upon both retention time (fig. 27) and spectra (fig. 28), compared to reported pigment retention times and spectra (Jeffrey et al. 1997).

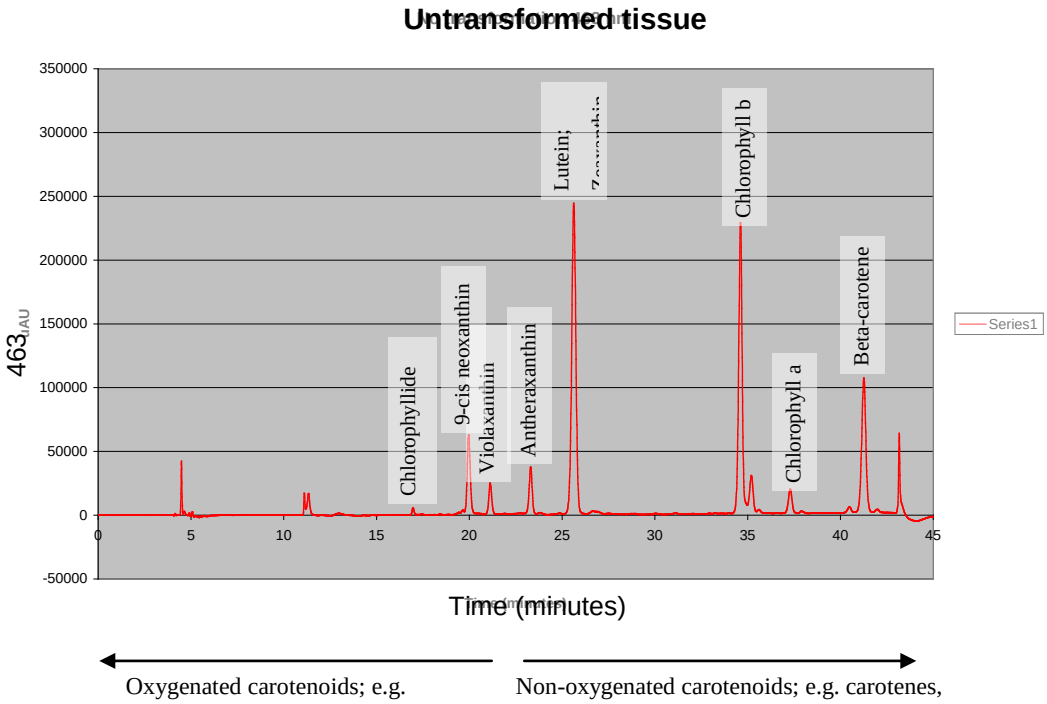


Figure 27. HPLC trace data – detail of uninfiltrated tissue

Polar molecules are generally eluted from the HPLC column in the earlier, more aqueous solution, while later eluting molecules are generally increasingly hydrophobic. Porphyrins without a phytol group will elute earlier than those with a phytol group, such as chlorophylls a and b. Oxygenated carotenoids, such as the xanthophylls and lutein will elute from the column earlier than non-oxygenated carotenoids such as α-carotene and β-carotene (ketones or epoxides).

Carotenoids Chlorophylls

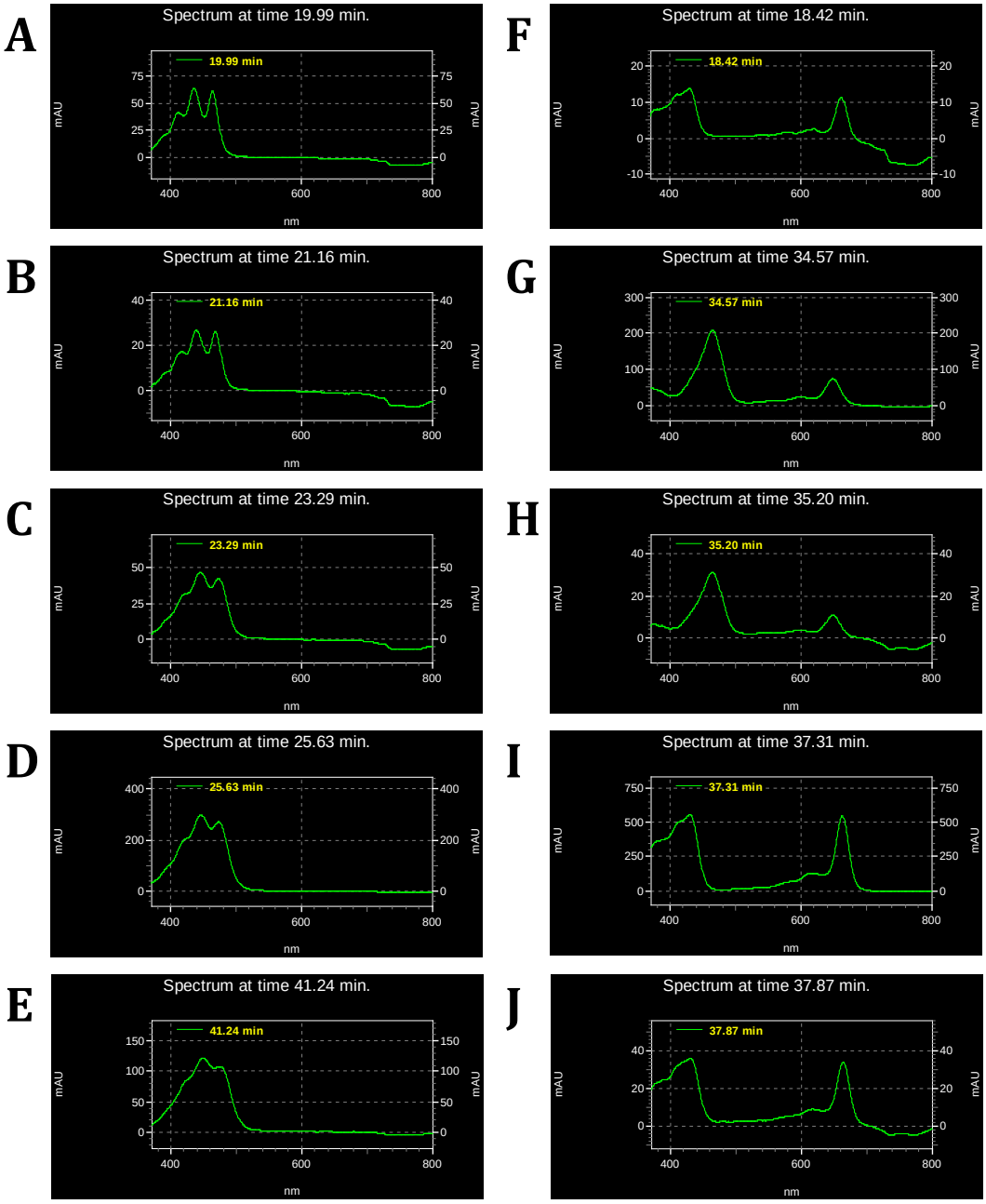


Figure 28. HPLC peak identification

The spectra shown in this figure are specific to the pigments eluted from the column at a given time. Identification of the spectra was based upon reported characteristics of relative retention time by reverse phase HPLC and comparison to published spectra (Jeffrey et al. 1997). The carotenoids were identified as follows: A. 9-cis neoxanthin, elution time 19.99 minutes; B. violaxanthin, elution time 21.16 minutes; C. antheraxanthin, elution time 23.29 minutes; D lutein, likely combined with zeaxanthin, as these do not separate easily, elution time 25.63 minutes; E. beta-carotene, elution time 41.24 minutes. The chlorophylls are: F. chlorophyllide a, elution time 18.42 minutes; G. chlorophyll b, elution time 34.57 minutes; H. chlorophyllin b, elution time 35.20 minutes I. chlorophyll a, elution time 37.31 minutes; J. chlorophyllin a, elution time 37.87 minutes. H and J are likely extraction artefacts

Although Proto was detected, there was no significant accumulation of this, or any other tetrapyrrole biosynthetic intermediate. In tissues exhibiting a chlorotic phenotype, chlorophyll degradation products were clearly seen, e.g. Chlorophyllide and unidentified peaks with a Qy absorbance. In all infiltrated tissue the 9-cis neoxanthin peak was reduced relative to the other carotenoids and extra unidentified peaks were observed in the HPLC traces as shown in figure 29. Interestingly, leaf tissue over-expressing Mg-chelatase H appeared to have enhanced carotenoid production relative to the chlorophylls, even though it did not display a chlorotic phenotype (fig. 29C).

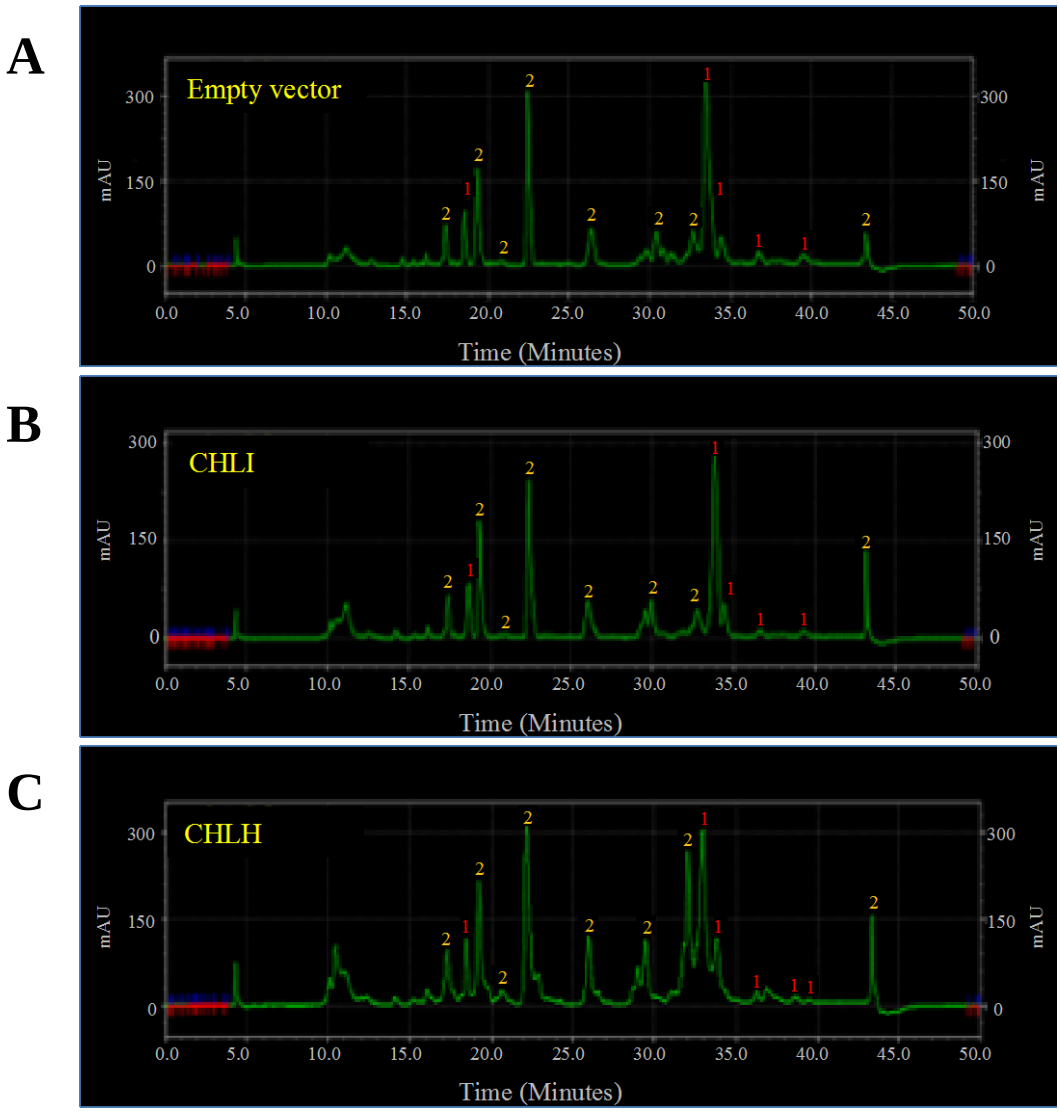


Figure 29. HPLC analysis of infiltrated tissues

This figure shows HPLC chromatograms for Nicotiana benthamiana *leaf tissue over-expressing* Zea mays *magnesium chelatase subunits CHLI and CHLH (B and C, respectively), compared to leaf tissue infiltrated with empty pBin61 binary vector (control tissue). Peaks for chlorophylls and carotenoids were identified based on spectra and retention times and are shown as having a red '1' and a yellow '2', respectively. With respect to the peak intensities, HPLC chromatogram derived from tissue over-expressing CHLI (B) shows very little difference to that derived from the control (A) while leaf tissue over-expressing magnesium chelatase H shows greater peak intensities for identified carotenoids. , depicted with , relative to the chlorophyll a and b peaks (C).*

4.2.2. Magnesium chelatase activity assay optimization

Development of the assay for Mg-chelatase activity in isolated chloroplasts was modified from a previously described method (Walker and Weinstein 1991). The experiments described here were scaled down significantly from the original protocols, and were therefore optimized for a smaller volume reaction.

The initial optimization tests included variations to the tissue sample/buffer ratios, additives such as DTT (a reduction agent) and dipyridyl, (an inhibitor of cyclase activity), Proto IX (substrate) concentrations (fig. 30), and Mg-ATP

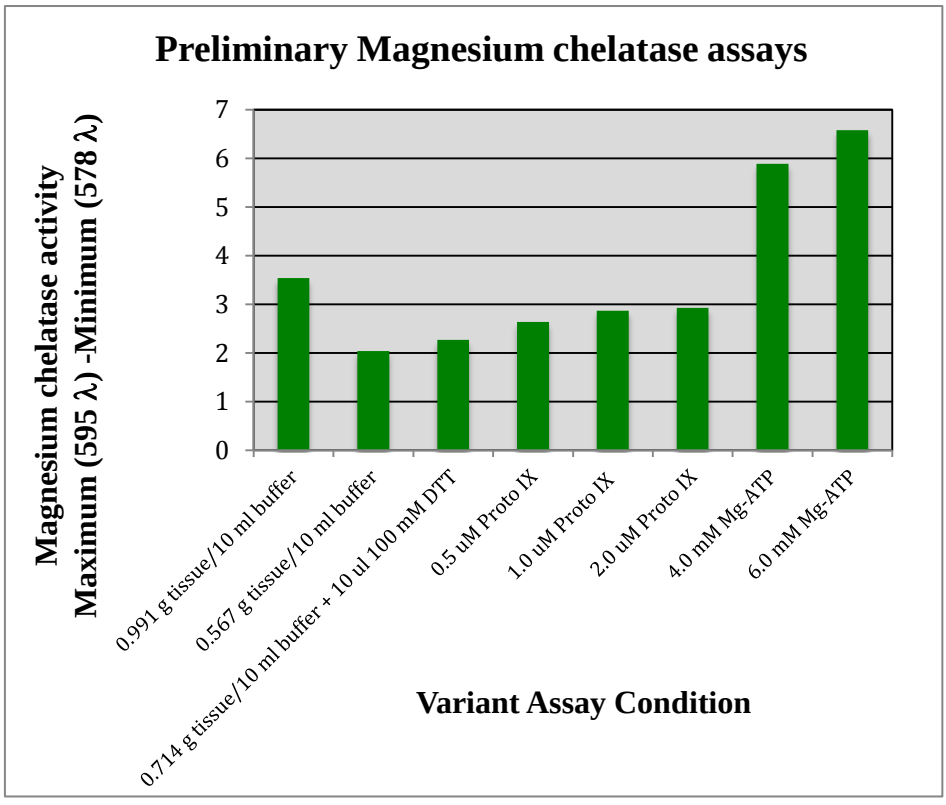


Figure 30. Mg chelatase assay optimization – preliminary condition variants

The initial tests for magnesium chelatase activity assay are shown, with variant plastid concentration (shown in grams of tissue used in the plastid prep/10 ml buffer), and substrate (Proto IX) as well as with the addition of DTT and, separately, Mg-ATP. These tests were performed only once, to identify factors that significantly increase the magnesium chelatase activity in the assay. Only the addition on Mg-ATP resulted in significantly increased activity, and thus further optimization of the assay focused on this component.

concentrations (figs. 30, 31). Of the different variables tested, the addition of Mg-

ATP had the greatest impact. In contrast, variations to tissue weight per 10 ml buffer, addition of DTT, and variations to the substrate (Proto) concentration did not appreciably increase the Mg-chelatase activity in the assay. Therefore, to optimize the assay, we focused on the Mg-ATP component.

The Mg-ATP concentration was found to be limiting up to approximately 10 mM (fig. 31). This was consistent with the results of Fuesler et al. (Fuesler et al. 1981). Thus, an ATP regenerating system was utilized in the final optimized assay, 0.1 U of Creatine Phosphokinase (CPK) and 20 mM phosphocreatine, along with 6 mM

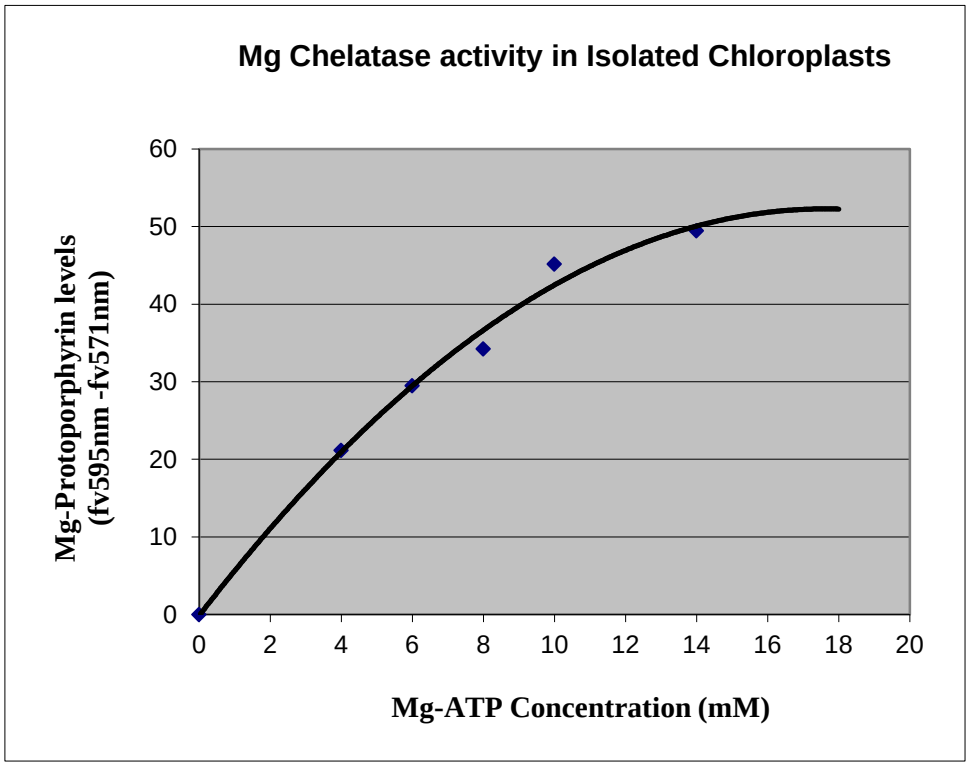


Figure 31. Mg Chelatase assay - ATP optimization

Assay optimization for Mg-ATP is shown here, with an almost linear trendline up to 10 mM Mg-ATP. The Mg-Proto levels were determined by measuring the fluorescence at 595 nm wavelength, less the baseline value measured at 571 nm wavelength.

ATP, This is consistent with the optimized conditions developed by Fuesler (Fuesler et al. 1981).

Assay optimization was tested in a time course. A large volume reaction was set up and sampled at 0, 5, 10, 20 and 30 minutes. Reaction samples were stopped and processed as described (see section 4.1.6). The optimized assay tested over full reaction run time (30 minutes) gives a plot with trendline R^2 value of 0.9905 (fig. 32)

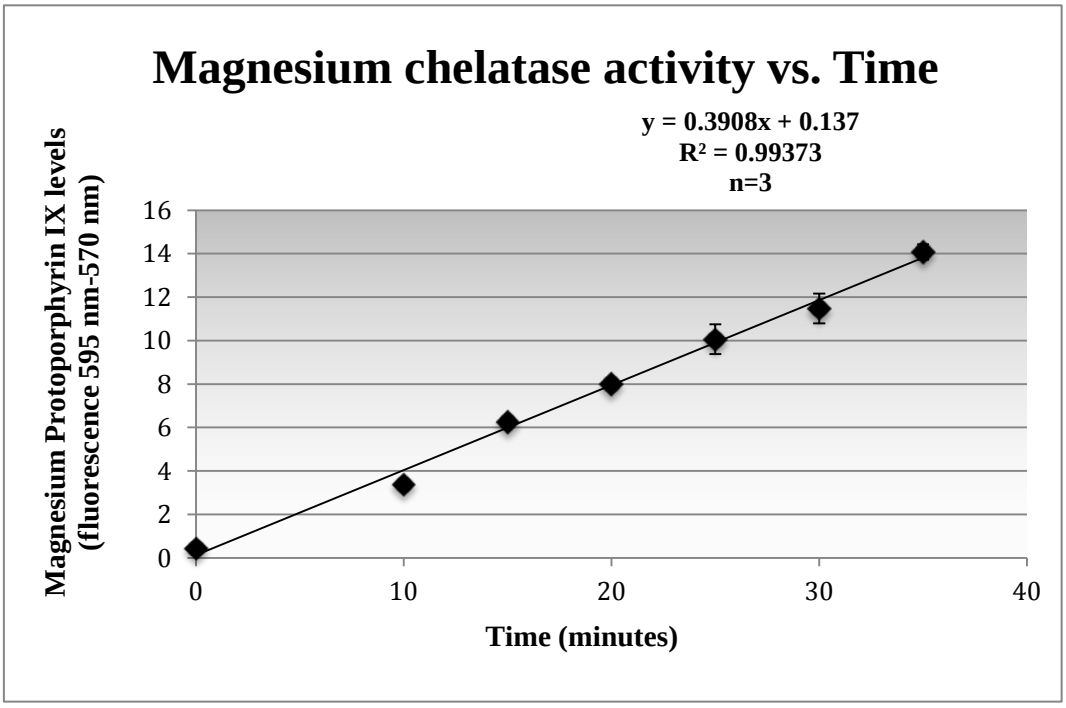


Figure 32. Optimized magnesium chelatase activity assay

The optimized magnesium chelatase activity assay was tested for reaction efficacy for the full duration of experimental run time (30 minutes), resulting in a trendline of plotted data with an R^2 value 0.9905, n=3.

4.2.3. Magnesium chelatase activity in isolated chloroplasts

Chloroplasts isolated from infiltrated tissues were assayed for Mg-chelatase activity. The CHLI semi-dominant mutant constructs pOY4 and pOY5 were used as a negative control for this assay.

As seen in figure 33, *Nicotiana benthamiana* tissue expressing the *Z. mays Oy1-N1989* and *Oy1-N700* mutants have between approximately 20 and 30% Mg-chelatase activity, compared to the Empty vector control. Furthermore, the CHLD subunit with mutated MIDAS domain (DM9) had activity comparable to the previously demonstrated null mutants (pOY4 and pOY5) (Sawers et al. 2006b), while tissue expressing the wild type CHLD protein had increased activity, relative to the mutants.

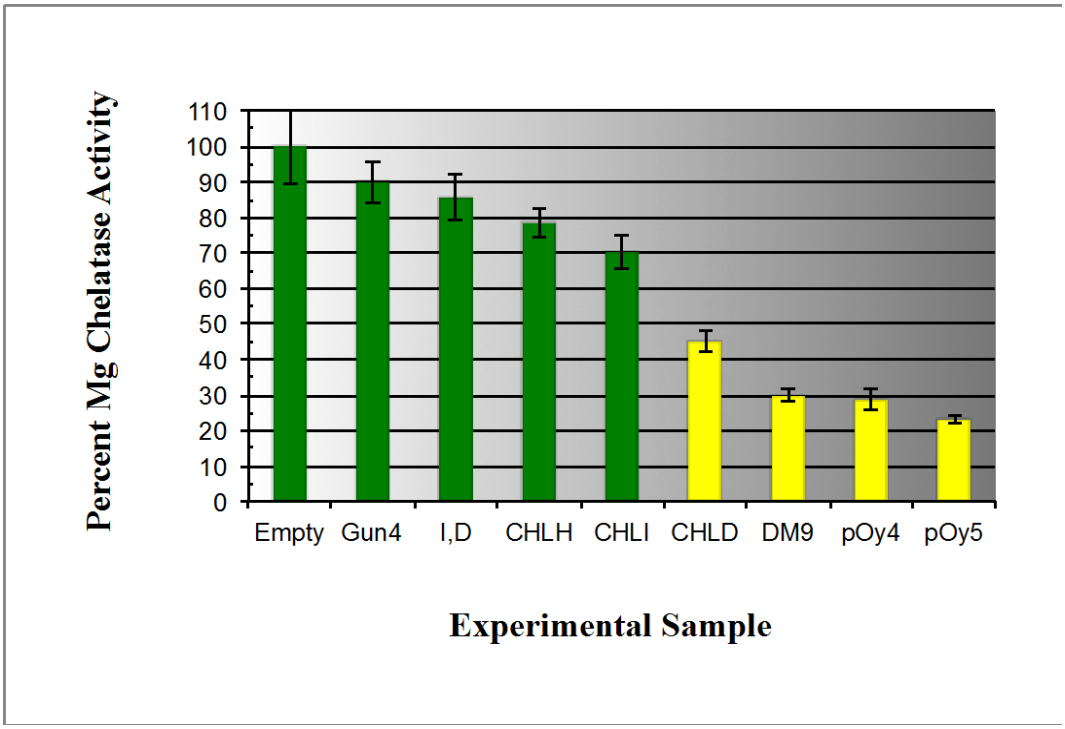


Figure 33. Magnesium Chelatase activity in *N. benthamiana* leaf tissue

This figure shows magnesium chelatase activity in Nicotiana benthamiana *leaf tissue expressing wild type and mutant magnesium chelatase subunits. Tissues having a chlorotic phenotype are depicted in yellow, while those with a wild type phenotype are shown in green. Note that leaf tissue expressing* Z. mays *wild type magnesium chelatase D has a chlorotic phenotype. Magnesium chelatase activity in tissue infiltrated with empty vector was 63 pmoles per 10 billion plastids per hour. Data shown is the average of 3 biological replicates (n=3). Error bars show standard error for each tissue type.*

4.3 Discussion

HPLC analysis was performed on leaf tissues expressing wild type subunits of Mg-chelatase. We were interested, specifically, in ascertaining the cause for the chlorotic phenotype elicited by the expression of CHLD. We hypothesized that the phenotype could be due to one of two things: 1. impaired or abolished Mg-chelatase activity derived from an inability to form functional enzymes, due to a detrimental excess of CHLD (too many scaffolds), or 2. generation of a retrograde signal derived from Mg-Proto accumulation; this accumulation, hypothetically, would be a result of accelerated Mg-chelatase activity due to an increase in CHLD concentration (plentiful scaffolds). According to Strand et al. (2003), accumulation of the tetrapyrrole intermediate Mg-Proto acts as a retrograde signal to the nucleus. To investigate if the CHLD chlorotic phenotype was brought about by a retrograde signal resulting from Mg-Proto accumulation, data derived from the HPLC were analyzed for accumulation of this intermediate, as well as various pigments found in leaf tissue. Failure to detect Mg-Proto accumulation might suggest that the chlorosis seen with CHLD expression can be attributed simply to a breakdown in chlorophyll biosynthesis.

HPLC analysis of *N. benthamiana* tissue expressing *Z. mays* Mg-chelatase subunits, specifically CHLD, shows no accumulated tetrapyrrole biosynthetic intermediate, including Mg-Proto (fig. 29B). After we did these experiments, however, evidence was published that, in fact, Mg-Proto accumulation is not involved in retrograde signaling (Mochizuki et al. 2008; Moulin et al. 2008). Therefore, we cannot rule out retrograde signaling as a cause for CHLD chlorosis.

Mg-chelatase activity assays were performed on isolated chloroplasts, and normalized based on plastid count. To do this, the chloroplast count of each sample was measured using flow cytometry and the activity per 10 billion plastids determined. To control for the potential effects brought about by infiltration with *Agrobacterium*, percent activity was normalized against the Empty vector control. Activity of the control tissue was determined to be 200 fluorescence units. Using a standard curve of Mg-Proto, the activity measurement was converted to a molar measurement of 63 pM. Thus, the activity seen in the Empty vector control is approximately 63 pmoles per hour per 10 billion chloroplasts.

Examination of enzyme activity in the DM9 CHLD mutant (mutation of MIDAS central Serine) expressing chloroplasts reveals a level of activity comparable to that of plastids expressing the semi-dominant CHLI mutant constructs, pOY4 and pOY5. *In vitro* analysis of *Synechocystis* CHLI mutants expressing *Oy1-N1989* (pOY4) and *Oy1-N700* (pOY5) showed that these mutations ablated Mg-chelatase activity (Sawers et al. 2006b). In contrast to the reported activity, these mutants show approximately 20-30% activity (fig. 31), compared to the control. Because we are unable to determine the exact rate of infection accomplished by *Agrobacterium* in these tissues, we speculate that the measured activity of the CHLI null mutants and potentially the DM9 mutant is due to Mg-chelatase activity accomplished by the chloroplasts of un-infected cells.

Due to the fact that CHLI functions primarily as a motor for the enzyme, it seems reasonable that over-expression of this subunit might result in an increased level of activity. The result was, in fact, the opposite. Expression of CHLI led to enzyme activity that measured approximately 70% that of the control. This

finding is in accordance with previously reported activity in transgenic tobacco plants over-expressing CHLI. In a study by Papenbrock et al. (2000), the activity of Mg-chelatase was diminished in plants both over-expressing and under-expressing CHLI.

Interestingly, the activity of CHLD expressing tissue was observed to be intermediate between the null mutants and the CHLI expressing tissue. To determine the reduction of activity needed to elicit a chlorotic phenotype, as is seen with CHLD expression, we compared the activities of chlorotic vs. wild type tissues. The activity of CHLD was measured at 45% that of the control, and CHLI expressing tissue was measured at 70%. Using the averaged activities of pOY4 and pOY5 (25%) as a base-line, and the activity of the Empty vector control infiltrated tissue as an upper limit, we determined that if the reduced activity of Mg-chelatase is solely responsible for the chlorosis seen in CHLD expressing tissue, then the enzymatic activity must be reduced to 27 to 60% that of the Empty vector control.

To investigate the phenotypic rescue of CHLD by co-expression with CHLI, we measured the activity in tissue co-infiltrated with the two subunits. The activity was found to be 85% that of the control, which is well above the estimated range necessary to maintain a wild type phenotype, based on the calculations resulting from the CHLD activity analysis.

A final observation of Mg-chelatase activity in the experimental samples is that, compared to the control tissue, they all demonstrated reduced Mg-chelatase activity, without exception. This was particularly interesting in the case of Mg-chelatase subunits I (which behaves as a molecular motor to drive enzyme

activity, as previously discussed) and H, which has been shown to act, essentially, as a substrate for the I:D double hexamer in *Rhodobacter* and *Synechocystis* (Reid and Hunter 2004; Sawicki and Willows 2008). Furthermore, the GUN4 protein has been shown to be a positive regulator of Mg-chelatase activity (Larkin et al. 2003). In light of these characteristics, one might expect leaf tissue over-expressing CHLI, CHLH, and GUN4 to have increased activity, compared to the control, rather than a reduction to 70%, 78%, and 90% that of control.

Based on this work, our conclusions are as follows. Over-expression of Mg-chelatase D does not result in an accumulation of Mg-Proto. Due to recent evidence that this molecule is not responsible for retrograde signaling, however, we cannot eliminate the involvement of a retrograde signal in the elicitation of this phenotype.

Furthermore, based on our data, perturbation of the correct stoichiometric ratio of Mg-Chelatase appears to impair catalytic efficiency of the enzyme pool. It seems that CHLI is not simply recruited as a molecular motor, and CHLH may not simply behave as a substrate, but rather they must be present at the correct ratio to promote the most efficient activity of the enzyme.

Chapter 5. iTRAQ Analysis of CHLD infiltrated tissues – over- and under-expressed proteins

In the previous transient expression experiments, CHLD was seen to elicit a chlorotic phenotype. Although we have demonstrated that CHLD over-expression leads to impaired Mg-chelatase activity, we were not able to rule out the involvement of a signal in the generation of this phenotype.

The experiments described in this chapter were performed to facilitate comparison of protein expression levels from leaf tissue over-expressing CHLD, CHLI and CHLH. The aim of this work was to ascertain what kind of changes were brought about by over-expression of these proteins, and specifically, to help determine if the phenotype seen with over-expression of the CHLD subunit was a result simply of a breakdown in chlorophyll biosynthesis, or if there is some kind of signaling involved.

We did a proteomic comparison of CHLI, CHLD and CHLH over-expressing material compared to wild type, using iTRAQ isobaric tags. The iTRAQ labeling method, along with HPLC-MS/MS allows multiplexed quantitation of peptides via reporter ion signals in the low mass region of MS/MS and quantitation of changes in protein expression levels by comparison of the experimental samples (Casado-Vela et al. 2010).

5.1 Materials and Methods

5.1.1. Plant tissue production

Plants were grown, and infiltration experiments performed as described previously (See sections 3.1.1 and 3.1.2). Directly after infiltration, the perimeter of the infiltrated area was lightly marked using a waterproof marker. This was done immediately, because directly after infiltration the injected area is clearly discernable as darker, saturated leaf tissue. The visible saturation disappears approximately 20 minutes post-inoculation, and as many infiltrations had wild type phenotypes, this was the most straightforward way of differentiating between infiltrated and un-infiltrated tissue on a leaf when harvesting.

Due to the fact that CHLD over-expressing tissue begins to show chlorosis approximately 3 days after infiltration, with full chlorosis at approximately 4 days, tall tissues were harvested 2 days after infiltration, to capture the proteomic changes brought about by the treatment, and circumvent any potential secondary effects brought about by the advent of chlorosis. Samples were cut from the area outlined in permanent marker, and were weighed to measure approximately 0.1 – 0.15 grams of tissue. Harvested tissue was placed in prepared foil packets, and the weight of each was recorded on the packet prior to flash freezing in liquid nitrogen.

5.1.2. Protein extraction

Total protein was extracted from leaf tissue using an initial trichloroacetic acid (TCA)-acetone precipitation, followed by a phenol extraction. Care was taken not to introduce any primary amines at any time in the protein preparation, as primary amines can react with iTRAQ reagents, competing with peptide derivitization.

5.1.2.1 TCA-Acetone precipitation

The protein extraction employed was modified from Wang et al. (2003). Using samples of similar weights, leaf tissue was first ground in liquid nitrogen and transferred to pre-labeled chilled 2.0 microfuge tubes containing 1.0 – 2.0 ml cold 100% Acetone. Samples were then mixed thoroughly for 30 seconds. In all wash steps, samples were re-suspended into the wash reagent, and the pellet was broken up, Mixing was done without vortexing, to prevent oxidation of the samples. After brief centrifugation at 4°C, 10,000 x g, samples were re-suspended in cold 100% acetone, mixed well and re-pelleted for three minutes at 4°C, 10,000 x g, and supernatant discarded before air drying.

Next, samples were ground in a mortar and pestle with a small amount of acid-washed quartz sand, washed in 10% TCA/Acetone three to four times, until the supernatant was colorless, and re-suspended in cold aqueous 10% TCA. The samples were centrifuged and washed again in cold aqueous 10% TCA, prior to being washed twice in cold 80% acetone/water. After drying the pellet at room temperature, the pellets were stored at -80°C or used directly for the final, phenol extraction.

5.1.2.2 Phenol extraction

Using similar weights for each sample, 0.05 – 0.1 g dried leaf tissue powder was re-suspended in buffer containing 0.8 ml tris-buffered phenol (pH 8.0; Sigma-Aldrich) and 0.8 ml dense SDS buffer (30% sucrose, 2% SDS, 0.1 M Tris-HCl, pH 8.0, 5% beta-mercaptoethanol). After thorough mixing, the samples were centrifuged for 3 minutes at 10,000g.

The upper phenol phase of each sample was then divided into three 200 µl aliquots, with careful attention not to disturb the SDS interphase. After visual verification that each aliquot was clear (not cloudy), a minimum of 5 volumes (1 ml) cold 100% methanol was added to each tube. Samples with cloudy aliquots were combined and re-centrifuged until clear aliquots were obtained. After the addition of methanol, samples were mixed and incubated at 20°C for 30 minutes, then centrifuged for 5 minutes at 10,000 x g. Samples were washed twice with 100% cold methanol, then twice with 80% aqueous acetone. All residual liquid was removed, but the pellet was not dried further, to facilitate re-suspension. For each sample, one aliquot was used for quantification, one aliquot was used for the iTRAQ experiment, and the final aliquot was held in reserve, as a backup for aliquots 1 and 2. Aliquot 1, for protein quantification, was re-suspended in Urea/CHAPS buffer (5M urea, 2M thiourea, 65mM DTT, 2% CHAPS, 2% sulfobetaine 3-10, 40 mM tris, 1% carrier ampholytes). Aliquot 2, for iTRAQ, was re-suspended in 0.5M triethylammoniumbicarbonate (TEAB), 0.1% SDS, pH 8.5.

5.1.3. Protein quantification

Because aliquots were made with equal volumes of Phenol supernatant during the phenol protein extraction, there are equivalent protein concentrations in each aliquot. The separate aliquots were generated in this manner to facilitate protein quantification using Bradford assay, which may be inaccurate in the presence of TEAB/SDS buffer. Therefore, aliquot two, for iTRAQ experimentation, was re-suspended in TEAB/SDS, while aliquot one was re-suspended in CHAPS/Urea buffer, and used in a standard Bradford assay (Bradford 1976). Because of the potential difference in protein solubility in the two buffers, the protein quantity in aliquot two was then verified using gel electrophoresis.

5.1.3.1 Bradford assay

For each sample, quantitation of aliquot one (re-suspended in CHAPS/Urea buffer) was performed using a standard Bradford reagent (BioRad; Hercules, CA). The Bradford reagent was diluted 1:5, as per the manufacturer's instructions, and 200 µl diluted reagent added to each well in a 96-well microtiter plate. One microliter of water was used for the negative control, while 1 microliter of each BSA standard used to generate a standard curve, and 1 microliter of each unknown used for protein quantification. Each was repeated at least twice, and measurements obtained using a plate reader (Thermo Scientific; Waltham, MA) were averaged.

Averaged measurements of BSA standards were then used to generate a standard curve (using a straight line of best fit) and the curve was used to determine the

protein quantities for aliquot two (sample in TEAB/SDS buffer for iTRAQ experimentation) of each sample.

5.1.3.2 Gel electrophoresis

Using the quantitation obtained with aliquot one, each sample was then loaded into NuPAGE 4-12% precast Bis-Tris gels (Invitrogen; Carlsbad; CA) as follows: for each gel, a sample with know protein quantity (CHAPS/Urea buffer) was loaded into lanes 1 and 2 at a concentration of 5 µg and 15 µg, respectively. Each sample to be used for iTRAQ analysis (TEAB/SDS buffer) was then loaded into two of the following lanes at a concentration of 5µg and 15µg, respectively, based on the concentration of the corresponding Bradford assayed samples (aliquot one, CHAPS/Urea buffer), and gel was run at 200v (fig. 34).

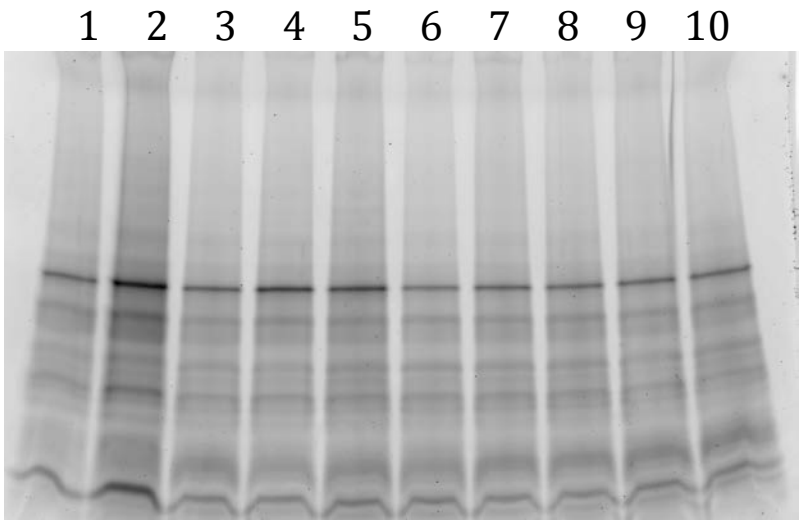


Figure 34. Protein quantification gel for iTRAQ analysis

Lanes 1 and 2 show 5 µg and 15 µg of control sample protein measured using a Bradford assay. Lanes 3 and 4 are loaded with 5 µg and 10 µg of protein extracted from leaf tissue infiltrated with empty vector, lanes 5 and 6 are loaded with 10 µg and 5 µg of protein from CHLI expressing leaf tissue (the volumes were accidentally transposed), lanes 7 and 8 show protein extracted from CHLD infiltrated leaf tissue, and lanes 9 and 10 are loaded with 5 µg and 10 µg of protein from CHLH infiltrated leaf tissue

Electrophoresed gels were incubated in (colloidal Coomassie dye/10% Methanol) for at least one hour, and de-stained in water prior to scanning. Due to variations introduced by the lighting, imaging on a gel documentation system causes inconsistencies that make it impossible to use ImageJ analysis software. Therefore, the gels were run on a Typhoon Trio laser scanner (GE Healthcare; Waukesha, WI), to give the best representation of the dark-intensity on the gel.

The gels were scanned in black and white to minimize any error that could be introduced during a secondary conversion to black and white prior to ImageJ gel analysis. Images were then analyzed using ImageJ software (Miller 2007), which measures the intensities of the bands, allowing sample quantitation.

5.1.4. iTRAQ labeling

Each sample was quantified and 100 µg aliquoted into a fresh microfuge tube. The samples were then reduced and alkylated, digested with trypsin, and labeled for iTRAQ using iTRAQ Multiplex kits (Applied Biosystems; Foster City, CA).

5.1.4.1 Reduction and Alkylation

Reduction was accomplished using tris-(2-carboxyethyl) phosphine (TCEP). To each sample, 2 µl freshly made up (in water) 100 mM TCEP was added. Samples were vortexed and briefly centrifuged to collect volume at the bottom of tube, and then incubated at 60°C for one hour with gentle shaking. After another brief centrifugation, alkylation was performed by adding one µl of freshly made 200

mM Methyl methane Thiosulfonate (MMTS), and incubating the samples at room temperature for 10 minutes.

5.1.4.2 Sample digestion and drying

Trypsin enzyme was prepared immediately prior to digestion by adding 100 µl ultrapure water to a new vial of lyophilized trypsin (Promega; Madison, WI) to obtain a final trypsin concentration of 1 mg/ml. The trypsin was then briefly vortexed and 20 µl added to each sample. The samples were incubated at 37°C overnight.

To maximize labeling efficiency, the sample volumes must be less than 50 µl for a 4-plex experiment and less than 33 µl for an 8-plex experiment. Therefore, after overnight digestion, samples were dried in a centrifugal vacuum concentrator, and reconstituted with 30 µl 0.5M TEAB buffer.

5.1.4.3 iTRAQ labeling and verification

Protein labeling with iTRAQ dyes was performed according to the manufacturer's recommendations using iTRAQ multiplex kits from Applied Biosystems (Foster City, CA). Each vial of iTRAQ reagent was allowed to come to room temperature before the addition of ethanol to obtain at least 70% ethanol, v/v. Each vial was then vortexed to mix the reagent with the alcohol, and the solution in each vial was transferred to a sample tube. The tubes were incubated at room temperature for one hour, whereupon the reaction was quenched by adding 100 µl Milli-Q water to each sample tube.

Efficacy of the iTRAQ labeling was verified by matrix assisted laser desorption/ionization time of flight/time of flight (MALDI TOF/TOF). To do so, we combined 2µl of each sample and the mixture was desalted, concentrated and purified using a C-18 tip according to the manufacturer's instructions (Eppendorf). To use the C-18 tip, the sample solutions were acidified by adding 3 µL of 1% trifluoroacetic acid (TFA) to 8µl of sample prior to loading on the C-18 tips. Peptides were eluted with 5 mg/mL α-cyano-4-hydroxycinnamic acid (CHCA) in 70% ACN/0.1% FA and spotted in duplicate on the MALDI target.

Spectra were obtained by the Australian Proteome Analysis Facility (APAF), in reflector positive mode on the matrix assisted laser desorption/ionization time of flight mass spectrophotometer (MALDI TOF/TOF; Applied Biosystems 4700) within a mass range of 750-3500 Da, a focus mass of 2100 Da, and using 2400 laser shots/spectrum. Fragmentation data were acquired on the top 10 intense ion signals using collision-induced dissociation conducted at 1 kV using "atmospheric" collision gas and a total of 7500-9000 laser shots. Spectra were externally calibrated with angiotensin I (*m/z* 1296.5), bradykinin (*m/z* 1060.2), ACTH 18-39 (*m/z* 2465.7), and neurotensin (*m/z* 1672.9). Baseline correction and a 5-point Gaussian smooth were performed on peaks using the Data Explorer software version 4.6 (Applied Biosystems). Ions having a signal-to-noise ratio (S/N) above 10 were included in the APAF-generated peak list for each sample preparation. Finally, APAF examined the reporter mass ion region in the MS-MS spectra to verify labeling.

We then pooled and dried all samples before submission to the APAF for strong cation exchange (SCX) fractionation and liquid chromatography mass

spectrometry (LC-MS-MS) as described previously (Song et al. 2008), and in detail below.

5.1.4.4 Strong Cation Exchange Chromatography

In the service provided by APAF, iTRAQ labelled peptides were fractionated by strong cation exchange liquid chromatography (SCX) using a PolySulfethyl A 200 mm × 2.1 mm, 5 µm, 200 Å column (PolyLC, Columbia, MD). Buffer A was 5 mM phosphate 25% (v/v) acetonitrile, pH 2.7, and buffer B was 5 mM phosphate, 350 mM potassium chloride, 25% (v/v) acetonitrile, pH 2.7. The dried iTRAQ labelled sample was resuspended in buffer A and applied to the SCX column. After sample loading and washing with buffer A for 37.5 min, buffer B concentration increased from 10% to 60% in 46 min and then ramped to 100% at a flow rate of 200 µL/min. Fifteen fractions were collected from the SCX separation then dried by centrifugal evaporation.

5.1.4.5 Reverse Phase NanoLC ESI MS/MS

In the reverse phase nano-liquid chromatography, also performed by APAF, dried SCX fractions were solubilized in 100 µL of 0.1% (v/v) formic acid, 2% (v/v) acetonitrile. Thirty microliters of sample was loaded onto a reverse phase peptide Captrap (2 µg capacity) (Michrom Bioresources, Auburn, CA) and desalted for 10 min with buffer A (0.1% (v/v) formic acid) at 10 µL/min. After desalting, the trap was switched in-line with a 150 µm × 10 cm C18, 3 µm, 300 Å ProteoCol column (SGE). The buffer B (90% (v/v) acetonitrile, 0.1% (v/v) formic acid) gradient started from 5% to 10% in 2 min and then to 50% in 80 min to elute peptides.

The LC eluent was subject to positive ion nano flow electrospray analysis using a QStar Elite MS/MS system (Applied Biosystems) in an information dependent acquisition mode (IDA). In IDA mode, a TOFMS survey scan was acquired (*m*/*z* 350−1600, 0.5 s), with the three most intense multiply charged ions (counts >25) in the survey scan sequentially subjected to product ion analysis. Product ion spectra were accumulated for 2 s in the mass range *m*/*z* 100−2000 with a modified Enhance All mode Q2 transition setting favouring low mass ions so that the reporting iTRAQ ion (114, 115, 116 and 117 *m*/*z* for 4-plex experiment or 113, 114, 115, 116, 117, 118, 119, 121 *m/z* for 8-plex experiment) intensities were enhanced for quantitation. Dynamic exclusion was used with a 2 min and 150 ppm window.

5.1.5. Data analysis

5.1.5.1 Analysis of Protein Pilot results

As part of the APAF service, MS/MS data were analyzed using ProteinPilot v3.0 (Applied Biosystems), which uses the Paragon algorithm (Shilov et al. 2007) to perform database matching for protein identification, protein grouping to remove redundant hits, and comparative quantitation. The resultant list of protein identifications was forwarded to me for subsequent analysis.

We assessed protein identifications (IDs) according to a set of criteria imposed to insure statistical significance. IDs with statistics falling outside of the criteria were eliminated from the data set. Criteria for this process are described here.

The NCBI Green plant reference protein set from 2010 protein database was used for initial searches. Additional searches also used a database constructed from all

Nicotiana benthamiana protein sequences, *Nicotiana sp.* protein sequences from NCBI where no *Nicotiana benthamiana* proteins were available, and the *Zea mays* Mg-chelatase proteins being infiltrated. This database had 498 entries and is included in the supplementary DVD.

Data were normalized for loading error by bias corrections calculated using ProteinPilot. All reported data were based on 95% confidence for protein identification as determined by ProteinPilot. A further requirement was a protein *p*-value, which ensured protein identification and quantitation was not based on a single peptide hit. The default ion intensity threshold in ProteinPilot for calculating peptide ratios was 40 counts.

Differentially expressed proteins were determined using the following criteria. First, proteins identified using non-unique peptides were eliminated (those with an unused value less than 1.3). Next, the error factor value (EF) was assessed. The EF assigned by Protein Pilot is determined by various criteria, including the variability in ratios of peptides used to identify a single protein, as well as the signal intensity of reporter ions. Ratios with an error factor of two or less are considered highly confident (Griffin et al. 2007). Therefore, proteins with an EF value over 2 eliminated. Identified proteins were then assessed according to p-value. Those IDs with a p-value of higher than 0.05 were eliminated so at to exclude individuals identified with less than 95% certainty, using the null hypothesis.

To determine under- or over expression of the identified proteins, we considered the iTRAQ ratios. These ratios represent differential protein expression in the experimental tissue (labeled with 115, 116, or 117 Da isotope tags) compared to

the control tissue (labeled with the 114 Da isotope tag). For example, our tissues were labeled as follows: leaf tissue infiltrated with empty vector (the control) was labeled with the 114 Da tag, while tissues over-expressing CHLI, CHLD, and CHLH were labeled with molecular weight tags of 115, 116, and 117 Da, respectively. To ascertain the differential expression levels of proteins detected in, for example, CHLD infiltrated tissue, the ratio 116:114 was examined. Protein identifications with a ratio over 1.2 were considered to be over-expressed, compared to wild type, while ratios less than 0.83 were considered to be under-expressed. Proteins with ratios lower than 1.2 and higher than 0.83 for all three experimental tissue types were considered to have no change in expression level, and were therefore eliminated from further analysis. Finally, proteins identified with 0% coverage (95) were eliminated. The remaining proteins were annotated, mapped and graphically represented using the Blast2GO software package (Conesa and Gotz 2007; Conesa et al. 2005).

5.1.5.2 Blast2GO analysis

Blast2GO is a software program that classifies proteins into gene ontology categories. Such categorization facilitates proteomic analyses by giving a broad overview of the biological schemes, such as protein function or biological pathway, in which the affected proteins are categorized. Data analysis proceeds through several stages, described here.

In the initial stage, amino acid sequences of the proteins identified and analyzed with Protein Pilot (section 5.3.5.1) are used in a blast search. The search identifies sequences similar to the query sequence(s), and assigns gene ontology

(GO) terms based on the similarity percentages (Conesa et al. 2005). The software then uses the blast results to perform a mapping step (described below). Because e-values are dependent upon database size (Altschul et al. 1990), the annotation of query sequences will ultimately be based on similarity percentages, which are independent of database size (Conesa et al. 2005).

Next, the mapping step is performed. In this step, three strategies are used to retrieve GO terms associated with blast search hits. The first strategy employs mapping files provided by NCBI (gene2acession and geneinfo files) to retrieve gene names from basic local alignment search tool (BLAST) hit accessions. In the second strategy, GI identifiers are used to retrieve UniProt identifications from the PIR (non-redundant reference protein)-database mapping file for each blast result. Finally, in the third strategy, accessions resulting from the blast are searched directly using the GO database DBXRef table (Conesa and Gotz 2007).

In the annotation step, functional terms are assigned to query sequence based on the pool of GO terms gathered in the mapping step. The algorithm used for blast2GO annotation considers multiple factors, including the similarity between query and hit sequences, the quality of the source of GO assignments, and the structure of the GO direct acyclic graphs (DAGs) (Conesa and Gotz 2007). Finally, any remaining protein matches with zero percent coverage were eliminated.

Due to the fact that the Blast2GO software annotation occurs on multiple levels (proteins are categorized with a range of functions, from broad to specific), the annotation was finalized with a manual curation step. In this step, each protein identification was examined and assigned a single functional category, based upon

the Blast2GO annotation results. This step was performed to give a more accurate graphical representation of the data, without inflating the number of sequences assigned to more general categories, such as ‘metabolic process’, due to multiple annotations of a single protein. The relative distributions of classified data set proteins were then visualized in graphical representations, with each protein represented only once.

5.2 Results

Proteomic analyses of *N. benthamiana* tissue over-expressing a single subunit of Mg-chelatase (CHLI, CHLD or CHLH), compared to *N. benthamiana* leaf tissue infiltrated with empty pBin61 vector, were performed as described in the materials and methods section. Although one 4-plex and two 8-plex experiments were performed using three separate biological replicates, due to quality issues with the 8-plex reagents, only the 4-plex experiment yielded high quality statistically significant data. The results presented here, are therefore preliminary, as the replicates are incomplete. Additional experimental replication was not done due to financial constraints.

Preliminary analysis of the iTRAQ data was performed on a smaller database consisting of the *Zea mays* proteins infiltrated into the tissue, all 460 *N. benthamiana* protein sequences in the NCBI database, as well as PSI, PSII and tetrapyrrole biosynthetic proteins from other *Nicotiana* sp. not represented in the *N. benthamiana* sequences. This allowed confirmation of *Zea mays* protein expression in the infiltrated tissue and analysis of the relative levels of these proteins compared to the endogenous *Nicotiana* proteins.

The data shown in Table 3 indicates that the *Zea m*ays proteins were expressed. Also, at the time this sample was taken the expression level was 5.7 times higher for *Zea mays* CHLI over the endogenous *Nicotiana benthamiana* CHLI protein, and 3.6 times higher for the *Zea mays* CHLD protein over the endogenous *Nicotiana* CHLD protein. Since there were no specific *Zea mays* peptides identified for the CHLH subunit, the slight increase in the quantitation indicated

for the shared peptide implied some expression of this protein, but at a lower level than the CHLI and CHLD proteins.

Table 3. 4-plex experiment. Analysis of *Zea mays* protein expression relative to endogenous protein. Peptide confidence above 90%

iTRAQ Label	114	115	116	117
Infiltrated *Zea mays* gene	Empty vector	ChlI	ChlD	ChlH
Number of *Zea mays* specific peptides detected (total spectra)	-	29 (221)	2 (5)	0
Quant. relative to 114 ions*	N/A	18.6	14.4	-
Number of *Nictotiana* ChlI, ChlD or ChlH specific peptides (total spectra)	N/A	14 (87)	1 (1)	3 (9)
Quant. relative to 114 ions	N/A	1.0	0.95	1.1
Shared CHLI, CHLD or CHLH specific peptides (total spectra)	N/A	1(11)	5 (12)	1 (2)
Quant relative to 114** (from shared spectra)	N/A	5.7	3.6	1.25

** These values have not been background subtracted and are dependent on the absolute concentration of the protein.*

*** These values give an estimate of the expression of the* Zea mays *proteins relative to the endogenous* Nicotiana benthamiana *protein.*

**** The data for this table were generated using trans-proteomic pipeline with XTandem as the search engine. Data shown were at 90% confidence interval, which corresponded to a false discovery rate of less than 2%.*

In the overall analysis of the 4-plex data, 88404 spectra passed spectra quality checks in protein pilot and were used in searches against the databases indicated. In searches against the small database 16873 spectra identified 4237 unique peptides below a global false discovery rate of 1%, which represented a protein pilot peptide confidence score of 87.6%. These peptides identified 162 proteins in the database of 498 proteins at a false discovery rate below 1%, which was at a protein pilot protein confidence score of 99%.

Searches against the larger, non-redundant green plant database were used in further analyses employing filtering criteria described in the materials and methods section. This resulted in the identification of 1173 proteins. After

statistical analysis of the data as described, tissues over-expressing CHLI, CHLD, and CHLH were found to differentially express 21, 117, and 54 proteins, respectively.

Assignation of GO-terms for differentially expressed proteins was accomplished using Blast2GO, as described in the materials and methods section. Distribution of GO-terms can be found in Appendix B, with differentially expressed protein identifications for CHLI, CHLD, and CHLH listed in appendices C, D and E, respectively. The Blast2GO analysis allows the classification of proteins according to biological process, facilitating grouping of functionally related proteins. Using Blast2GO, graphical representation of differentially expressed proteins in a given tissue type (e.g. CHLD over-expressing leaf tissue) enables visualization of over- and under-expressed protein distribution resulting from experimental treatment (e.g. over-expression of Mg-chelatase D-subunit). In this way, we can determine which biological pathways or functions, if any, are affected by the treatment.

Furthermore, the Blast2GO, generates graphical representation of the data using cutoffs for the annotation score or for the number of sequences. This feature facilitates examination of dataset annotation encompassing a large number of proteins. Due to the relatively low number of differentially expressed proteins in our experimental tissues, we chose correspondingly low thresholds for node populations, so as avoid too broad a categorization of the results. Cutoff values for graph nodes are therefore based on the number of proteins under- or over-expressed in CHLI, CHLD and CHLH infiltrated tissue.

5.2.1. Differential protein expression in CHLI over-expressing leaf tissue

Over-expression of CHLI had very little effect on global protein abundances. In total, 21 proteins were detected as being differentially expressed, with 11 under-expressed and 10 over-expressed. Details of the specific proteins differentially expressed in CHLI over-expressing leaf tissue can be found in Appendix C.

5.2.1.1 Under-expressed proteins in CHLI infiltrated tissue

The proteins detected as being differentially under-expressed in CHLI over-expressing leaf tissue are classified into seven gene ontology categories, with a minimum node population of one, and the highest node comprised of 3 individuals (fig. 35).

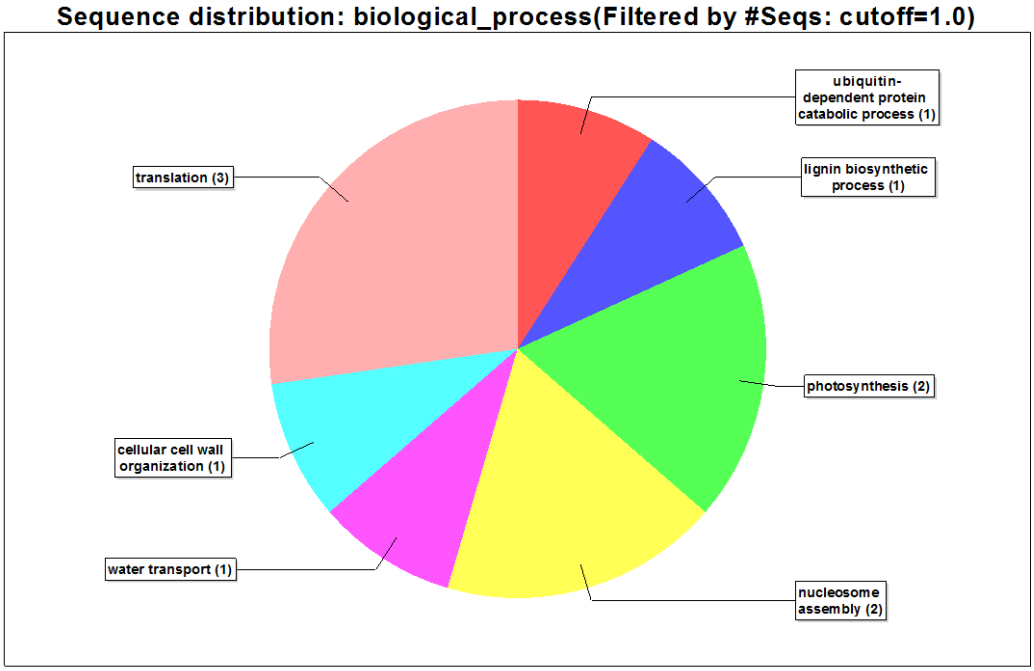


Figure 35. CHLI infiltrated tissue – differentially under-expressed proteins

Classification using gene ontology (GO) terms of under-expressed in CHLI infiltrated tissue, compared with wild type. GO term assignation was accomplished using the blast2go software package; graph nodes were defined by a minimal population of 1 sequence.

5.2.1.2 Over-expressed proteins in CHLI infiltrated tissue

As seen with the under-expressed proteins in CHLI infiltrated tissue, there are few detected differentially over-expressed proteins. Eight categories of GO terms are depicted in figure 36, with the largest category (translation) being comprised of two terms.

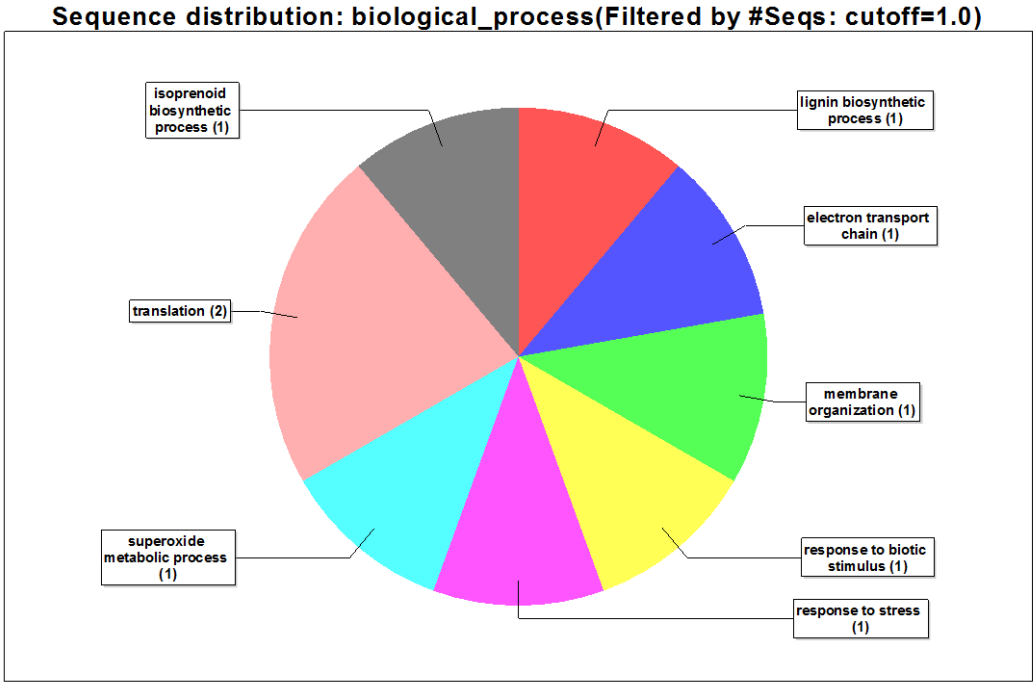


Figure 36. CHLI infiltrated tissue – differentially over-expressed proteins

Classification using gene ontology (GO) terms of over-expressed in CHLI infiltrated tissue, compared with wild type. GO term assignation was accomplished using the blast2go software package; graph nodes were defined by a minimal population of 1 sequence.

5.2.2. Differential protein expression in CHLH over-expressing leaf tissue

Over-expression of the CHLH subunit of Mg-chelatase resulted in detection of 54 differentially expressed proteins, with 19 proteins under-expressed and 35 proteins over-expressed. Details of specific differentially expressed proteins can be seen in Appendix E.

5.2.2.1 Under-expressed proteins in CHLH infiltrated tissue

Under-expressed proteins detected in CHLH over-expressing tissue have been grouped into 11 classifications (fig. 37). Of these, the processes of translation, photosynthesis, and nucleosome assembly have node populations of 3 or more

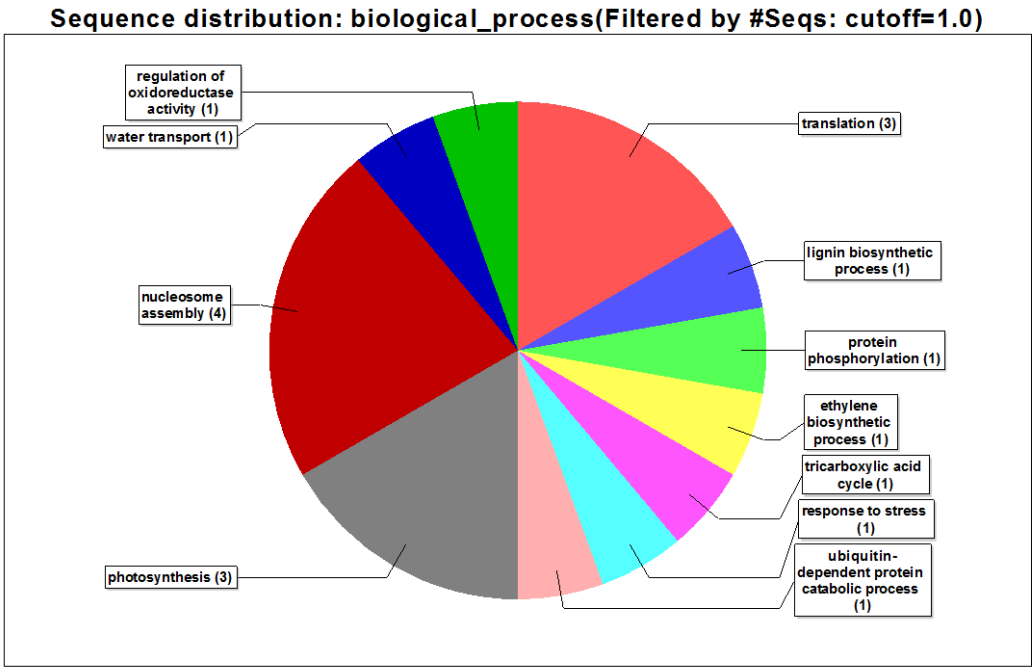


Figure 37. CHLH infiltrated tissue - differentially under expressed proteins

Classification using gene ontology (GO) terms of under-expressed in CHLH infiltrated tissue, compared with wild type. GO term assignation was accomplished using the blast2go software package; graph nodes were defined by a minimal node population of 1 sequence.

proteins. The proteins grouped into the ‘translation’ class are all cytosolic ribosomal proteins, while the ‘nucleosome assembly’ category consists of histones (h2, 2a, h3, h4). The proteins with the gene ontology term ‘photosynthesis’ are RBCS (Rubisco small subunit; nuclear encoded), LHCB3 (CAB; nuclear encoded), and photosystem I p700 apoprotein (PSAA; chloroplast encoded).

Also over-expressed were chloroplast protein 12, which is involved in the regulation of oxidoreductase activity, and a 14-3-3 protein.

5.2.2.2 Over-expressed protiens in CHLH infiltrated tissue

Of the over-expressed proteins detected in CHLH infiltrated tissue, 31% were classified as translation proteins, while 20% were allocated to photosynthesis (fig. 38).

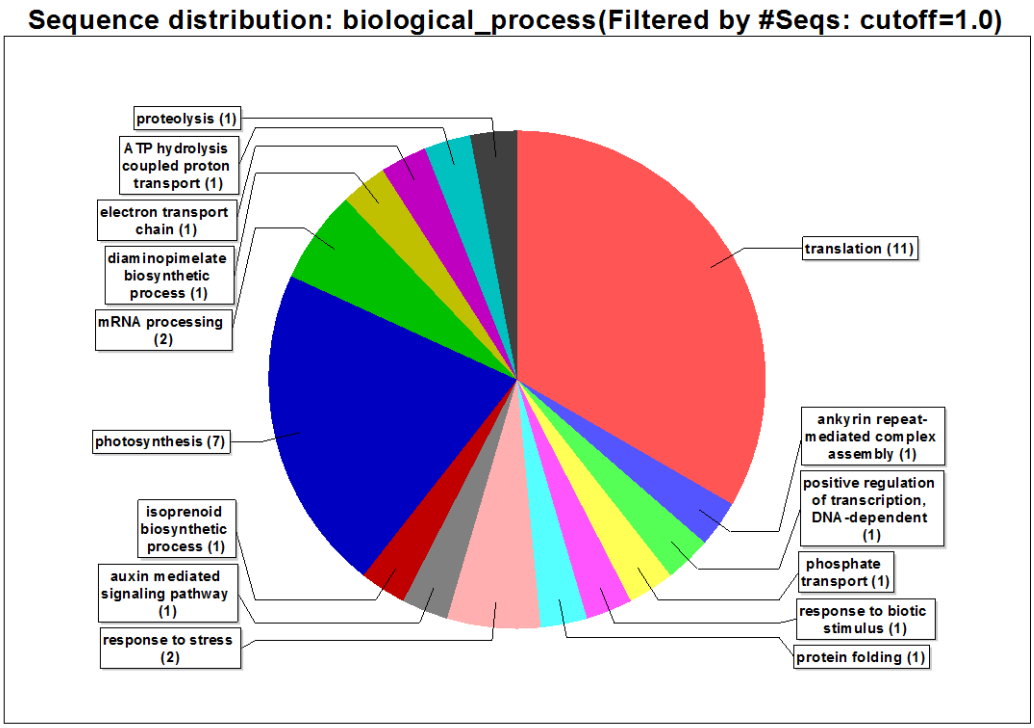


Figure 38. CHLH infiltrated tissue – differentially over-expressed proteins

Classification using gene ontology (GO) terms of over-expressed in CHLH infiltrated tissue, compared with wild type. GO term assignation was accomplished using the blast2go software package; graph nodes were defined by a minimal node population of 1 sequence.

5.2.3. Differential protein expression in CHLD over-expressing leaf tissue

In CHLD infiltrated tissue 117 proteins were detected as being differentially expressed. Of these, 45 were under-expressed and 72 were over-expressed. Details of all the proteins differentially regulated in CHLD infiltrated tissue can be found in Appendix D.

5.2.3.1 Under-expressed proteins in CHLD infiltrated tissue

Twelve gene ontology categories were generated using a node cutoff of 2 (fig. 39). Among these, translational elongation had the highest node population. Interestingly, many under-expressed proteins appear to be related to regulation or translation.

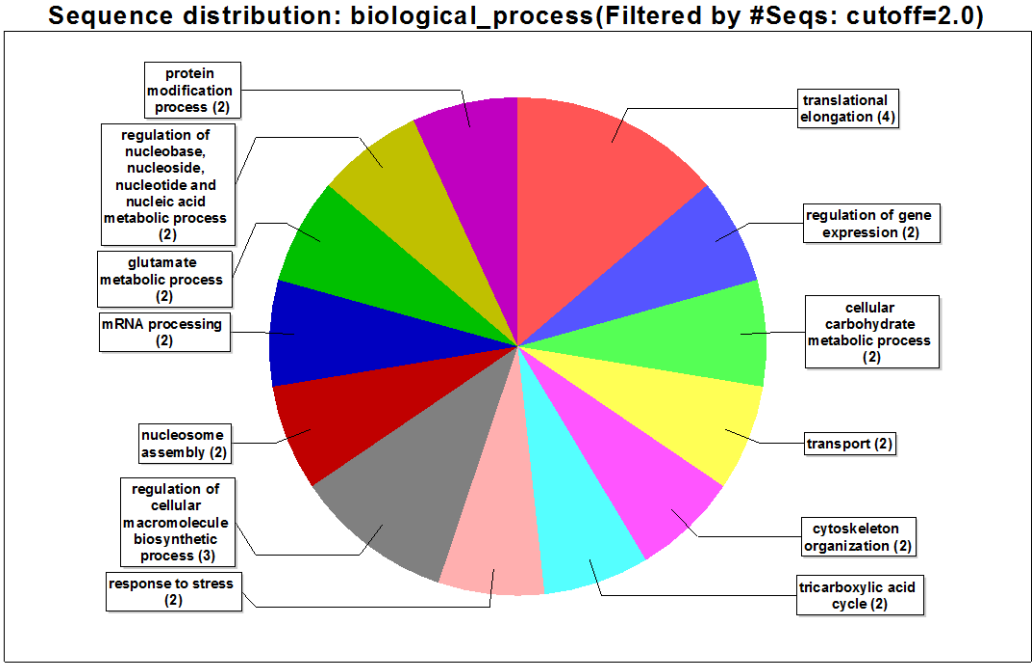


Figure 39. CHLD infiltrated tissue - differentially under-expressed proteins

Classification using gene ontology (GO) terms of under-expressed in CHLD infiltrated tissue, compared with wild type. GO term assignation was accomplished using the blast2go software package; graph nodes were defined by a minimal node population of 2 sequences.

5.2.3.2 Over-expressed proteins in CHLD infiltrated tissue

Fourteen categories of proteins were differentially over expressed in CHLD infiltrated leaf tissue (fig. 40). The biological system most dramatically affected was photosynthesis, with transport and generation of precursor metabolites and energy having similar node populations.

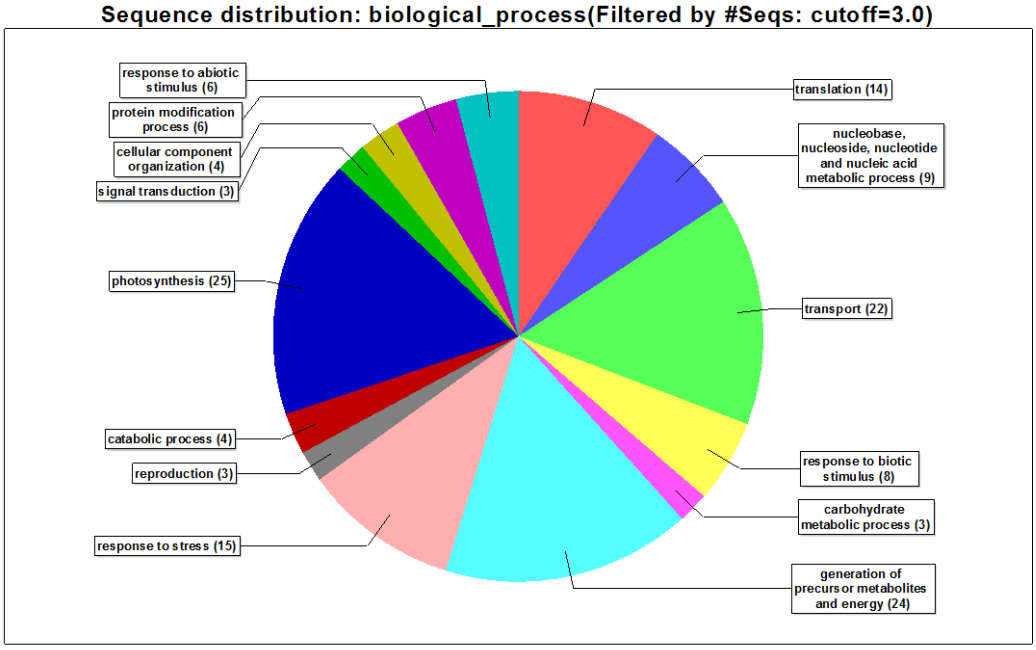


Figure 40. Over-expressed protein distribution in CHLD infiltrated tissue

Classification using gene ontology (GO) terms of over-expressed in CHLD infiltrated tissue, compared with wild type. GO term assignation was accomplished using the blast2go software package; graph nodes were defined by a minimal node population of 3 sequences.

5.3 Discussion

Using iTRAQ analysis, the protein profiles of *N. benthamiana* tissues over-expressing one of the subunits of Mg-chelatase, CHLI, CHLD, or CHLH were compared to that of leaf tissue infiltrated with empty vector. Tissue infiltrated with empty vector was used as a negative control in lieu of un-infiltrated tissue to account for effects brought about by the process of *Agrobacterium* infiltration.

Results for proteins detected as differentially expressed show that over-expression of CHLI or CHLH, compared to CHLD over-expressing tissue, does not result in changes to an overall metabolic process. In contrast, CHLD over-expressing leaf tissue exhibited over-expression of a plethora of photosynthetic proteins (figs. 41-44). These unexpected results are especially notable when the phenotypes of the leaf tissue are considered, as, while CHLI or CHLH over-expression results in leaf tissue with a wild type phenotype, over-expression of the CHLD subunit results in chlorosis (fig. 21). Because many of the proteins differentially over-expressed are nuclear encoded, and in chlorotic tissue should be down-regulated via a chloroplastic signal to the nucleus, these data imply a breakdown in retrograde signaling.

The most dramatic result obtained from the iTRAQ experiment was over-expression of many of the components of the photosynthetic machinery. In photosystem II, the main proteins of the reaction center were over-expressed, along with components of the light-harvesting antenna (fig. 41). Included in the

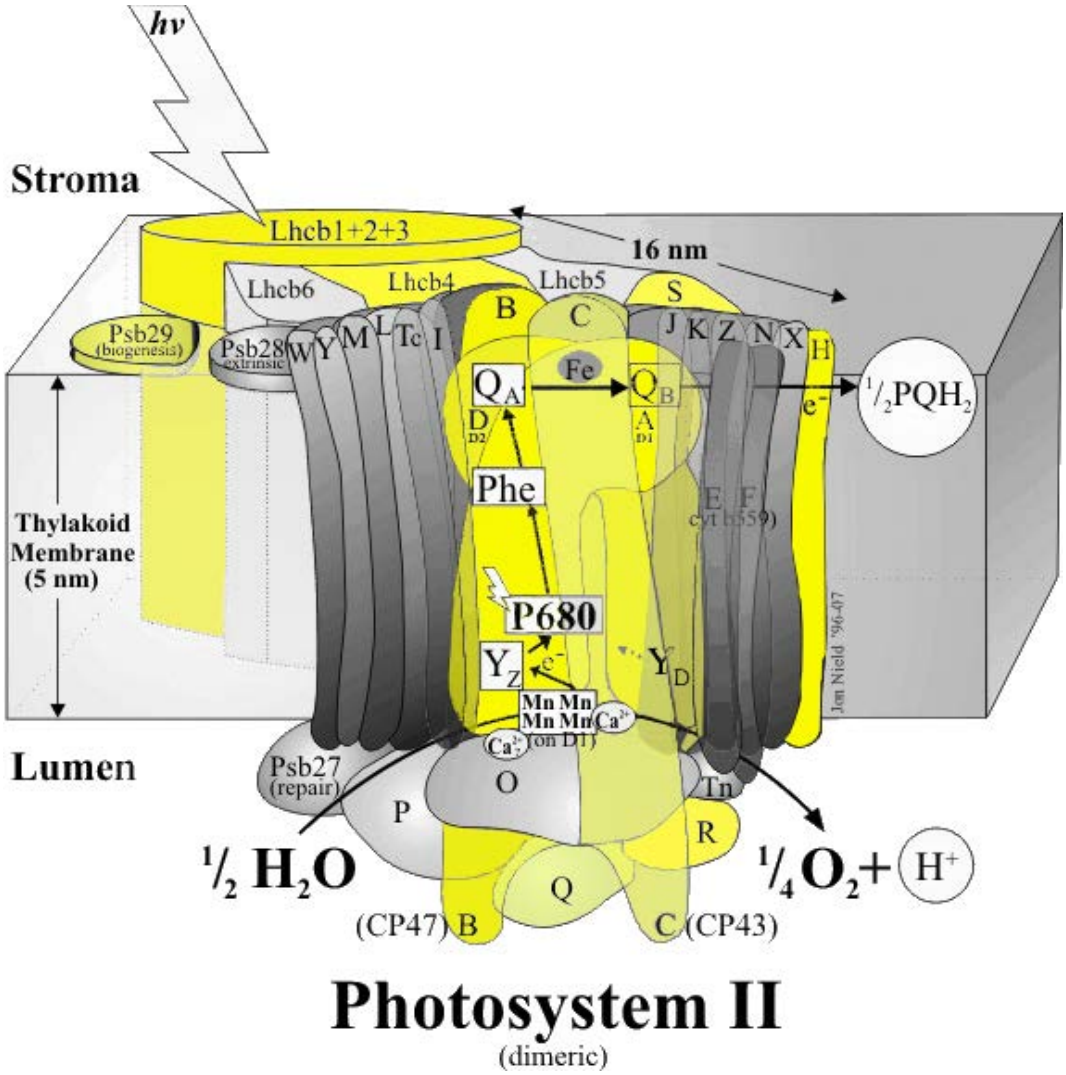


Figure 41. Over-expressed PSII proteins in CHLD infiltrated leaf tissue

In this figure PSII proteins over-expressed in CHLD expressing tissue are shown. Included are PSBA, PSBB, PSBC, PSBD, PSBH, PSBQ, PSBR, PSBS, LHCB1, and LHCB4.

Modified from www.jonnield.com/en/psiiimages/oxygenicphotosynthmodel.html

PSII over-expressed proteins are PSBA (D1), PSBB (CP47), PSBC (CP43), PSBD (D2), PSBQ, and PSBR, as well as LHCB1, LHCB4, and PSBS.

Over-expressed proteins of the cytochrome b_6f complex include plastocyanin, cytochrome f, cytochrome b_6 and the cytochrome b_6f complex iron-sulfur protein (ISPa; fig. 42)

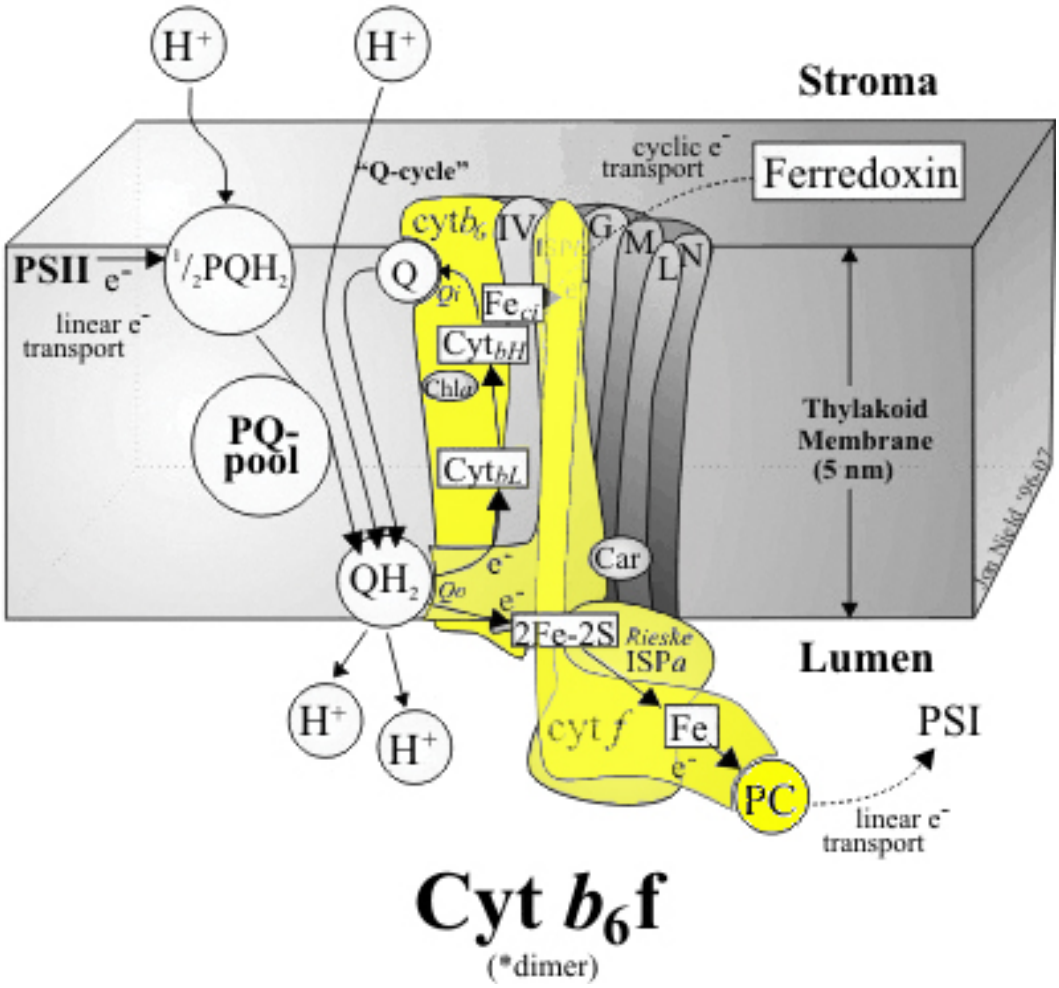


Figure 42. Cytochrome b6f complex proteins over-expressed in CHLD infiltrated tissue

Over-expressed proteins of the Cytochrome b6f complex are shown highlighted in yellow. These include plastocyanin, cytochrome f, cytochrome b_6, and plastocyanin.

Modified from www.jonnield.com/en/psiiimages/oxygenicphotosynthmodel.html

In photosystem I, over-expressed proteins included PSAD, PSAH, PSAE, PSAF, LHCA3 and PSAC (fig. 43).

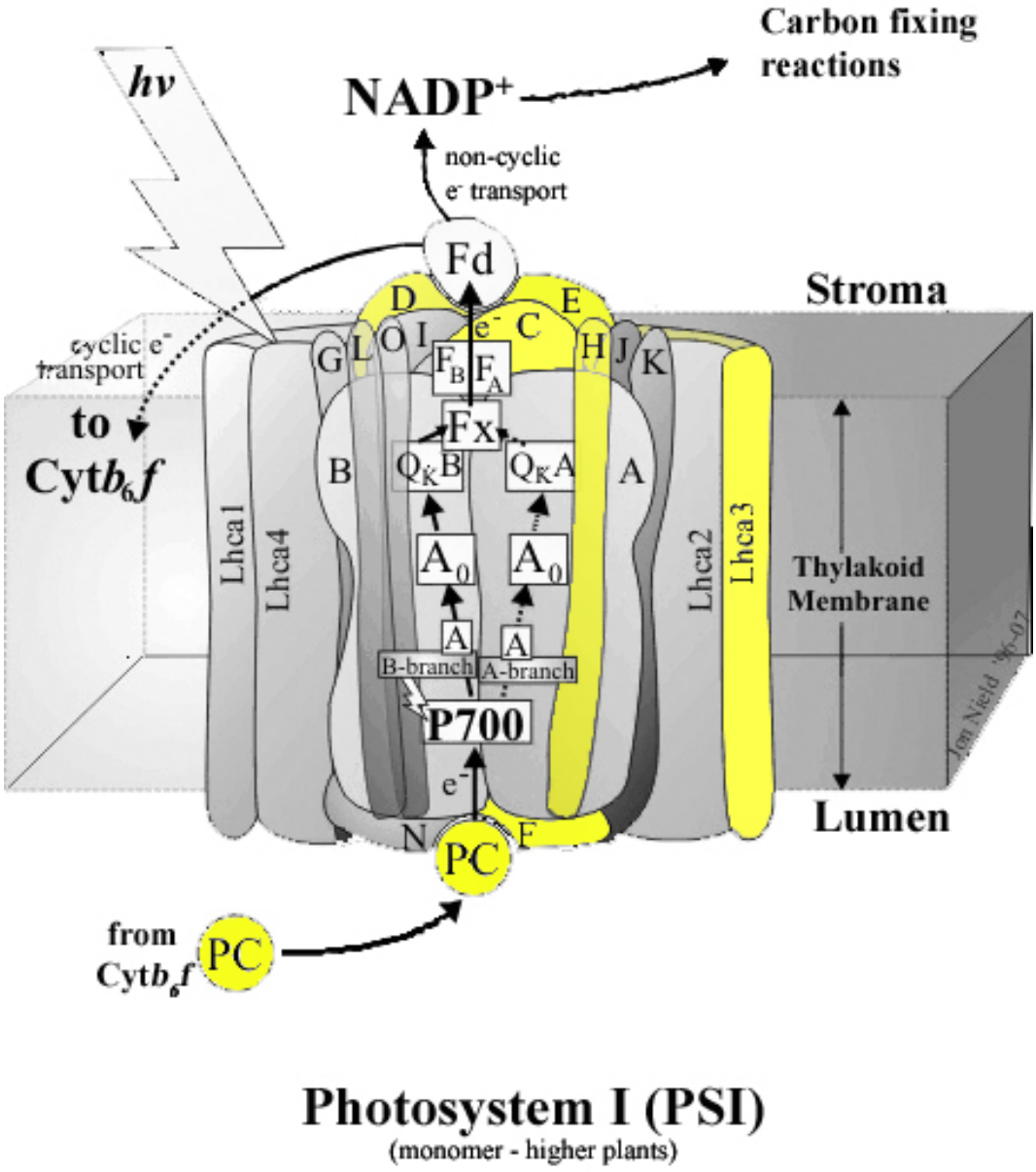


Figure 43. Photosystem I proteins over-expressed in CHLD infiltrated tissue

Cytochrome b6f proteins over-expressed in CHLD infiltrated tissue include PSAD, PSAH, PSAE, PSAF, LHCA3, and PSAC, highlighted in yellow.

Modified from www.jonnield.com/en/psiiimages/oxygenicphotosynthmodel.html

Finally, over-expressed proteins of the ATP synthase complex include the ATPase α□□□□□□□□□, the ATPase□β-subunit, the CF_1 epsilon subunit, the F(0) subunit b, and the ATPase delta chain (oligomycin sensitivity conferral protein;

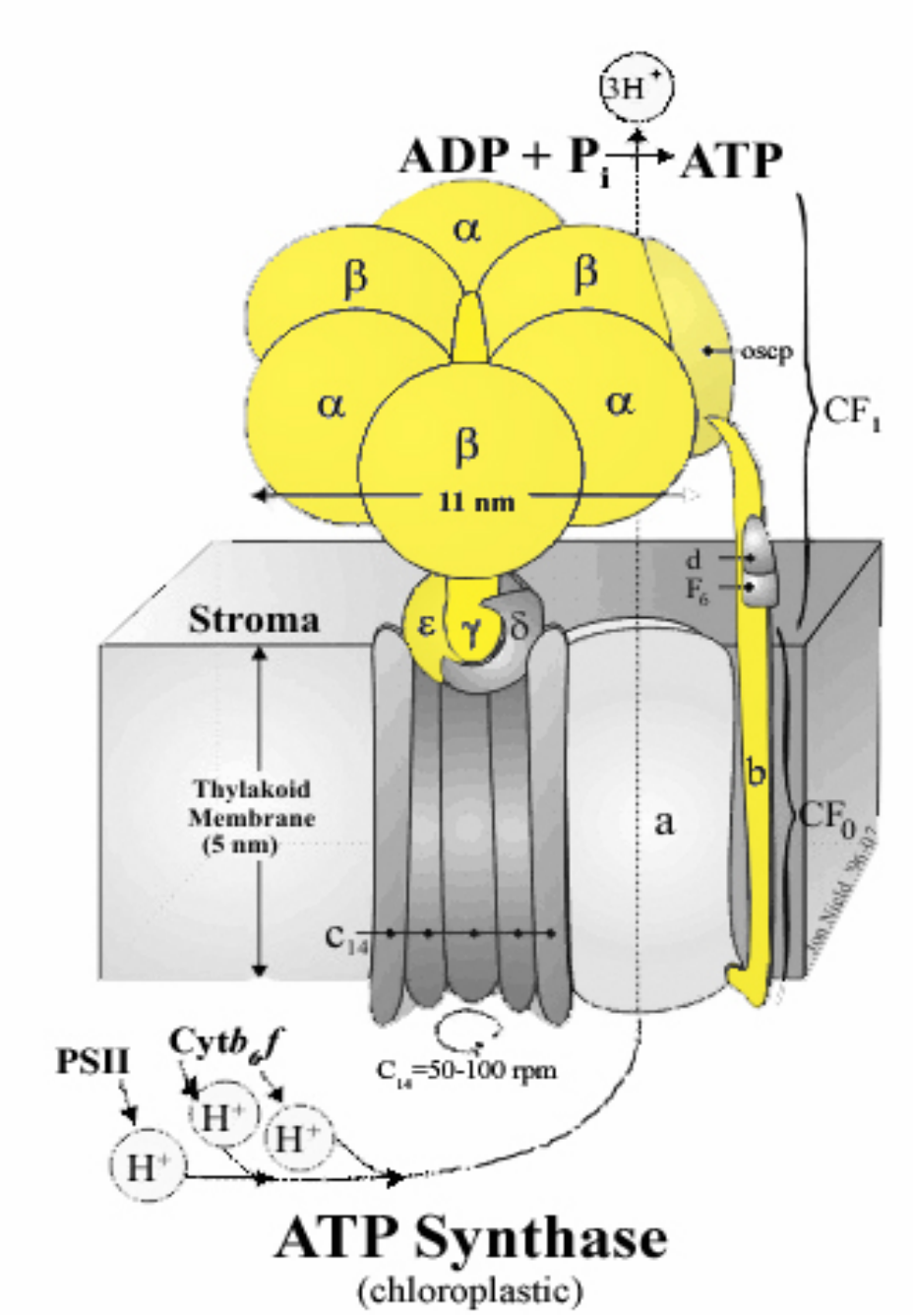


Figure 44. Over-expressed proteins of ATP Synthase in CHLD infiltrated tissue

The PSI proteins over-expressed in CHLD infiltrated tissue are shown here. Included are the ATPase α□□□□□□□□□ the ATPase□β-subunit, the CF_1 epsilon subunit, the F(0) subunit b, and the ATPase delta chain.

Modified from www.jonnield.com/en/psiiimages/oxygenicphotosynthmodel.html

oscp) (fig. 44)

As can be seen from figures 41-44, CHLD over-expression leads to increased production of an impressive portion of the photosynthetic apparatus. Although previous work demonstrated that expression of nuclear encoded, chloroplast localized genes correlates to the production of chlorophyll {Oelmuller, 1986 #133}, these data imply that expression of the photosynthetic machinery detected here correlates to the expression of the CHLD subunit of Mg-chelatase.

Apart from the differentially over-expressed photosynthetic proteins, one protein of particular interest in CHLD over-expressing tissue is the under-expressed protein BTF3. BTF3 was originally isolated as a general transcription factor required for RNA polymerase II dependent transcription, but later found to be a subunit involved in nascent-polypeptide associated complex (Yang et al. 2007). This complex binds to ribosome-associated nascent polypeptide chains as they emerge (Wiedmann et al. 1994), and controls targeting of nascent proteins to their correct subcellular locations by preventing inappropriate interaction of the nascent peptide with signal recognition particle (SRP), thereby inhibiting incorrect targeting (Lauring et al. 1995). BTF3 plays a role in translational control, and functions in directing nascent polypeptides to different translocation systems, including the chloroplast translocation complex(es) (Freire 2005). In tissue over-expressing the CHLD protein, which is chloroplast localized in addition to exhibiting over-expression of many photosynthetic proteins, this protein might be expected to be over-expressed, rather than under-expressed.

Over-expression of CHLH led to over-expression of chloroplast ribosomal proteins while select cytosolic ribosomal proteins appear to be under-expressed.

In addition, a number of photosynthesis proteins are over-expressed, including PSI proteins PSAH, PSAD, plastocyanin, LHCA3, and PSAC, as well as PSII proteins PSBQ and PSBH. Inconsistently, RBCS was significantly under-expressed, along with LHCII and PSI p700 apoprotein a1. The contradictory expression levels of these proteins may indicate a regulatory dysfunction with regard to photosynthesis. The observed discrepancy of photosynthetic protein expression is unexpected in this tissue, which displays a wild type phenotype (fig. 21C) and has Mg-chelatase activity of approximately 78% (fig. 33).

In CHLH infiltrated tissue, however, one notable differentially over-expressed protein is isopentenyl diphosphate isomerase (IPI1; GI 13603408). This enzyme is responsible for the isomerization of isopentenyl diphosphate to its more reactive isomer, dimethylallyl pyrophosphate (DMAPP) (Sun et al. 1998). It has been reported that the activity of this enzyme in *E. coli* is limiting for isoprenoid production as indicated by the accumulation of carotenoids in engineered strains producing these pigments. Introduction of isopentenyl diphosphate isomerase enhances the accumulation of carotenoid pigments by several fold (Kajiwara et al. 1997; Sun et al. 1998; Sun et al. 1996). IPI1 levels are increased in high light, in an ABA independent manner (Nakamura et al. 2001). Over-expression of this protein in CHLH infiltrated tissues could be responsible for the increase in carotenoids detected by HPLC (fig. 29).

Nine of the differentially regulated proteins were common to all three experimental tissues (Table 4). This could be due to a global effect seen with over-expression of a Mg-chelatase subunit, or, although the control tissue was infiltrated with Empty vector to reduce proteomic differences resulting from

Agrobacterium infiltration, it could be due to a change in protein expression of the control tissue.

Table 4. Commonly differentially regulated proteins

		Differential expression ratios		
GI number	*Biological process*	*CHLI:WT*	*CHLD:WT*	*CHLH:WT*
20020	*translation*	*1.4382*	*1.4941*	*1.3363*
548746	*translation*	*1.2409*	*1.4875*	*1.2434*
585876	*translation*	*0.7573*	*0.5847*	*0.7663*
730558	*translation*	*0.7797*	*0.5210*	*0.6888*
2497757	*response to biotic stimulus*	*1.6053*	*1.6613*	*1.5655*
34481799	*ubiquitin-dependent protein catabolic process*	*0.8170*	*0.6083*	*0,7833*
117307343	*nucleosome assembly*	*0.6963*	*0.6883*	*0.6709*
158053020	*nucleosome assembly*	*0.7888*	*0.7589*	*0.7562*
183393002	*electron transport chain*	*1.2815*	*1.5577*	*1.2961*

Due to the small number of proteins differentially regulated in CHLI and CHLH expressing tissues, it is difficult to draw conclusions with respect to those subunits. The most dramatic difference seen here, over-expression of a large portion of the photosynthetic apparatus, is contrary to what might be expected in this tissue. Because over-expression of Mg-chelatase D leads to a dramatic decrease in Mg-chelatase activity (fig. 33), and the tissue phenotype is severely chlorotic (fig. 21B) presumably chlorophyll biosynthesis is correspondingly reduced. It has been demonstrated that chloroplast biogenesis, including the establishment of the photosynthetic machinery, is impaired in the absence of chlorophyll. Our results imply, however, that expression of photosynthetic components correlate positively with expression of the CHLD subunit of Mg-chelatase.

Another interesting observation is that while the nuclear encoded small subunit of Rubisco (RBCS) is over-expressed in CHLD over-expressing tissue, the chloroplast encoded large subunit is under-expressed. This is an example of a

clear breakdown in retrograde signaling, as expression of these subunits is normally coordinated. Moreover, RBCS is transcriptionally regulated in response to light, and has been used previously as a reporter of retrograde signaling breakdown. In contrast, other nuclear genes that encode for proteins incorporated into the photosynthetic apparatus (PSII proteins LHCB1, LHCB4, PSBQ, PSBR, PSBS; PSI proteins PSAD, PSAE, PSAF, PSAH, LHCA3) are over-expressed in CHLD over-expressing tissue, despite its chlorotic phenotype. These data provide support for the notion that there is some signaling activity in CHLD expressing tissue.

In conclusion, iTRAQ proteomics of single infiltrations reveal that the chlorosis seen with CHLD over-expression may be more complicated than a breakdown in Mg-chelatase activity. Higher photosynthetic protein synthesis with lower chlorophyll and disrupted Mg-chelatase activity are indications of genome uncoupling, since the there is up-regulation of nuclear encoded photosynthetic genes, regardless of an inability of the chloroplast to produce chlorophyll.

Chapter 6. Microscopy of Infiltrated tissues

There have been several contrasting reports on the localization of Mg-chelatase subunits (Gibson et al. 1996a; Gibson et al. 1996b; Nakamura et al. 2001; Nakayama et al. 1995; Walker and Weinstein 1995). In an attempt to visualize the compartmentalization of these subunits within the plastid, each was sub-cloned into a pBin61 expression vector with a GFP tag. The resultant GFP fusion protein was then visualized using confocal microscopy. This work was intended both to monitor import of the expressed *Z. mays* proteins into the chloroplasts of *Nicotiana benthamiana* leaf tissue, as well as a prelude to scanning electron microscopy, which allows visualization of protein sub-compartmentalization within the plastid. The surprising confocal microscopy results led to a more exhaustive investigation using this technique.

To aid visualization of protein import into the plastid, each subunit was sub-cloned into a pBin61 binary vector encoding a modified green fluorescent protein (mGFP) tag, henceforth, simply referred to as GFP. GFP is a protein found in jellyfish species *Aequorea* (*forskalea, victoria*). Capable of generating a highly visible fluorophore, this protein has been used widely in science as a marker of gene expression and protein targeting in intact cells and organisms (Tsien 1998).

With an emission spectrum peak at 508 nm, and excitation at 395 and 470 nm (Heim et al. 1994), GFP and its derivatives are a powerful tool for a plethora of study types, including transgenic complementation (Bendahmane et al. 2000; Johansen and Carrington 2001; Shao et al. 2003; Van der Hoorn et al. 2000), promoter analysis (Yang et al. 2007), (Yang et al. 2007) and protein production

(Vaquero et al. 1999; Vaquero et al. 2002). Furthermore GFP fusion is a commonly used technique to show protein localization to the chloroplast (Bhushan et al. 2003; Figueiredo et al. 2011). The GFP deriviative used in this study, modified GFP (mGFP), has slightly shifted excitation and emission spectra, with excitation maxima at 405 and 477 nm, and an emission maximum at 508 nm

6.1 Materials and Methods

6.1.1. mGFP Tagging of Magnesium Chelatase Subunits

Sequence verified Mg-chelatase constructs (CHLI, CHLD, and CHLH) in pGEM-T Easy vector were double digested with *Xba*1/*Bam*H1 restriction enzymes and run out on agarose gels (0.8-1.0% w/v) containing ethidium bromide (0.5 µg/ml) or Gel Red (Biotium, Hayward, CA, USA) according to the manufacturer's recommendation. Digested DNA fragments were excised from the gel and gel purified using the Qiagen gel extraction kit prior to ligation to pBin61 binary vector encoding an mGFP tag (kindly provided by Thomas Brutnell and Peter Moffett) and digested with the same enzymes. Transformants were selected by growth on agar plates with tetracycline and kanamycin (15µg/ml and 50 µg/ml, respectively) as selection reagents (1% Tryptone, 0.5% Yeast Extract, 1.0% Sodium Chloride, 1.5% agar).

6.1.2. Confocal microscopy sample preparation

6.1.2.1 Harvesting and fixation of infiltrated leaf tissue

Tissue growth and infiltration were performed as described in sections 3.1.1 and 3.1.2, respectively. Leaves were infiltrated with Mg-chelatase subunits fused to GFP, alone as well as in conjunction with non-GFP tagged subunits. Initial tests showed detectable, but low, GFP expression approximately 72 hours after infiltration, with stronger GFP expression four days after infiltration. Thereafter,

leaves were harvested four days after infiltration, after testing positive for GFP expression by excitation with a hand held 365 nm UV light source.

To facilitate tissue stability prior to confocal microscopy, samples were fixed by vacuum infiltration. Leaf sections 1.8 mm by 5.0 mm in size were vacuum infiltrated in 4% paraformaldehyde solution (4% paraformaldehyde in 1x phosphate buffered saline; 4% paraformaldehyde, 0.137M NaCl, 0.0027M KCl, 0.010M $Na_2HPO_4 \cdot 2H_2O$, 0.002M KH_2PO_4, pH 7.4). Vacuum infiltration was accomplished by placing multiple leaf sections in a 5.0 ml syringe with approximately 1 ml 4% paraformaldehyde solution. Keeping the leaf sections immersed, and covering the syringe tip with a gloved finger, the syringe plunger was drawn out of the syringe, creating a vacuum. The vacuum was held briefly (approximately 1 second) before release. This process was repeated 8 – 10 times.

Infiltrated samples were then incubated in 4% paraformaldehyde solution two to four hours at room temperature, or at 4°C overnight, prior to three successive 30 minute washes in 1x Phosphate Buffered Saline (PBS; 0.137M NaCl, 0.0027M KCl, 0.010M $Na_2HPO_4 \cdot 2H_2O$, 0.002M KH_2PO_4, pH 7.4). After paraformaldehyde fixation, leaf tissue can be stored for a maximum of 10 days at 4°C in 1xPBS, however, for these experiments, agarose embedding was performed within 24 hours, and sectioning was performed immediately after embedding or on the following day.

6.1.2.2 Embedding leaf sections in agarose gel

Fixed leaf tissue was embedded in 6% aqueous agarose solution in a small metal mold. Multiple leaf sections were embedded per mold and then cut into separate

blocks for sectioning on the vibratome. Blocks were then trimmed into pyramid shapes such that the long dimension of the 1.8 mm by 5 mm leaf section was oriented vertically within the block.

6.1.2.3 Vibratome sectioning of leaf samples

Vibrating microtome, or vibratome, sectioning allows fresh leaf tissue to be cut precisely into very thin cross sections, facilitating visualization of GFP-fused proteins within the internal cellular structures using confocal microscopy. Using a Leica VT 1000S vibratome, 70 µm-thick sections were cut from infiltrated leaf tissue expressing a single GFP-tagged subunit or a co-infiltration of a GFP-tagged subunit co-expressed with a subunit with an alternate tag, such as HA or Myc.

The pyramid shaped blocks containing fixed samples were mounted onto a round metal plate using superglue. This mounting plate was then attached to the stage of the vibratome, and the stage area filled with ultra-pure water to keep the agarose/sample hydrated. The vibratome was then programmed for start and stop positions, and the knife speed adjusted to obtain a clean cut. The knife speed varied depending upon the sample, but section thickness was found to be best at 70 – 90 micrometers. Cut samples were manipulated with a small soft water color brush and stored at 4°C in 1xPBS until mounted.

6.1.2.4 Mounting sections for confocal microscopy

Using a small watercolor paintbrush, sections were manipulated into a droplet of water placed on a clean glass slide. Using the paintbrush or the slivered end of a broken wooden swab, the section was positioned so as to eliminate any folding.

The water droplet was then removed using small triangular sections of Whatman filter paper, and approximately 10 µl Gel Mount (Sigma-Aldrich) was added to the positioned sample(s) prior to placement of the cover slip. After sectioning, samples were examined under the confocal microscope as soon as possible (preferably immediately, and not to exceed 24 hours), to preserve chlorophyll fluorescence.

6.1.3. Confocal microscopy

Using an Olympus Fluoview 300 confocal microscope (Tokyo, Japan), vibratome-cut sections of fixed fresh tissue were imaged for chlorophyll auto-fluorescence and GFP fluorescence. Excitation of GFP and chlorophyll employed a Multi-line Argon laser (excitation wavelengths 457 nm, 488 nm, 514 nm).

With this microscope, GFP fluorescence was detected using a NIBA filter (blue excitation (470-495 nm)/green emission (510-550 nm)). Tissue samples were illuminated using an Argon laser, with excitation wavelength of 488 nm. Chlorophyll auto-fluorescence was detected after illumination using the Helium Neon green laser (543 nm), utilizing the WIG filter (green excitation (530-550 nm)/red emission (575 nm)).

6.1.3.1 Analysis of confocal images

Confocal images were analyzed to determine the localization of the GFP fusion proteins expressed in infiltrated *Agrobacterium* tissues. Specifically, detection of GFP co-localizing with chlorophyll autofluorescence was used to determine that the fusion protein was associated with or imported into the chloroplast, as

opposed to being retained in the cytosol. In cases where the tissue was harvested relatively early (approximately 72 hours) after infiltration, the nascent fusion proteins were seen to be localized cytosolically in rare cases (presumably prior to chloroplastic compartmentalization), while the majority of cells in the same tissue showed co-localization of GFP with chlorophyll (fig. 45). In tissue harvested later (4 days after infiltration) the cytosolic localization of GFP was no longer observed.

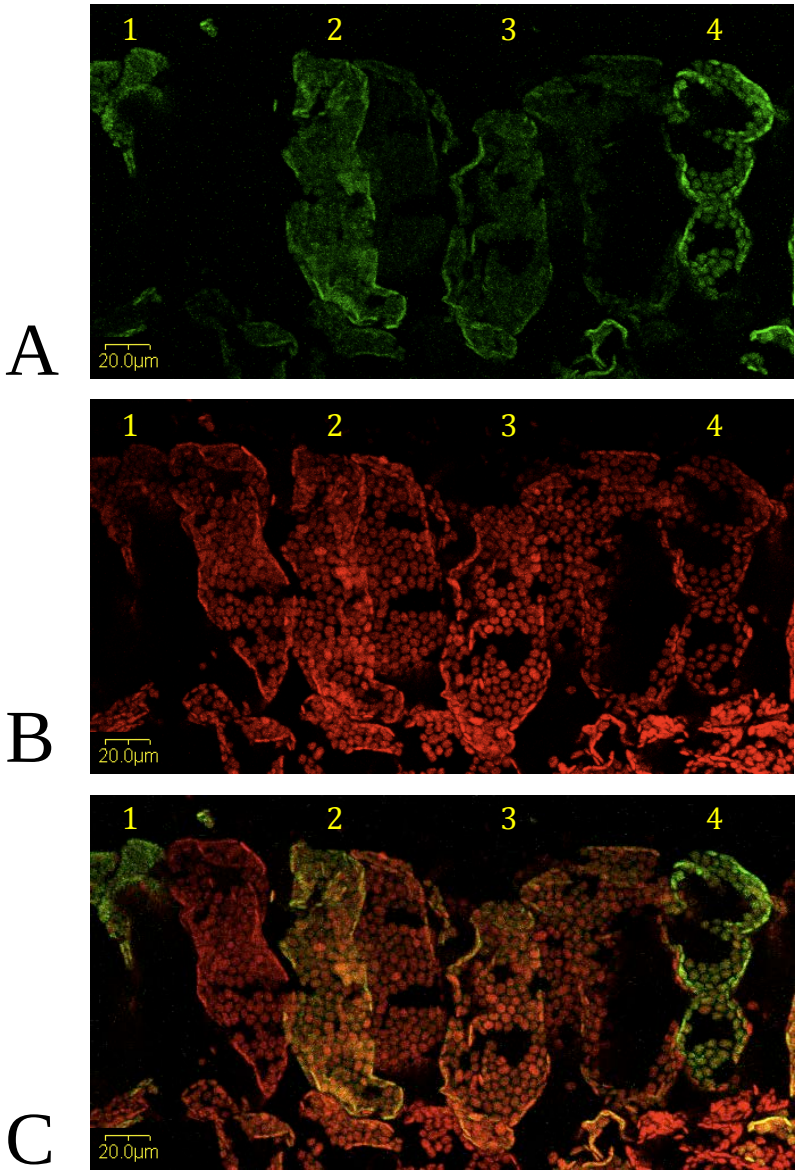


Figure 45. Confocal image analysis

In this figure examples of the GFP detection (A), chlorophyll autofluorescence (B) and overlay (C) are shown. The progression of GFP localization can be seen, with cytosolic localization of the nascent proteins shown in cell 1, the beginning of chloroplastic import shown in cells 2 and 3, and full chloroplastic localization shown in cell 4.

6.2 Results

Experiments expressing Mg-chelatase subunits fused to the GFP reporter protein were replicated with a minimum of three biological replicates. A summary of the data for all of the constructs is provided in table 5. The following sections examine the results for each of the constructs in detail.

Table 5. Percentage of cells expressing GFP fusion constructs in infiltrated tissues

		Percentage GFP expression per biological replicate										*Number of Biological replicates*	*Averaged Percentage GFP expression*
CHLI		*40*	*40*	*45*	*80*	*70*	*80*	*80*				*7*	*62*
CHLD		*10*	*15*	*30*	*17*							*4*	*18*
nuclear CHLH (nCHLH)		*25*	*30*	*60*	*30*	*40*	*30*	*30*	*60*	*37*	*50*	*19*	*41*
		55	*25*	*50*	*28*	*50*	*80*	*25*	*43*	*33*			
chloroplastic CHLH (cCHLH)		*35*	*40*	*30*	*45*	*40*	*50*	*30*				*7*	*39*
CHLD + GUN4		*10*	*10*	*33*	*20*							*4*	*18*
CHLD-GFP + CHLI-HA	*punctate*	*2.5*	*10*	*15*	*7*	*8*	*15*	*22*				*7*	*11*
	diffuse	*2.5*	*5*	*0*	*7*	*4*	*10*	*8*				*7*	*5*
	total	*5*	*15*	*15*	*14*	*12*	*25*	*30*				*7*	*17*
nCHLH + GUN4		*30*	*20*	*20*	*40*	*43*	*25*	*38*				*7*	*31*
cCHLH + GUN4		*30*	*15*	*30*								*3*	*25*

6.2.1. CHLI-GFP Expression

Zea mays Mg-chelatase I subunits fused to mGFP (referred to here simply as GFP) were expressed in *Nicotiana benthamiana* leaf tissue using *Agrobacterium* mediated transient expression. The fusion protein appears to localize in the chloroplasts, as seen in figure 46. The pattern of GFP expression shows that the protein is imported into and dispersed throughout the plastid. Column A shows mGFP detection, with chlorophyll auto-fluorescence depicted in column B. The overlay is shown in Column C. Differential GFP fluorescence intensity between

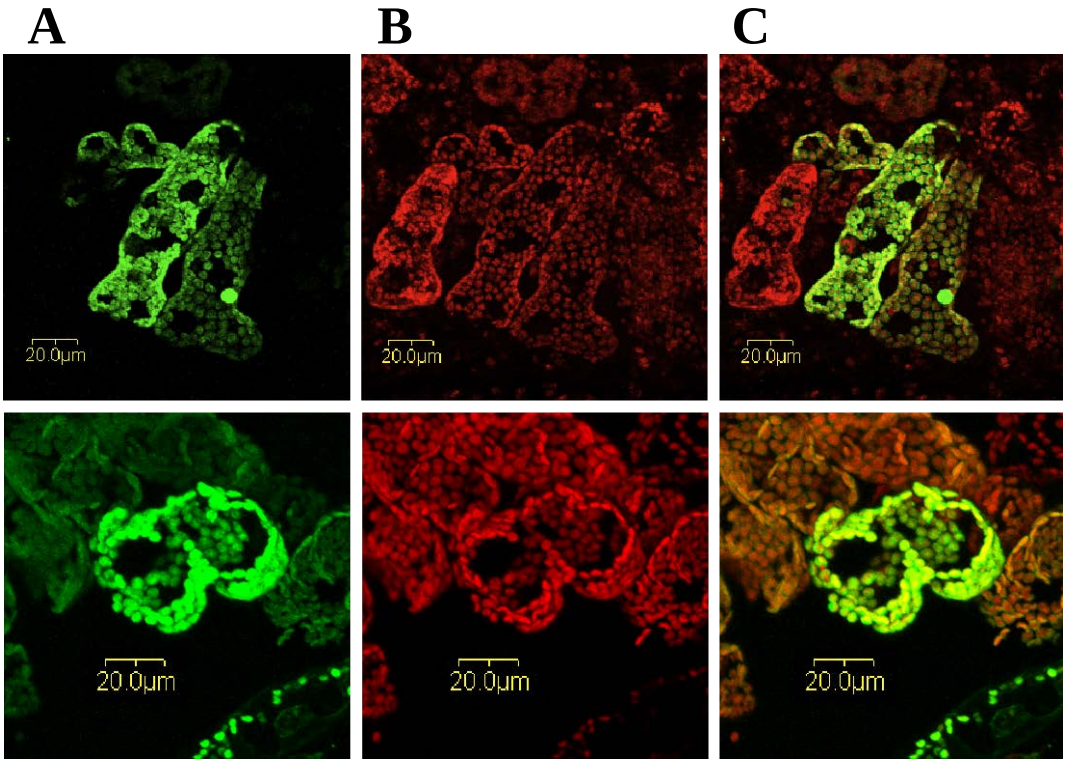


Figure 46. Magnesium chelatase I -GFP fusion protein expression

Shown in this figure are two examples of transiently expressed CHLI-GFP fusion protein. Column A shows chloroplast localized GFP fluorescence, and is dispersed throughout the plastid. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay. As can be seen specifically in the lower region of both examples, some cells were not infected, and thus can only be detected in column B.

individual cells in the same leaf can be observed, along with an infection rate of approximately 50 – 70%. Tissue infiltration led to CHLI-GFP detection in both palisade and spongy mesophyll cells, as well as in the epidermis (stomatal guard cells).

6.2.2. CHLD-GFP

As done with Mg-chelatase I (section 6.2.2), Mg-chelatase D from *Z. mays* was expressed as a GFP fusion protein in *N. benthamiana*. Similarly to the CHLI subunit, CHLD-GFP appears to be localized within (or near), and limited to, the chloroplasts (fig. 47). In this case, however, rather than being dispersed

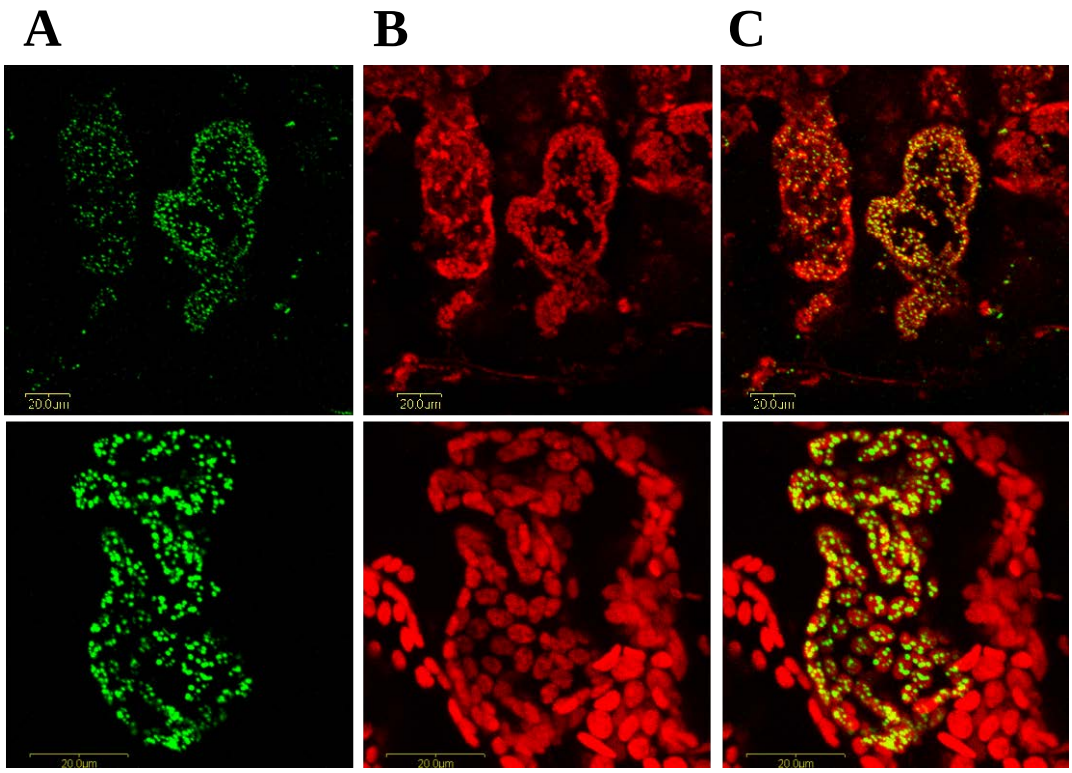


Figure 47. Magnesium chelatase D-GFP fusion protein expression

Two examples of CHLD-GFP expression are shown here. The upper panel shows a wide view of CHLD-GFP expressing tissue, with the lower panel showing expression in a single cell, under greater magnification. CHLD-GFP fusion proteins appear to associate with the chloroplasts. Higher magnification of chlorophyll auto-fluorescence reveals that the plastids are pitted (column B, lower) Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

throughout the plastid, GFP labeling is punctate in appearance (fig. 47A). Close examination of the GFP expression reveals that the GFP localization occurs at or just within the chloroplast surface. Z-stacks of the plastid show that the GFP labeling is limited to the surface of the plastid (data not shown). Under high magnification, visualization of chlorophyll auto-fluorescence shows pits in the chloroplast surface (fig. 47B, 48).

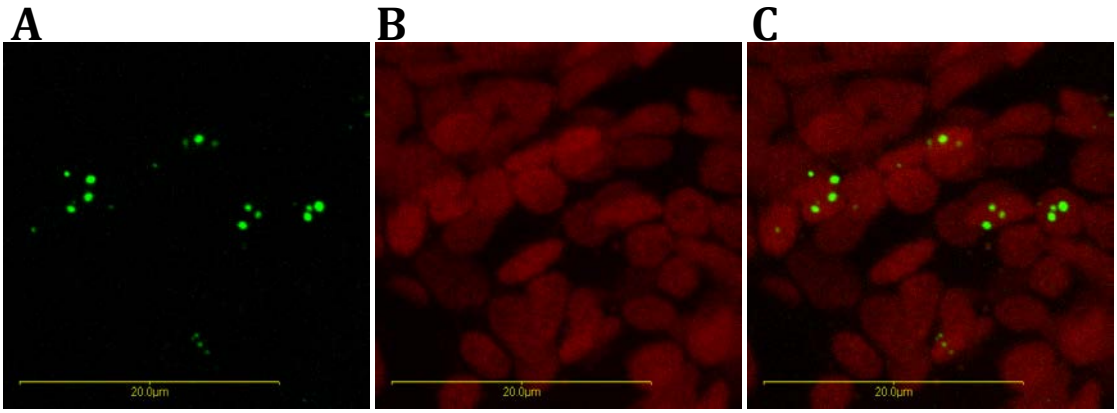


Figure 47. High magnification image of CHLD-GFP

The figure shows the GFP (A), chlorophyll autofluorescence (B) and overlay (C) of plant tissue over-expressing the CHLD-GFP fusion protein.

6.2.3. CHLI-HA co-infiltrated with CHLD-GFP fusion

Co-expression of *Z. mays* mGFP tagged CHLD and HA tagged CHLI (CHLI-HA) reveals two expression phenotypes. Because only CHLD is expressed as a fusion protein, it is not possible to determine the extent to which CHLI is expressed in these confocal experiments. The two phenotypes seen (which pertain only to the CHLD-GFP fusion protein), exhibit both punctate and diffuse forms (fig. 49).

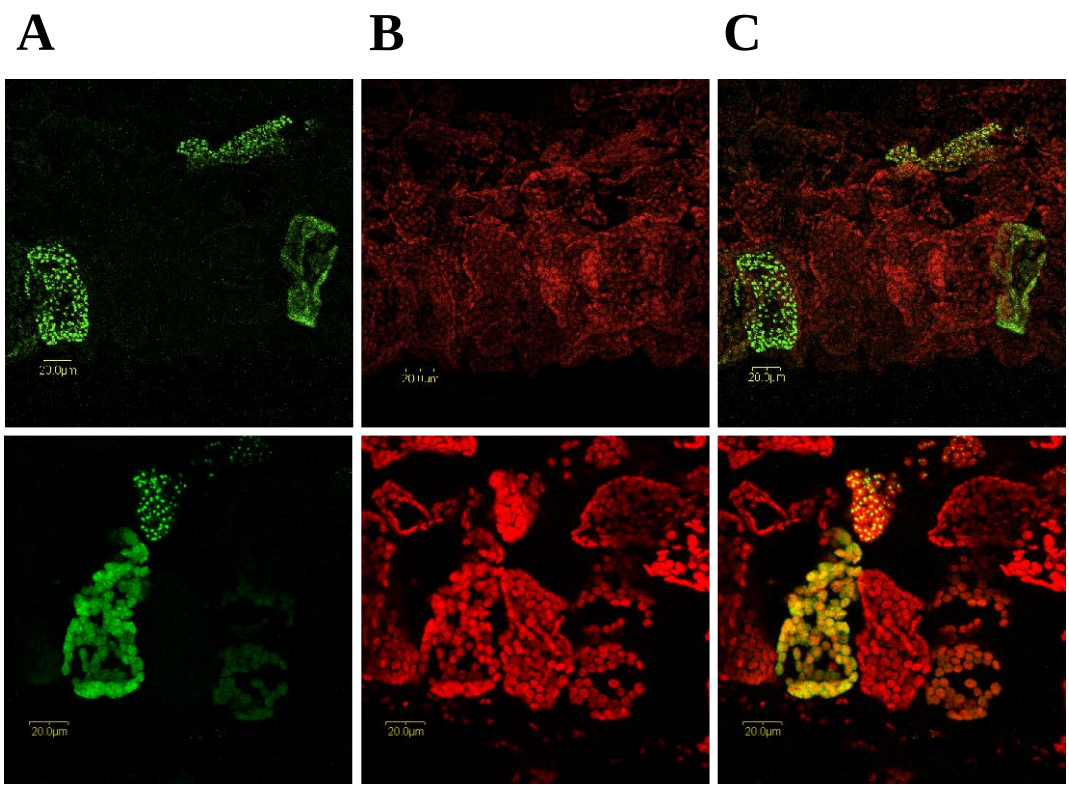


Figure 48. Co-expression of CHLD and CHLI

This figure shows the two expression phenotypes of CHLD-GFP, when the fusion protein is co-expressed with an HA-tagged magnesium chelatase I. Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

6.2.3.1 Co-expression of CHLD and CHLI - punctate GFP expression

In tissue co-infiltrated with GFP-tagged Mg-chelatase D and HA-tagged CHLI, two phenotypic forms are observed. The punctate form observed in co-infiltrated tissue (figs. 49, 50) is identical in appearance to the phenotype seen with expression of the CHLD-GFP fusion protein alone (figs. 47, 48).

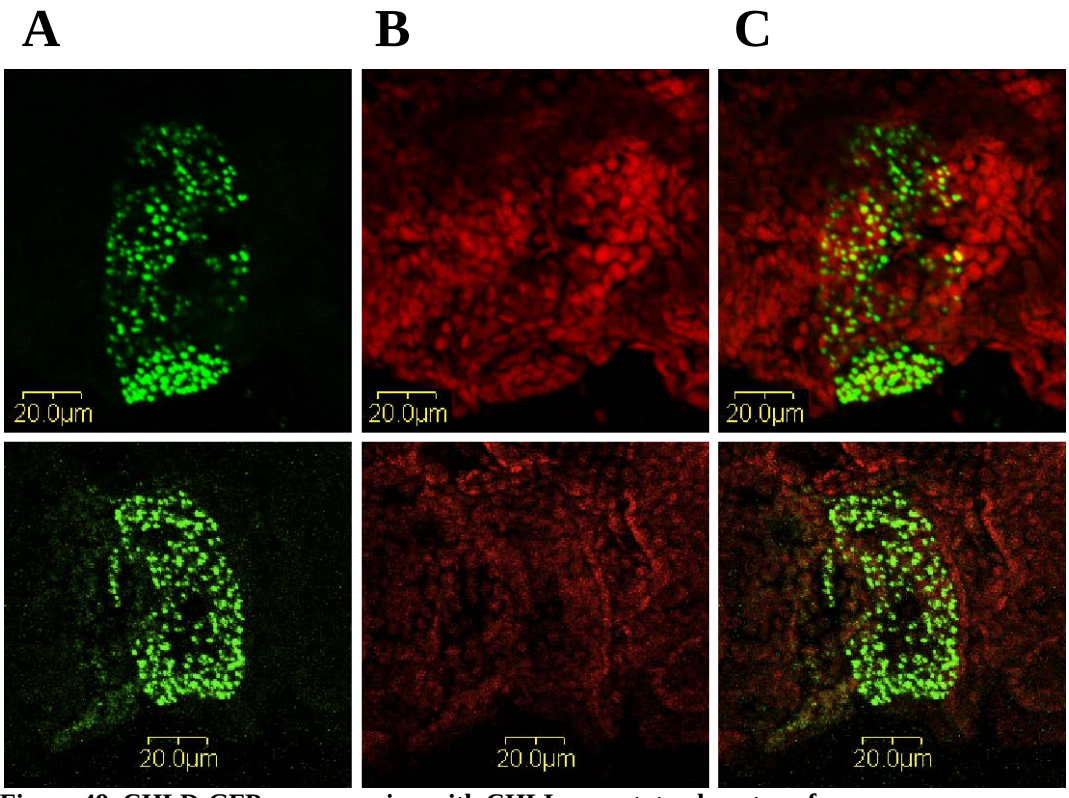


Figure 49. CHLD-GFP co-expression with CHLI - punctate phenotype form

The punctate form of GFP labeling seen with CHLD-GFP/CHLI-HA co-expression is shown. Two examples are given of the punctate labeled cells, which are identical to the expression phenotype of CHLD-GFP expressed alone. Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

6.2.3.2 Co-expression of CHLD and CHLI - diffuse GFP expression

Co-expression of CHLD-GFP and CHLI-HA resulted in an additional phenotype, compared to expression of CHLD-GFP alone. A portion of the cells expressing GFP did not show a punctate form of labeling. Instead, the CHLD-GFP fusion protein appeared to be imported into, and diffusely localized within the chloroplast (figs. 46, 48). Although this phenotype closely resembles that seen with GFP tagged CHLI (fig. 44), in this case the CHLI protein is tagged with the HA epitope, and thus is not directly responsible for this GFP phenotype.

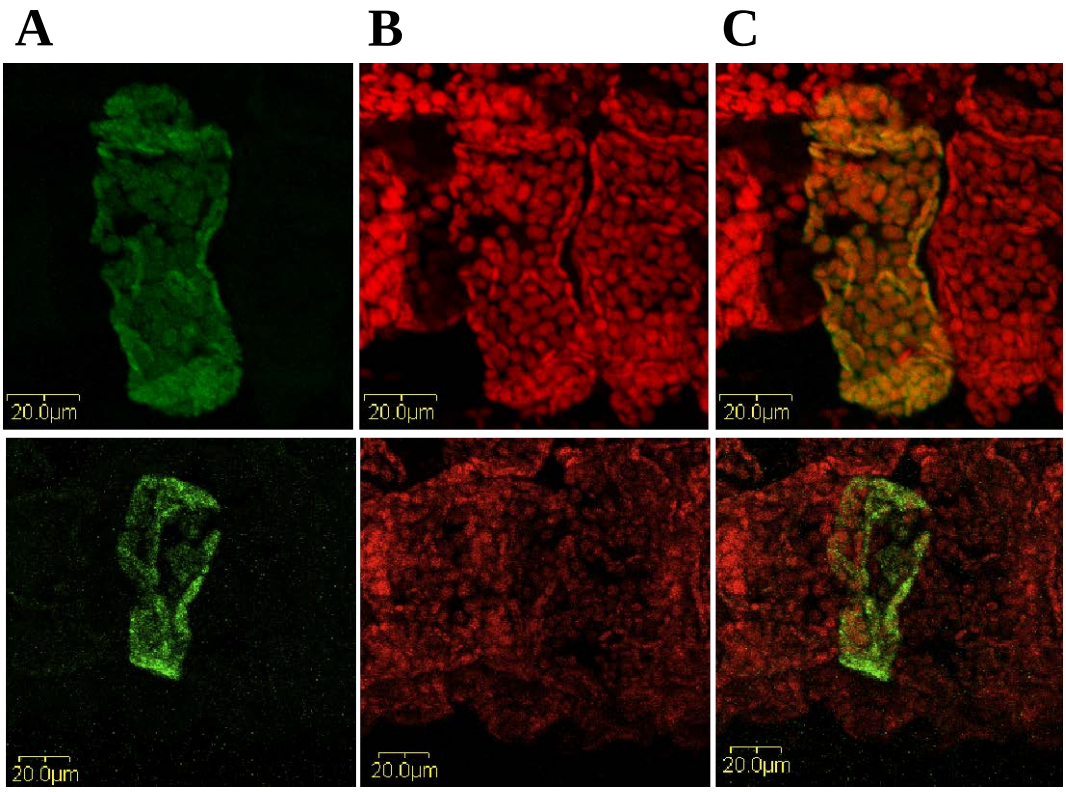


Figure 50. CHLD-GFP co-expression with CHLI - diffuse phenotype form

In leaf tissue co-expressing CHLD-GFP with the magnesium chelatase I subunit, an alternate, diffuse form of GFP labeling. Shown in this figure are two examples of cell exhibiting the diffuse GFP phenotype. Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

6.2.4. CHLH-GFP fusion over-expression

As with the CHLD and CHLI subunits, sub-cloning of Mg-chelatase H was performed as described in section 6.1.1. Ligation to the pBin61-mGFP vector was done and the ligation reaction product used to transform *Agrobacterium.* Multiple individual transformants were cultured and infiltrated into *N. benthamiana* to test for successful expression of CHLH-mGFP. Because the sub-cloning excluded any PCR amplification, all constructs were assumed to be identical.

Infiltrated plants were screened for GFP expression using a hand-held UV light. Tissue expressing GFP was outlined on the leaf with a wide-tipped marker pen and processed for confocal microscopy. Plants that had weak GFP fluorescence were omitted from the processing. Using confocal microscopy, two different expression phenotypes were observed. The two clones used are referred to here as CHLH3 and CHLH4. Due to the fact that these clones were derived from a single sub-cloning procedure, localization to different compartments of the cell was highly unexpected. To investigate this further, the clones were sequenced using 454 pyrosequencing, as described below.

6.2.4.1 CHLH-GFP – chloroplast localized

The first CHLH-GFP fusion protein, referred to as CHLH3, showed GFP localization in the chloroplast. The GFP expression pattern is similar to that seen with Mg-chelatase I, being diffuse in appearance (fig. 52). Detection of CHLH3 was intermittently problematic, with low GFP expression in the pre-processing screen. In addition, some samples exhibiting GFP fluorescence in the pre-processing screen (using a hand-held UV light) did not have detectable GFP fluorescence after processing for confocal microscopy.

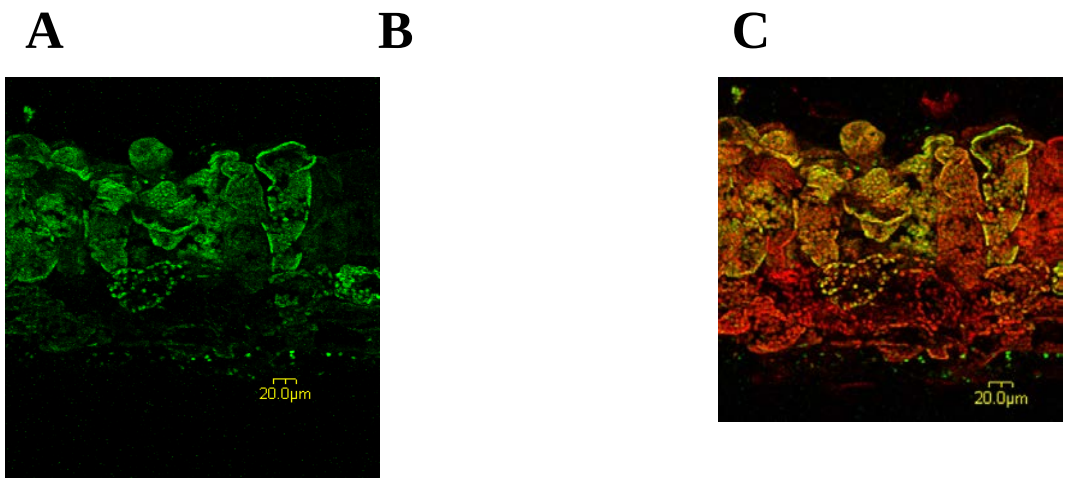


Figure 51. CHLH - chloroplast localized

This figure shows the chloroplast localized form of CHLH-GFP. Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

6.2.4.2 CHLH-GFP – nuclear localized

The phenotype observed with an additionl clone of CHLH, referred to here as CHLH4, showed nuclear localization of the CHLH-GFP fusion (fig. 53). Because these results were unexpected, this experiment was performed exhaustively to

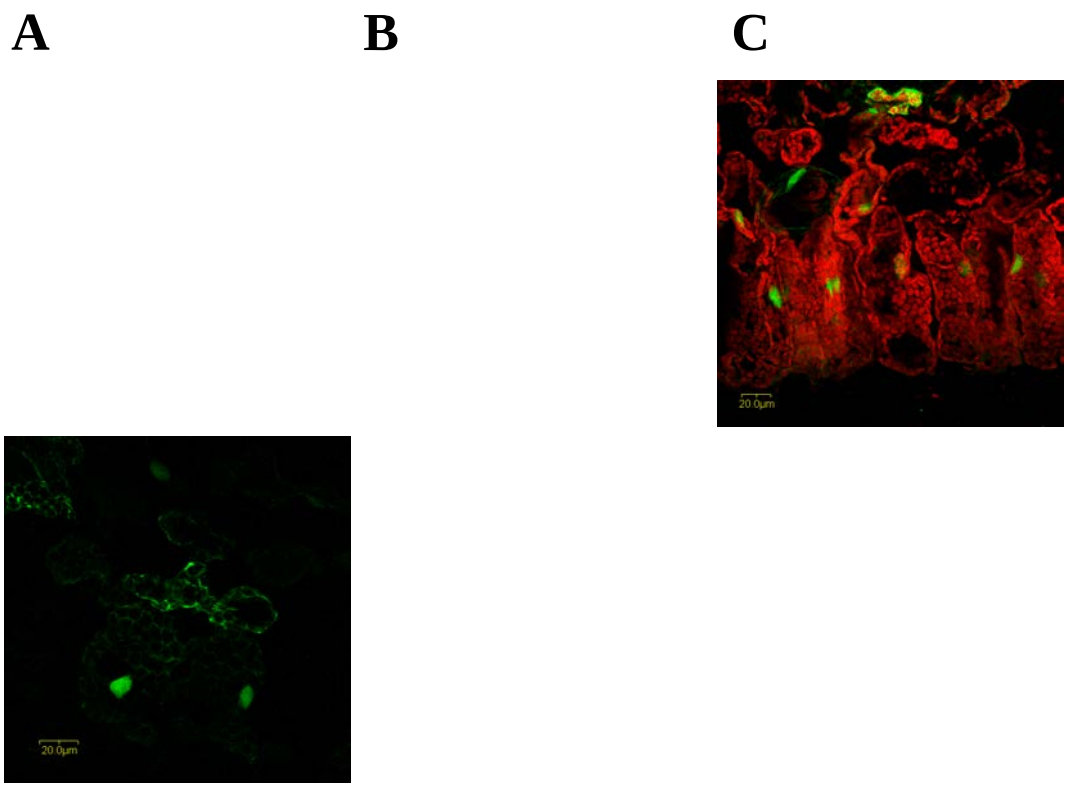


Figure 52. CHLH - nuclear localized

This figure shows the nuclear-localized form of CHLH-GFP. Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

verify the result (table 5). Additional evidence of nuclear localization of the CHLH-GFP fusion was provided using differential interference contrast (DIC) microscopy (fig. 54). In figure 54, chlorophyll autofluorescence can be seen in structures that line the perimeter of the cells, presumably the chloroplasts, while GFP fluorescence can be seen in a central structure within the cell, presumably the nucleus. Another approach to provide evidence of CHLH-GFP nuclear

localization would involve 4',6-diamidino-2-phenylindole (DAPI) staining of the cells.

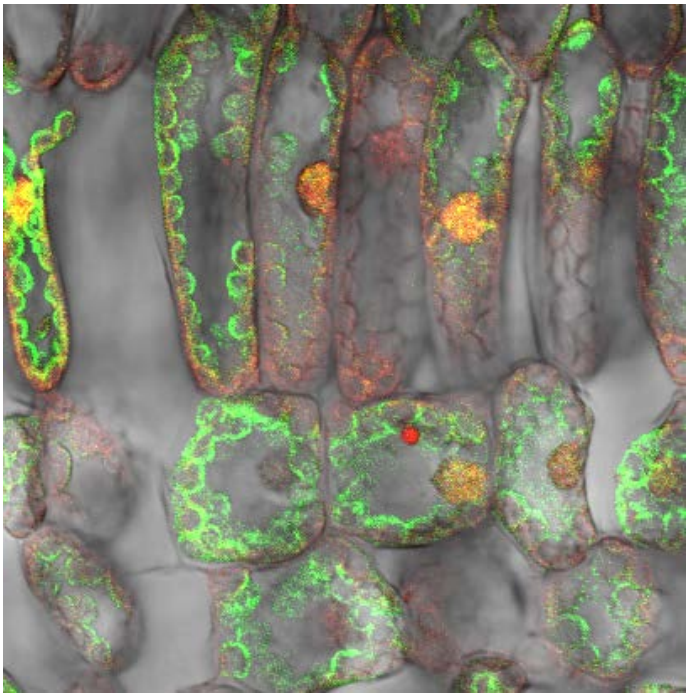


Figure 53. Nuclear localized CHLH - DIC image

A differential interference contrast (DIC) image of cells over-expressing nuclear localized CHLH-GFP fusion protein is shown here, with GFP fluorescence represented in yellow and chlorophyll autofluorescence represented in green.

6.2.4.3 Sequence verification of Mg-chelatase H expression constructs

Because of the surprising nature of the results, 454 sequencing was undertaken using several clones from both the construct resulting in CHLH-GFP chloroplast localization and the construct resulting in CHLH-GFP nuclear localization.

The complete sequences of the pBin61 CHLH-GFP constructs were determined by 454 pyrosequencing. In total, 5 separate clones (2 of the chloroplast localized and 3 of the nuclear localized forms) were labeled with unique sequence tags, multiplexed together, and sequenced on 1/8 of a 454 plate. The reactions were accomplished using a Roche 454 GS FLX sequencer with titanium reagents (Roche; Penzberg, Germany). The sequencing was performed at the Ramaciotti Centre for Gene Function Analysis.

High-level coverage of each plasmid was generated and construct sequences were assembled using the MIRA assembler (Chevreux et al. 2004). Each plasmid assembled into a single contig and the sequences were examined for mutations. Pairwise alignment of each of the plasmid construct sequences using ClustalW revealed that all five plasmids were identical in sequence.

6.2.4.4 Chloroplastic localized CHLH-GFP co-expression with GUN4

Co-expression of the GUN4 protein with the diffuse form of CHLH leads to a much more consistent signal from the GFP fusion protein. GFP fluorescence in leaf tissues expressing the fusion were detectable using both the hand held UV light, as well as the confocal microscope (fig. 55), in contrast to tissue expression the CHLH-diffuse-GFP alone.

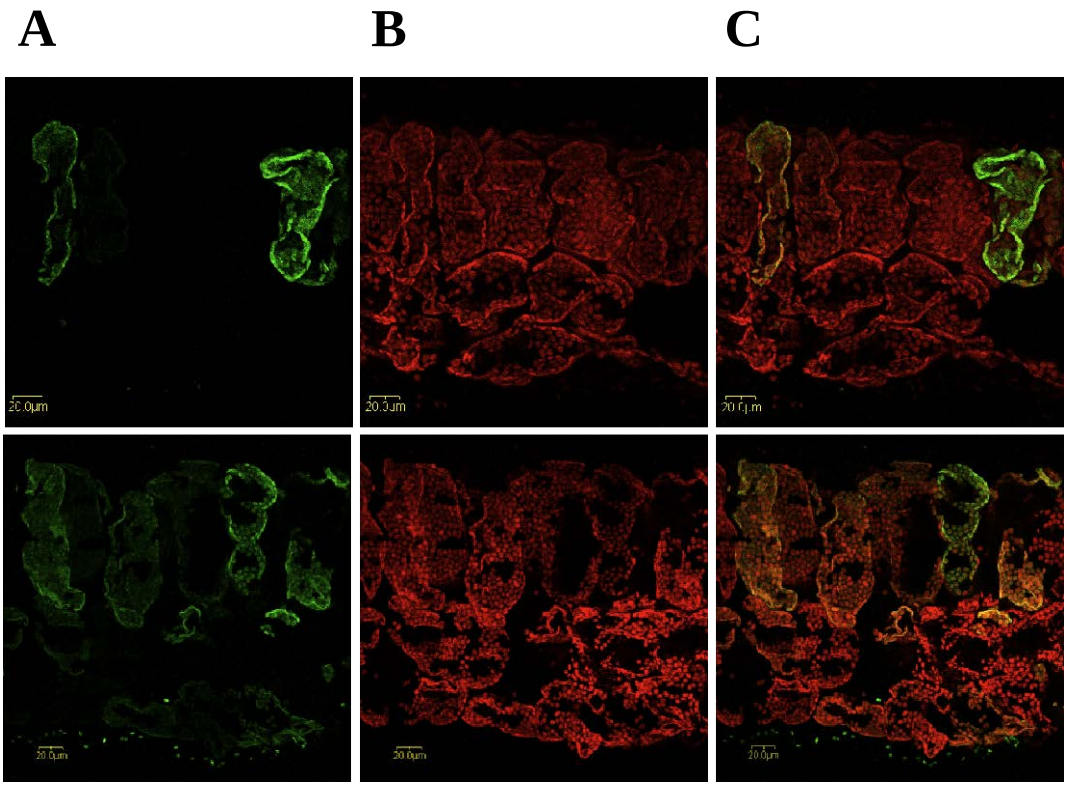


Figure 54. Chloroplast-localized CHLH-GFP co expression with GUN4

Chloroplast localizing form of CHLH-GFP is shown co-expressed with the GUN4 protein. Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

6.2.4.5 Nuclear-localized CHLH-GFP co-expression with GUN4

Localization of the CHLH-nuclear-GFP fusion protein did not change upon co-expression with the GUN4 protein (fig. 56).

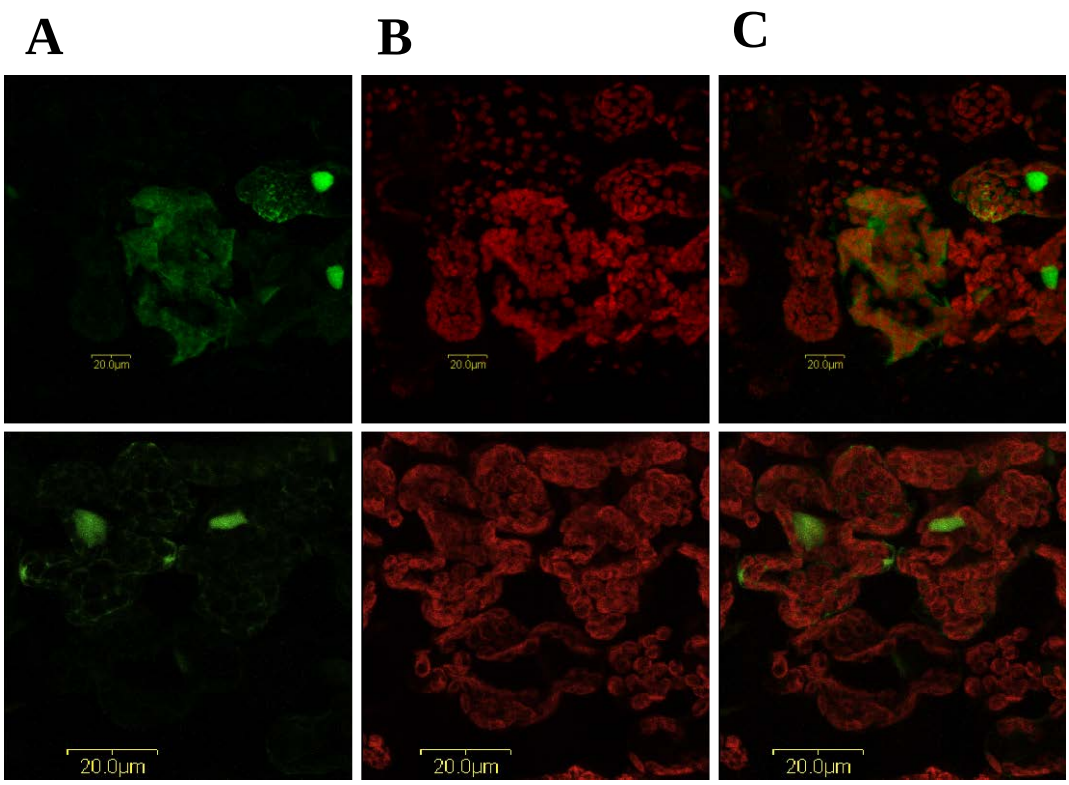


Figure 55. Nuclear-localized CHLH-GFP co-expression with Gun4

Co-expression of GUN4 with the CHLH-GFP fusion does not affect compartmentalization of the protein, as seen here. Column A shows GFP fluorescence. Column B shows chlorophyll auto-fluorescence, and column C shows the overlay.

6.3 Discussion

Expression of Mg-chelatase I fused to GFP shows that this subunit is localized to, and appears to be imported into the chloroplast without issue. In contrast, while the nascent CHLD-GFP fusion protein appears directed to the chloroplast, rather than aggregating in the cytosol or remaining diffused throughout the cytosol, in tissues expressing this fusion alone, CHLD-GFP appears to aggregate at or near the surface of the plastid (figs. 47, 48).

Phenotypic rescue of the CHLD chlorosis by co-expression with the CHLI subunit prompted examination of the same type of co-infiltration using the GFP reporter. When co-infiltrated with both the CHLD-GFP fusion and an HA epitope tagged Mg-chelatase I-subunit, tissues exhibited two GFP expression phenotypes (fig. 49). In these tissues, both a diffuse and a punctate form of GFP labeling were observed, with the punctate form being more common, in an approximate ratio of 2 punctate:1 diffuse.

We speculate that the punctate form of GFP labeling is attributable to infection only by *Agrobacteria(um)* carrying the CHLD-GFP construct. Cells exhibiting a diffuse GFP labeling form, however, were likely infected by at least two separate *Agrobacteria*, carrying the two different constructs, CHLD-GFP and CHLI-HA. The subsequent transfer of both constructs then results in co-expression of CHLD-GFP, and CHLI-HA. The diffuse form of GFP expression could be a result of this co-expression. If CHLD requires CHLI to be properly incorporated within the plastid, transient expression could lead to a surfeit of the CHLD protein, and thus, issues with import or sub-localization. This hypothesis is supported by the fact

that the diffuse expression pattern of CHLD-GFP is not observed in the absence of CHLI co-expression. Note that the CHLI protein was labeled with HA, rather than GFP, and is thus not directly responsible for the diffuse GFP phenotype.

These data imply that the punctate form of GFP expression seen with CHLD expressing leaf tissue may be due to a problem with import of this protein into the plastid. A study by Friedman and Keegstra (1989) demonstrated saturation of the TOC machinery. This finding provides support for the existence of specific protein import sites at the chloroplast surface (Jarvis 2008), and may explain the punctate appearance of CHLD-GFP seen here. Furthermore, rescue by CHLI resulting in the diffuse phenotype seen in CHLD/CHLI expressing tissues could be resulting from co-import of the two subunits, as is seen with PORA and its substrate (Jarvis and Robinson 2004). Although in this case the co-import would be of two subunits, rather than an enzyme and its substrate, it is conceivable that a similar mechanism exists to control stoichiometry ratios of Mg-chelatase.

Alternatively, it could be that the CHLD is forming aggregates, due to a high concentration of the protein. In that case, one might expect to see aggregates forming in the cytosol, unless a threshold level of protein concentration is needed for aggregation formation. Collection of the CHLD proteins within the chloroplast envelope, immediately after import, may accomplish the concentration of proteins needed for the formation of aggregates, and in this case, due to the diffuse phenotype seen with co-expression of the CHLI protein, the data implies that CHLI is involved in the correct sub-localization of CHLD-GFP.

Expression of the CHLH-GFP fusion produced unexpected results. Generation of the two constructs used should have resulted in identical plasmid sequences (see

materials and methods section), but after screening and processing, one clone localized to the plastid, as expected, while the other localized to the nucleus. Because localization of the CHLH protein in the nucleus has never been reported, these results were initially met with uncertainty.

To verify that the correct sequence was present in both clones, multiple plasmid preps were performed for each and sequenced using 454 sequencing. Subsequent alignment of the returned sequences (3 clones for the nuclear localizing form and 2 clones for the chloroplast localizing form) showed these clones to have 100% identity. There are no indications in the sequence, therefore, as to why these clones exhibit the two, very distinct phenotypes.

An interesting aspect of GFP stability observed with the chloroplast localized form, however, was that the signal intensity varied noticeably from one experimental sample to the next. This is in contrast to expression of the CHLI-GFP, CHLD-GFP, and CHLH-nuclear-GFP fusions. In these other cases, expression seems to be more stable and reliably detectable after sample processing. With the chloroplast localized form of CHLH-GFP, fluorescence was detected with a hand held UV lamp prior to processing, but undetectable after processing, using the confocal microscope. Furthermore, co-expression of the CHLH-chloroplast-GFP with the GUN4 protein resulted in what seemed to be a stabilization of the H-subunit, as detection of the GFP signal was much more reliable in these tissues. Based on these data, we conclude that the GUN4 protein may serve to stabilize the CHLH subunit, which may otherwise be degraded prior to chloroplastic import. This conclusion is supported by recent findings by Peter and Grimm (Peter and Grimm 2009), who found that GUN4 stabilizes enzymes in

tetrapyrrole biosynthesis. In addition, it seems that GUN4 is not involved in localization of the CHLH subunit in the chloroplast, as co-expression of GUN4 with the nuclear localizing form of CHLH did not change compartmentalization of the H-subunit (fig. 56).

Localization of the CHLH protein within the nucleus is puzzling. This has never been reported, and only occurs with one of the clones. However, because this result was strongly questioned, the experiment was repeated exhaustively, and the binary vector (along with the CHLH insert) sequenced using the powerful 454 technique. The results were consistent throughout repeated testing. We speculate, therefore, that the differences seen in expression pattern may result from differences in the efficacy of the two *Agrobacterium* clones. Because of the inconsistency with signal observed with the chloroplast localized form, along with the strength of the signal observed with the nuclear localized form of CHLH-GFP, we suggest that localization of this subunit may be influenced by its concentration.

Chapter 7. Conclusions

The original aim of the project was to examine the effects of over-expression of Mg-chelatase subunits, expressed either singly, or in combinations of subunits. This work employs a modified *Agrobacterium tumefaciens*, the plant pathogen responsible for crown gall disease. (Pizschke and Hirt 2010; Smith and Townsend 1907). With this system, protein production is accomplished using the binary vector pBin61, which includes a tandem 35S promoter to drive constitutive protein expression (Bendahmane et al. 2002). Because the construct is not incorporated into the plant's genome, but instead is simply expressed upon infiltration into the experimental plant, the direct effects of construct expression can be ascertained without the confusion that might result from the pleiotropic effects of herbicide use (Oelmuller et al. 1986) or from global proteomic changes resulting from plant development.

To this end, each of the subunits of Mg-chelatase were cloned into the pBin61 binary vector to enable their transient expression in *Nicotiana benthamiana* (Bendahmane et al. 2002). Phenotypic analysis of the cloned subunits of Mg-chelatase transiently expressed in *Nicotiana benthamiana* tissues led to a novel and surprising result, namely that over-expression of the wild type CHLD subunit led to a chlorotic phenotype. Investigation of the basis for this chlorosis became an additional focus of inquiry in this work.

Our data have resulted in several new findings with respect to the Mg-chelatase enzyme. First, over-expression of any of the subunits of Mg-chelatase leads to a reduction in enzyme activity, suggesting that efficient functioning of this enzyme

is dependent upon a strict maintenance of the correct stoichiometric ratio. These data are consistent with previous work, where the activity of Mg-chelatase was decreased with both over- and under-expression of the CHLI subunit in tobacco (Papenbrock et al. 2001), although *in Vitro* work has shown cooperativity for Mg-chelatase H (Sawicki and Willows 2008).

Co-infiltration experiments showed that the chlorotic phenotype observed with CHLD over-expression is rescued by co-expression of the I-subunit. Furthermore, we determined that the interaction of the I- and D-subunits may involve the MIDAS motif of the CHLD integrin I domain, using truncated CHLD proteins and subsequent phenotypic analysis. There is considerable evidence that CHLD and CHLI interact, based on both *in Vivo* and *in Vitro* studies (Guo et al. 1998; Jensen et al. 1998; Lake et al. 2004; Papenbrock et al. 1997; Willows et al. 1996; Willows et al. 1999). Furthermore, the predominant theory of integrin I involvement in protein-protein interactions is that all integrins use the MIDAS motif as the central anchorpoint for their ligands (Takagi 2007).

Although it appears that co-expression of the CHLI subunit is involved in rescue of CHLD chlorosis, the reason for this chlorosis was still unexplained. To investigate the phenotype further, Mg-chelatase activity assays were performed in leaf tissue over-expressing the CHLD subunit, and contrasted to the activity of tissue over-expressing a mutant form of CHLD, called DM9. The DM9 construct had a single amino acid substitution of the central serine of the MIDAS motif, and when over-expressed, had an activity level 30% that of wild type tissue, as compared to the 45% activity level seen with the CHLD over-expressing tissue (fig. 33). Co-expression of the CHLD and CHLI subunits, which results in rescue

of CHLD chlorosis, results in Mg-chelatase activity 85% that of wild type. These data indicate that although the CHLD subunit is functional in the host leaf tissue, over-expression of the subunit in the absence of corresponding CHLI has a strongly deleterious effect on enzymatic activity. The question that remains is: can the chlorosis seen in CHLD over-expressing tissue be attributed solely to this deleterious effect, or is it possible that there is some sort of regulatory signal involved that elicits this phenotype?

To investigate the possibility of a signal resulting from CHLD over-expression, we used iTRAQ analysis to examine proteomic changes in this tissue. Interestingly, the analysis detected over-expression of almost half the proteins involved in photosynthesis, including PSBA, B, C, D, H, Q, R, S, LHCB1 and 4, cytochrome f, cytochrome b_6, and plastocyanin, etcetera (figs. 41-44). Retrograde signaling is defined as the process by which the chloroplast exerts control over nuclear encoded, chloroplast localized genes (Nott et al. 2006). In the absence of normal chlorophyll production, for example, the plastids are subject to photo-oxdative damage, resulting in down-regulation of nuclear genes (Oelmuller and Mohr 1986). Although leaf tissue over-expressing CHLD is chlorotic, many of the over-expressed proteins detected by iTRAQ analysis are nuclear encoded, demonstrating a clear breakdown in retrograde signaling. Furthermore, these data suggest that expression of the photosynthetic apparatus proteins detected is correlated with the expression of CHLD, rather than with chlorophyll production.

The final investigation of CHLD chlorosis explored localization of the transiently expressed fusion protein, CHLD-GFP, using confocal microscopy. Microscopic analysis of *N. benthamiana* leaf tissue infiltrated with the CHLD-GFP construct

resulted in detection of GFP expression in approximately 18% of the cells, as compared to 62% and 41% of cells expressing CHLI-GFP and CHLH-GFP, respectively (table 5). Although the expression level seen in CHLD-GFP infiltrated tissue may be adversely affected by the GFP tag, in light of the possible involvement of a signal, it is interesting to note that while only 18% of cells were seen to express GFP, the chlorosis seen in tissue expressing CHLD-T7 (fig. 22) and CHLD-GFP (data not shown) was uniform and quite severe throughout the infiltrated region.

Based on these results, is not possible to say if CHLD chlorosis is due to a decrease in chelatase activity resulting simply in a breakdown in chlorophyll biosynthesis, or if it is the result of some kind of signaling mechanism. Phenotypic analysis experiments, however, suggest that the chlorosis seen with CHLD expression is elicited by the N-terminal half of the protein, while rescue with CHLI co-infiltration may involve interaction of the two subunits via the MIDAS motif.

Finally, in leaf tissue transiently expressing the D-subunit, it appears that import of the CHLD protein is impaired in the absence of correspondingly high levels of CHLI. This finding is demonstrated most readily by the diffuse phenotype of GFP expression seen in leaf tissues co-expressing the CHLD and CHLI subunits (fig. 44). A reasonable conclusion to this work is that the CHLD protein is imported into the chloroplast in a controlled manner, which may serve to maintain the correct enzyme stoichiometry, and that the import of the D-subunit may be facilitated by CHLI.

In addition to the basis behind the CHLD phenotype, another question that remains pertains to the nuclear localization of the H-subunit. Overexpression of the H-subunit led to a wild type leaf phenotype, a slight decrease in Mg-chelatase activity and a comparatively small set of differentially expressed proteins, when contrasted with CHLD expressing leaf tissue. The most curious aspect of CHLH expression, however, was the difference in localization between the two constructs. The first construct localized to the plastid, but was observed to be at times difficult to detect. Co expression with the GUN4 protein seemed to stabilize the subunit, however. The second construct localized to the nucleus, and appeared to be unaffected by GUN4 expression. This difference in localization cannot be attributed to sequence of the construct or the binary vector, as, based on 454 sequence analysis, the sequences for the chloroplast-localizing and nuclear-localizing forms match with 100% identity.

We speculate that the GUN4 protein serves to stabilize the CHLH subunit destined for the plastid. In addition, to explain the localization difference, we hypothesize that there is some disparity in the efficacy of *Agrobacterium*-mediated expression, leading to differences in CHLH concentrations in the infiltrated plants. This difference in concentration may, in turn, be responsible for the inconsistency seen with localization of this protein.

Appendices

Appendix A. Details of Oligonucleotide Primers

Name	Sequence	Description
pf300	GTCGTTGGGCAGGATGCTATCAAAACTG	Zm cDNA Chel D Forward
pf301	TGAGCTCTTCAAGTCTGTCAGGGCGGTCT	Zm cDNA Chel D Reverse
pf302	GCAATGCCTCCGATCTCACGATCAAT	Zm cDNA 5' RACE Mg Chelatase D
pf303	CCACTGCTTCAACGCGGTCATCAAAA	Zm cDNA 5' RACE Mg Chl D Reverse
pf304	CGTCACACCAAGTGGAATCTGCACAAAA	Zm cDNA 5' Race Mg chl D rev nested
pf305	ATGAGAGCCAAAAGAATGGCTCGAA	Zm Mg Chelatase D sequencing primer
pf307	GCTGCTATGGAAGGACGTGAAAAAGT	Zm Mg Chelatase D sequencing primer
pf308	CCTGATCTCTGCTGGTGTAGCTTTCTG	Zm Mg Chelatase D sequencing primer
pf309	CAACCTGGTCTTCTTGCTGAAGCAC	Zm Mg Chelatase D sequencing primer
pf310	TCTTGATTTTCTGGAGGTGGTGGT	Zm Mg Chelatase D sequencing primer
pf311	GAGCCGTCGTTGGGCAGGAT	Zm Mg Chelatase D sequencing primer
pf312	TGGCGAAAGCTAATGCCCTCTC	Zm Mg Chelatase D sequencing primer
pf313	GCAGAGCAGTTTTGATAGCATCCTG	Mg chelD RACE 5' nest rev comp to genrcr5' nest
pf314	AGGTCATGCGCCTCAACAAGCTC	Mg chel H 5' internal primer
pf315	ACAAGCTCGGCTCCTTCAGCATGT	Mg Chel H 5' nest internal
pf316	TCCCAATGAGGCAGAGAGTTCCA	Mg Chel H 3' UTR
pf317	CAAATCAATTTACCTGTCAATCCCCTCAA	Mg Chel H 3' spans stop (TAA)
pf318	ATGCTGAAGGAGCCGAGCTTGTTG	Mg Chel H 5' RACE GSP
pf319	CGAGCTTGTTGAGGCGCATGAC	Mg Chel H 5' RACE nested GSP
pf320	TCTAGATGGCGACGCCCACCGCGCTCT	Mg Chelatase D FLF + Xba1
pf322	GAGCTCCGCGACGAGGACACCTACG	Mg Chel H 1a
pf323	TGTGCAATGCATGTGACTTCCCTGTCC	Mg Chel H 1b
pf324	TGCAGAAGCTCGACGTGCCGTACA	Mg Chel H 2a
pf325	GACAGGCTCCACACGGTTGACACG	Mg Chel H 2b
pf326	TACCTGACCCCACCGGCCGAGAAC	Mg Chel H 2d
pf327	ATGATCTTGGCGTACACCTTCCCAACGAC	Mg Chel H 2c
pf328	GACAGGGAAGAACATCCACGCTCTCG	Mg Chel H 3a
pf329	GAGTTCCAAGCTCGTTATGGGTTGAATGC	Mg Chel H 3b
pf331	GGATCCGTGATTTGAGCTCTTCAAGTCTGTCAGG	Mg Chl D FLR + Bam H1
pf332	TCTAGATGTCGTCGTCCCTAGTGTCC	Mg Chel H FL 5'+ Xba1
pf333	GGATCCCCTGTCAATCCCCTCAATCTTG	Mg Chel H FL 3' + BamH1
pf334	GGAAGGATCTGTACGTGAGCACTTG	Mg chel D central region Fwd
pf335	AGCCATTCTTTTGGCTCTCATGTC	Mg chel D central region Rev
pf336	ACGCGGTCATCAAAACTCATTGGAAG	Mg chel D 5' region Rev use with pf320 (xba1)
pf337	CAAGAAAGGTTTTTGTTGAAAAGACTGACAT	Mg chel D 3' region Fwd use with pf 331 (bamH1)
pf340	GTAAAACGACGGCCAGTG	M13 Fwd
pf341	GGAAACAGCTATGACCATG	M13 Rev
pf342	GTGGTGTCGGCTATGTACGAGGAC	Chel H seq loc start 3820 rev
pf344	ATGTTATCGGTGCCCCACAGG	Chel H seq loc start 3205 rev Tm=66.24
pf345	AAAGGAGCGTGACCTTGTCGTT	Chel H seq loc start 2450 Tm=64.44
pf346	TCAGACCTCAAGAAGGACGGCTAC	Chel H seq loc start 1686 Tm=65.19
pf347	GTCGAGGGCAGCTATGTTCACC	Chel H seq 2597 rev Tm=65.35
pf348	AGCGACTGGTACTCCCTCACATTC	Chel H seq 1837 rev Tm=65.01
pf349	GTACGACGACCCCGTTCTCTACCT	Chel H Chel H 917 fwd Tm=65.81
pf350	AGGTCGTGGGCTACCTCGT	Chel H seq 421 Tm=62.04
pf351	CTTGGGGTCCTTGAGCCTGTC	Chel H seq 1053 rev Tm=65.48
pf352*	TCTAGATGTCGTCGTCCCTAGTGTCCACCCC	CHLH FL fwd Tm=72.68+Xba1
pf368	CTACCAGCAGCGTCCACAACAAATA	directed mut Chl D re: Integrin1 DASGS->DAAGS
pf369	TATTTGTTGTGGACGCTGCTGGTAG	directed mut Chl D re: Integrin1 DASGS->DAAGS
pf372	CTCCTCTCTAGAATGGCTTCCACCTTCTC	Chel I coding + Xba1
pf373	GAACAGATCTGCTAAAGACTTCATAAAACTTCTC	Chel I coding + BgI2
pf374	TCTTCGGATCCTTAGAAAACTTCTTTGC	direct mut Chel D MH-g37 K311stop + Bam H1
pf375	TCTTCCACGGATCCGAAAACTTCTTTGC	DM chel D MH-g37 K311 stop + BamH1
pf392	GGATCCAGGTGGTGGCGGCGGGGGTG	CHL D N-term + poly-Proline region + BamH1 cutsite R
pf393	TCTAGATGAATCAAGATTCTTCAGAAGACCAAGATGAGGA	CHL D C-term + Xba1+ ATG
pf394	AGAATCTTGATTCTCCCTCCCATACCCGCGCCA	CHL D tape TP to C-terminus
pf395	GGGAGGGAGAATCAAGATTCTTCAGAAGACCAAGATGAGGA	CHL D tape C-terminus to TP
pf398	AGGAAGAGGCTCATGGACACC	C-terminal Zm CHLH fwd sequencing primer - to vector
pf399	TGTATCCACGGGATTTGCCAAG	CHLD sequencing primer to check 3' end plus vector tag
pf400	TCTAGATGGCGACGCCCACCG	CHLD 5'end FLF Xba1 - lower Tm than pf320
pf401	GGGTGGTTGCTCCCTCCCATACCCGCGCCA	CHLD Transit Peptide - to connect to (Poly P + C-terminus)
pf402	GGGAGGGAGCAACCACCCCCACCCCCG	CHLD Poly P + C-terminus - connect to the Transit peptide
pf403	TAATACGACTCACTATAGGG	T7
pf404	ATTTAGGTGACACTATAGAA	SP6
pf420	AGCTTTCTAGAATGGCTTCCACCTTCTC	CHLI coding + Xba1 - clone into pBin-GFP
pf421	GGAAGATCTGCTAAAGACTTCATAAAACTTCTC	CHLI coding + BgI2 - clone into pBin-GFP

Appendix B. Distribution of GO-terms for differentially expressed proteins in CHLI, CHLD and CHLH over-expressing leaf tissue

Sequence Distribution				
Graph Level	**GO Term**	**#Seq**	**Score**	**Parents**
1	biological_process	148	148.82	
2	metabolic process	126	109.39	biological_process
2	cellular process	126	88.61	biological_process
3	cellular metabolic process	90	54.15	metabolic process, cellular process
2	response to stimulus	41	42.72	biological_process
3	biosynthetic process	59	37.83	metabolic process
3	primary metabolic process	79	35.69	metabolic process
3	response to stress	35	35	response to stimulus
4	photosynthesis	32	32	cellular metabolic process
4	generation of precursor metabolites and energy	30	30	cellular metabolic process
4	transport	29	29	establishment of localization
6	translation	27	27	gene expression, cellular macromolecule biosynthetic process, cellular protein metabolic process
5	cellular protein metabolic process	39	24	protein metabolic process, cellular macromolecule metabolic process
4	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	25	23.08	primary metabolic process, cellular nitrogen compound metabolic process
4	protein metabolic process	46	22.4	primary metabolic process, macromolecule metabolic process
3	macromolecule metabolic process	49	20.28	metabolic process
3	establishment of localization	29	17.4	localization, biological_process
4	cellular macromolecule metabolic process	42	16.2	cellular metabolic process, macromolecule metabolic process

5	cellular macromolecule biosynthetic process	27	16.2	cellular macromolecule metabolic process, cellular biosynthetic process, macromolecule biosynthetic process
4	gene expression	27	16.2	macromolecule metabolic process
3	catabolic process	15	15	metabolic process
3	response to abiotic stimulus	14	14	response to stimulus
4	cellular nitrogen compound metabolic process	25	13.85	cellular metabolic process, nitrogen compound metabolic process
6	protein modification process	13	13	macromolecule modification, cellular protein metabolic process
2	cellular component organization	12	12.22	cellular component organization or biogenesis
3	response to biotic stimulus	12	12	response to stimulus
2	localization	29	10.44	biological_process
4	macromolecule biosynthetic process	27	9.72	macromolecule metabolic process, biosynthetic process
4	cellular biosynthetic process	27	9.72	cellular metabolic process, biosynthetic process
3	nitrogen compound metabolic process	25	8.31	metabolic process
4	macromolecule modification	13	7.8	macromolecule metabolic process
null	cellular component organization or biogenesis	12	7.33	biological_process

Appendix C. Under- and Over-expressed proteins in CHLI infiltrated leaf tissue

Unused	N	%Cov(95)	GI number	Blast2GO name	Biological process	Ratio I:WT	p-value	EF
2.1	950	4.08	19900	Osmotin-like protein	response to stress	1.2757	0.0114	1.19
5.97	421	53.93	20020	ribosomal protein	translation	1.4382	0.0018	1.24
21.92	111	29.95	396589	caffeic acid 3-o-methyl transferase	lignin biosynthetic process	1.3274	0.0155	1.20
2.95	729	4.80	548491	polygalacturonase	cellular wall organization	0.7403	3.06E-07	1.09
44.97	37	55.38	548746	ribosomal protein l12	translation	1.2409	2.96E-05	1.09
14.9	186	35.06	585876	60s ribosomal protein l23a-like	translation	0.7573	2.03E-06	1.09
4.51	526	10.83	730558	60s ribosomal protein l34	translation	0.7797	0.01	1.18
4.19	568	19.30	2497757	non-specific lipid transfer protein	response to biotic stimulus	1.6052	0.0051	1.31
2.19	909	8.75	2511737	caffeoyl-o-methyltransferase	lignin biosynthetic process	0.7583	2.61E-08	1.08
10.34	270	21.20	4580920	annexin p35	membrane organization	1.2038	0.0053	1.13
1.79	1055	1.75	34481799	cullin-1-like isoform1	ubiquitin-dependent protein catabolic process	0.8170	0.0012	1.10
6.9	370	18.09	37625025	60s ribosomal protein	translation	0.8071	0.0001	1.09
6.45	393	17.59	48375048	light-harvesting complex II protein lhcb3	photosynthesis	0.7918	5.91E-05	1.10
23.63	102	14.93	94730418	photosystem I p700 apoprotein a1	photosynthesis	0.6691	4.75E-09	1.09
4.27	554	16.95	116739352	formate dehydrogenase	oxidation reduction	1.4090	0.0011	1.05
10.39	268	50.49	117307343	histone h4	nucleosome assembly	0.6963	1.16E-10	1.07
9.48	286	37.93	148279693	farnesyl diphosphate synthase	isoprenoid biosynthetic process	1.3637	0.0413	1.34
10.9	252	31.91	158053020	histone h2	nucleosome assembly	0.7888	1.44E-11	1.06
11.84	235	71.91	159517163	superoxide dismutase	superoxide metabolic process	1.3089	7.20E-06	1.11
4.71	501	16.96	183393002	cytochrome c	electron transport chain	1.2815	0.0051	1.18
9.42	287	16.72	227473473	water channel protein	water transport	0.7855	0.0007	1.13

Appendix D. Under- and Over-expressed proteins in CHLD infiltrated leaf tissue

Unused	N	%Cov(95)	GI number	Blast2GO name	Biological process	Ratio D:WT	p-value	EF
16.63	159	49.06	19702	eukaryotic translation initiation factor 5a2	translation	0.7436	0.0073	1.23
36.44	51	57.35	19748	PSI reaction center subunit II	photosynthesis	1.2569	7E-07	1.09
11.64	237	50.56	19837	light harvesting chlorophyll a b-binding protein	photosynthesis	1.3576	2E-07	1.10
5.97	421	53.93	20020	ribosomal protein l12	translation	1.4941	2E-06	1.14
1.72	1074	44.94	170210	light harvesting chlorophyll a b-binding protein	photosynthesis	1.4990	0.0074	1.25
12.72	222	56.52	225237	ribosomal protein s19	translation	1.2725	8E-07	1.09
2.22	896	5.56	225796	peroxidase	lignin biosynthetic process	1.4050	0.0176	1.29
16.43	161	25.14	266936	50s ribosomal protein l27	translation	1.2658	0.0001	1.11
2.8	745	17.81	296473	beta-glucanase	response to biotic stimulus	1.3281	0.0055	1.18
4.57	521	36.81	396591	caffeic acid 3-o-methyltransferase	lignin biosynthetic process	0.6541	0.0101	1.33
31.89	68	26.32	401249	cytochrome b6-f complex iron-sulfur subunit	photosynthesis	1.5522	6E-28	1.06
13.86	201	48.97	407355	photosystem I psaH protein	photosynthesis	1.3191	2E-08	1.09
13.93	199	31.45	416681	atp synthase delta chain	ATP sythesis coupled proton transport	1.4305	4E-08	1.12
3.49	663	59.73	439577	elongation factor 1-alpha	translational elongation	0.6742	2E-06	1.15
13.52	206	95.96	548526	plastocyanin chloroplast	photosynthesis	1.5470	0	1.04
44.97	37	55.38	548746	ribosomal protein l12	translation	1.4875	9E-05	1.19
32.27	63	95.12	576417	ribulose bisphosphate carboxylase	photosynthesis	1.3218	2E-10	1.08
14.9	186	35.06	585876	60s ribosomal protein l23a-like	translation	0.5847	9E-08	1.14
23.54	104	51.05	632724	PSI reaction center subunit III	photosynthesis	1.5114	3E-14	1.10
4.51	526	10.83	730558	60s ribosomal protein l34	translation	0.5210	0.0104	1.54
28.81	77	33.67	2494076	nadp-dependent glyceraldehyde-3-phosphate dehydrogenase	glycolysis	0.7333	2E-09	1.09
4.19	568	19.30	2497757	non-specific lipid transfer protein	response to biotic stimulus	1.6613	0.003	1.29
14.97	185	31.16	2507328	40S ribosomal protein S6	translation	0.6925	2E-07	1.13
12.71	223	20.00	2632098	plastid-lipid-associated protein	fruit development	1.3089	8E-06	1.11
14.35	193	27.91	2764992	plasma-membrane associated cation-binding protein 1	DREPP2 protein - function unknown	1.2935	0.0027	1.17
5.02	478	4.39	2916727	polyphenol oxidase	oxidation reduction	1.2501	0.0028	1.13
6.3	402	11.83	3021512	isocitrate dehydrogenase	tricarboxylic acid cycle	0.7182	0.014	1.28
11.95	231	56.55	3219765	actin	cytoskeleton organization	0.7721	0.0025	1.16
2.08	956	21.17	3252854	glutamate dearboxylase	glutamate metabolic process	0.7403	0.0473	1.34
28.22	78	70.05	3493460	glycolate oxidase	photorespiration	1.2825	5E-14	1.06

Unused	N	%Cov(95)	GI number	Blast2GO name	Biological process	Ratio D:WT	p-value	EF
4.24	559	43.64	3717987	vacuolar atp synthase subunit g	ATP hydrolysis coupled proton transport	1.5780	0.026	1.47
55.45	23	59.73	3869088	elongation factor 1-alpha	translational elongation	0.6980	0.0002	1.17
5.61	441	7.92	3913240	magnesium-chelatase subunit chld	chlorophyll biosynthetic process	1.5127	0.0359	1.47
10.87	254	8.79	3914361	phospholipase d	phosphatidylcholine metabolic process	1.2004	0.0072	1.14
10.41	267	19.12	3914472	photosystem II 10 kda chloroplastic (psbR)	photosynthesis	1.3794	4E-18	1.06
7.12	356	18.18	3947727	proliferating cell nuclear antigen	regulation of DNA replication	0.7753	0.0015	1.14
10.34	270	21.20	4580920	annexin p35	membrane organization	1.3079	0.0055	1.19
12.48	225	20.70	6094336	spermidine synthase	spermidine biosynthetic process	0.7075	0.0016	1.21
3.55	656	3.46	7263569	pyruvate kinase	glycolysis	1.2196	0.3001	1.52
20.09	127	65.99	10798636	elongation factor 2-like	translational elongation	0.7006	2E-07	1.12
24.34	95	56.22	11994924	ribulose-bisphosphate carboxylase oxygenase large subur	photosynthesis	0.6572	0.0025	1.24
4.79	492	9.20	15226852	glutamate decarboxylase	glutamate metabolic process	0.7253	0.0431	1.36
3.14	705	1.73	15227525	calmodulin-domain protein kinase cdpk isoform 2	protein amino acid phosphorylation	1.3707	0.0333	1.33
2.94	731	0.83	20177485	1a protein	RNA processing	0.6976	0.0029	1.22
2.29	876	3.67	23343585	serine acetyltransferase	cysteine biosynthetic process from serine	1.4182	0.0488	1.41
6.26	408	51.52	25990286	chloroplast protein 12	regulation of oxidoreductase activity	0.4860	0.0321	1.90
4.35	549	1.14	28268678	respiratory burst oxidase	oxidation reduction	1.2488	0.0252	1.21
5.79	430	7.93	28866015	calcium-dependent protein kinase	protein amino acid phosphorylation	0.6767	0.0399	1.44
18.79	136	29.28	30013659	at1g31330 t19e23_1 (psaF)	photosynthesis	1.4262	2E-19	1.07
10.91	251	19.07	31580934	ribosomal protein s2	translation	1.2128	0.0005	1.10
17.11	152	34.29	31711507	germin-like protein	auxin mediated signaling pathway	1.2746	1E-18	1.05
8.08	322	47.69	33340517	s-adenosylmethionine synthetase	one-carbon metabolic process	0.7351	0.0026	1.19
1.79	1055	1.75	34481799	cullin-1-like isoform 1	ubiquitin-dependent protein catabolic process	0.6083	0.0019	1.28
27.06	87	28.79	37625023	60s ribosomal protein l3	translation	0.7114	0.0023	1.22
6.9	370	18.09	37625025	60s ribosomal protien l3-like	translation	0.7273	0.0025	1.20
55.33	24	30.90	38154485	molecular chaperone hsp90-1	protein folding	0.6303	0.0002	1.23
64.34	15	52.01	38325815	heat shock protein 70	response to stress	0.7694	8E-10	1.08
13.00	213	15.33	40036995	tubulin beta	microtubule polymerization	0.7589	0.0004	1.15
38.46	49	89.62	40455718	plastocyanin chloroplast	photosynthesis	1.4223	1E-26	1.05

Unused	N	%Cov(95)	GI number	Blast2GO name	Biological process	Ratio D:WT	p-value	EF
16.03	171	30.59	40455722	light-harvesting complex I chlorophyll a b binding protein	photosynthesis	1.5367	5E-17	1.08
5.45	449	20.19	46020014	dehydrin	response to stress	0.6158	0.0416	1.58
16.75	155	36.03	48474196	stromal 70 kda heat shock-related chlorplastic-like	protein folding	1.2762	0.0057	1.18
14.61	189	30.59	51949800	adenosine kinase	AMP biosynthetic process	0.7781	0.0079	1.17
26.39	90	38.37	53854350	acetyl-c-acetyltransferase	isoprenoid biosynthetic process	0.7678	3E-05	1.12
3.26	688	44.18	55977763	atp synthase cf1 alpha subunit	ATP synthesis coupled proton transport	1.4366	0.0197	1.31
2.06	966	16.67	57791555	ubiquitin extension protein	translation	0.6533	0.0106	1.32
38.77	48	49.12	58700507	chloroplast oxygen-evolveing protein 16 dka subunit	photosynthesis	1.6549	1E-11	1.11
148	2	88.35	60391818	atpase beta subunit	ATP synthesis coupled proton transport	1.5107	9E-12	1.11
2.02	985	2.98	68300897	soluble starch synthase I	starch biosynthetic process	0.7734	0.0416	1.24
6.39	397	25.00	75248032	ribosomal protein s16	translation	1.2995	0.0117	1.21
11.54	240	14.25	75303624	nadh dehydrogenase subunit 7	photosynthesis	1.2184	0.0032	1.13
3.82	623	1.77	75326538	two-pore calcium channel	calcium ion transmembrane transport	0.7350	0.0046	1.19
18.65	139	34.48	76262913	40s ribosomal protein	translation	0.6103	3E-12	1.11
3.22	692	1.48	76880152	dna polymerase	plastid DNA replication	1.3400	0.0328	1.30
20.36	125	81.20	78102541	atp synthase cf1 epsilon subunit	ATP synthesis coupled proton transport	1.3241	1E-18	1.05
47.13	32	50.00	78102549	cytochrome f	photosynthesis	1.4516	0	1.04
12.1	230	17.85	81301541	photosystem II protein d1	photosynthesis	1.2401	6E-24	1.04
28.82	76	25.21	81301558	photosystem II protein d2	photosynthesis	1.2336	2E-26	1.03
45.72	36	34.71	81301559	photosystem Ii cp43 chlorophyll apoprotein	photosynthesis	1.4299	0	1.04
40.67	44	37.20	81301594	photosystem II 47 kda protein	photosynthesis	1.2942	1E-23	1.05
6.47	391	41.10	81301597	photosystem Ii phosphoprotein	photosynthesis	1.7901	6E-11	1.15
0	200	26.05	81301598	cytochrome b6	photosynthesis	1.2619	3E-07	1.08
2.05	971	11.11	81301629	photosystem I subuit VII	photosynthesis	1.3131	0.0005	1.12
13.84	202	36.77	81301641	ribosomal protein s7	translation	1.2282	0.0021	1.13
4.72	500	1.52	82570694	glycoprotein precursor		1.3129	0.0227	1.26
51.63	28	36.86	84620802	photosystem Ii 22 kda protein	photosynthesis	1.3180	1E-07	1.09
4.26	555	35.74	84620804	chloroplast photosystem II 22 kda component	photosynthesis	1.3479	1E-09	1.07
17.4	145	55.17	85701227	ribosomal protein s15	translation	1.2520	0.0004	1.12

Unused	N	%Cov(95)	GI number	Blast2GO name	Biological process	Ratio D:WT	p-value	EF
8.32	313	20.62	90823167	transcription factor btf3	regulation of translational initiation	0.7996	0.0218	1.21
1.34	1191	6.72	91118960	agamous-like protein	regulation of transcription, DNA-dependent	0.5692	0.0449	1.70
14.38	192	27.52	91207772	ribosomal protein s3	translation	1.2351	0.0038	1.15
35.36	54	61.41	94730361	atp synthase c/0 b subunit	ATP synthesis coupled proton transport	1.2743	7E-05	1.11
13.52	208	46.72	94730419	ribosomal protein l14	translation	1.3170	2E-10	1.07
4.26	556	0.64	109690139	replicase	RNA processing	1.4534	0.0375	1.42
52.8	27	47.02	110377766	chlorophyll a-b binding protein	photosynthesis	1.3917	3E-18	1.06
10.72	259	13.49	110623121	atp-dependent clp protease proteolytic subunit	proteolysis	1.3568	6E-05	1.13
7.01	360	1.04	116047933	myosin	cytoskeleton organization	1.2026	0.0104	1.15
10.39	268	50.49	117307343	histone h4	nucleosome assembly	0.6883	8E-10	1.08
14.56	190	17.77	119866037	importin subunit alpha-1	protein import into nucleus	0.7489	0.0102	1.23
27.25	84	35.76	131053151	1-aminocyclopropane-1-carboxylate oxidase	etylene biosynthetic process	0.7012	6E-05	1.13
1.4	1172	4.02	145234140	o-methyltransferase	methylation	1.2862	0.0124	1.06
3.88	613	1.15	145334819	myosin 2	cytoskeleton organization	1.2666	0.0232	1.22
7.83	331	14.98	152206078	isomerase peptidyl-prolyl cis-trans isomerase	protein peptidyl-prolyl isomerization	0.6605	0.0002	1.21
23.65	101	49.83	157142955	protein thylakoid chloroplastic-like	thylakoid membrane organization	1.2485	2E-08	1.07
10.9	252	31.91	158053020	histone h2	nucleosome assembly	0.7589	2E-05	1.13
11.84	235	71.91	159517163	superoxide dismutase	superoxide metabolic process	1.3686	2E-06	1.12
2.58	794	0.86	163914237	rr-like protein	signal transduction	0.8000	0.0015	1.12
29.08	75	35.25	175363751	elongation factor 1-gamma-like	translational elongation	0.7014	5E-11	1.09
4.71	501	16.96	183393002	cytochrome c	electron transport chain	1.5577	1E-08	1.09
6.19	413	9.94	189096126	50s ribosomal protein l1	translation	1.2474	5E-05	1.10
9.03	295	22.73	189096130	50s ribosomal protein l5	translation	1.2429	2E-06	1.08
12.97	214	23.77	189096131	50s ribosomal protein l6	translation	1.2260	7E-07	1.07
6.59	386	13.87	189096148	50s ribosomal protein l29	translation	1.4091	1E-06	1.12
1.48	1143	3.64	211586480	retrotransposon protein	DNA integration	1.9535	0.0347	1.59
27.45	82	40.96	219913732	malate dehydrogenase	tricarboxylic acid cycle	0.7864	5E-10	1.07
11.63	238	23.55	219934656	RNA-binding protein cp31	mRNA processing	0.6362	5E-14	1.07
7.54	341	21.90	219934658	nucleic acid-binding protein	mRNA processing	0.7475	0.0068	1.22

Appendix E. Under- and Over-expressed proteins in CHLH infiltrated leaf tissue

Unused	N	%Cov(95)	GI number	Blast2go name	Biological process	Ratio H:WT	p-value	EF
2.1	950	4.08	19900	osmotin-like protein	response to stress	1.2685	0.0047	1.15
5.97	421	53.93	20020	ribosomal protein l12	translation	1.3363	2E-07	1.09
16.43	161	25.14	266936	50s ribosomal protein l27	translation	1.2428	1E-05	1.08
4.57	521	36.81	396591	caffeic acid 3-o-methyltransferase	lignin biosynthetic process	0.7812	0.0006	1.10
13.86	201	48.97	407355	photosystem I psaH protein	photosynthesis	1.2191	1E-09	1.06
2.82	742	41.59	407769	PSI reaction center subunit II (D1)	photosynthesis	1.2056	0.0325	1.18
3.02	719	22.08	443960	40s ribosomal protein s9	translation	0.7625	0.0234	1.20
13.52	206	95.96	548526	plastocyanin chloroplast	photosynthesis	1.2268	1E-34	1.03
44.97	37	55.38	548746	ribosomal protein l12	translation	1.2434	0.0025	1.14
14.9	186	35.06	585876	60s ribosomal protein	translation	0.7663	6E-08	1.07
4.51	526	10.83	730558	60s ribosomal protein l34	translation	0.6888	0.0086	1.27
6.46	392	4.63	2300480	l-ascorbate oxidase homolog	oxidation reduction	1.2140	0.0205	1.17
4.19	568	19.30	2497757	non-specific lipid transfer protein	response to biotic stimulus	1.5655	0.0024	1.24
6.3	402	11.83	3021512	isocitrate dehydrogenase	tricarboxylic acid cycle	0.8254	0.0414	1.20
4.24	559	43.64	3717987	vacuolar atp synthase subunit g	ATP hydrolysis coupled proton transport	1.2224	0.0327	1.20
2.69	764	8.08	6014908	dihydrodipicolinate synthase	diaminopimelate biosynthetic process	1.2027	0.0105	1.14
2.12	939	23.83	13603408	isopentenyl diphosphate isomerase	isoprenoid biosynthetic process	1.3378	0.0359	1.30
4.58	520	26.43	20086364	transcriptional coactivator-like protein	positive regulation of transcription, DNA depend	1.2252	0.0053	1.13
15.44	179	26.57	24637568	ankyrin repeat domain	ankyrin repeat-mediated complex assembly	1.2286	4E-05	1.09
6.26	408	51.52	25990286	chloroplast protein 12	regulation of oxidoreductase activity	0.6402	0.0013	1.25
4.11	581	21.62	27529852	histone 2a	nucleosome assembly	0.7912	3E-06	1.07
5.79	430	7.93	28866015	calcium dependent protein kinase	protein amino acid phosphorylation	0.7599	0.0081	1.17
3.36	683	56.67	30013663	ribulose bisphosphate carboxylase (small)	photosynthesis	0.7943	0.0264	1.22
10.91	251	19.07	31580934	ribosomal protein s2	translation	1.2090	4E-05	1.08
17.11	152	34.29	31711507	germin-like protein	auxin mediated signaling pathway	1.2301	1E-21	1.03
1.79	1055	1.75	34481799	cullin-1-like isoform 1	ubiquitin-dependent protein catabolic process	0.7833	0.0004	1.09
3.78	627	8.42	38325811	heat shock 70 kda protein	response to stress	1.4372	0.0047	1.25
16.03	171	30.59	40455722	light-harvesting complex I chlorophyll a b bind prot 3	photosynthesis	1.2481	3E-13	1.05
2.42	828	37.21	44917161	14-3-3 protein	regulation of molecular function	0.8153	0.003	1.08

Unused	N	%Cov(95)	GI number	Blast2go name	Biological process	Ratio H:WT	p-value	EF
5.45	449	20.19	46020014	dehydrin	response to stress	0.6651	0.0112	1.32
6.45	393	17.59	48375048	light-harvesting complex II protein lhcb3	photosynthesis	0.8191	0.0002	1.09
16.75	155	36.03	48474196	stromal 70 kda heat shock-related chloroplastic-like	protein folding	1.2749	0.0002	1.13
38.77	48	49.12	58700507	chloroplast oxygen-evolving protein 16 kda subunit	photosynthesis	1.2732	5E-07	1.09
7.37	347	13.97	75138727	histone h3	nucleosome assembly	0.7348	1E-09	1.07
6.39	397	25.00	75248032	ribosomal protein s16	translation	1.2168	0.0048	1.13
6.47	391	41.10	81301597	photosystem II phosphoprotein	photosynthesis	1.4133	3E-10	1.09
2.05	971	11.11	81301629	photosystem I subunit VII	photosynthesis	1.2154	0.0109	1.15
17.4	145	55.17	85701227	ribosomal protein s15	translation	1.2302	2E-05	1.09
5.64	440	6.50	88659685	metacaspase type II	proteolysis	1.2012	0.0222	1.17
23.63	102	14.93	94730418	photosystem I p700 apoprotein a1	photosynthesis	0.6562	2E-09	1.09
13.52	208	46.72	94730419	ribosomal protein l14	translation	1.2128	2E-09	1.06
4.26	556	0.64	109690139	replicase	RNA processing	1.3279	0.0292	1.28
10.39	268	50.49	117307343	histone h4	nucleosome assembly	0.6709	1E-10	1.08
27.25	84	35.76	131053151	1-aminocyclopropane-1-carboxylate oxidase	ethylene biosynthetic process	0.8241	0.0024	1.12
10.9	252	31.91	158053020	histone h2	nucleosome assembly	0.7562	3E-10	1.08
4.71	501	16.96	183393002	cytochrome c	electron transport chain	1.2961	1E-06	1.07
6.19	413	9.94	189096126	50s ribosomal protein l1	translation	1.2421	3E-05	1.09
9.03	295	22.73	189096130	50s ribosomal protein l5	translation	1.2487	4E-09	1.06
12.97	214	23.77	189096131	50s ribosomal protein l6	translation	1.2032	2E-06	1.07
6.59	386	13.87	189096148	50s ribosomal protein l29	translation	1.3115	1E-06	1.10
17.54	144	39.73	219934652	RNA-binding protein	mRNA processing	1.2646	0.0313	1.23
4.03	595	41.58	219934654	RNA-binding protein	mRNA processing	1.4184	5E-05	1.14
9.42	287	16.72	227473473	water channel protein	water transport	0.8148	0.0124	1.17
2.15	927	3.06	227476174	inorganic phosphate transporter 1-1	phosphate transport	1.2810	0.001	1.11

References

Albiniak A, Baglieri J, Robinson C (2012) Targeting of lumenal proteins across the thylakoid membrane. Journal of Experimental Botany 63: 1689-1698

Aldridge C, Cain P, Robinson C (2009) Protein transport in organelles: Protein transport into and across the thylakoid membrane. FEBS Journal 276: 1177-1186

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignments search tool. Journal of Molecular Biology 215: 403-410

America T, Hageman J, Guera A, Rook F, Archer K, Keegstra K (1994) Methotrexate does not block import of a DHFR fusion protein into chloroplasts. Plant Molecular Biology 24: 283-294

Andres C, Agne B, Kessler F (2010) The TOC complex: Preprotein gateway to the chloroplast. Biochimica et Biophysica Acta 1803: 715-723

Andrew T, Pascale G, Dailey H (1990) Regulation of heme biosynthesis in higher animals. In: Dailey H (ed) Biosynthesis of Heme and Chlorophylls. McGraw-Hill, New York, pp 163-200

Apel K (1981) The protochlorophyllide holochrome of barley (*Hordeum vulgare* L.). Phytochrome-induced decrease of translatable mRNA coding for the NADPH:protochlorophyllide oxidoreductase. European Journal of Biochemistry 120: 89-93

Apel K, Hirt H (2004) Reactive oxygen species: metabolism, oxidative stress, and signal transduction. Annual Review of Plant Biology 55: 373-399

Asakura Y, Kikuchi S, Belcher S, Osborne E, Yano S, Terashima I, Barkan A, Nakai M (2004) Maize mutants lacking chloroplast FtsY exhibit pleiotropic defects in the biogenesis of thylakoid membranes. Plant Cell 16: 201-214

Axelsson E, Lundqvist J, Sawicki A, Nilsson S, Schroder I, Al-Karadaghi S, Willows R, Hansson M (2006a) Recessiveness and dominance in barley mutants deficient in Mg-chelatase subunit D, an AAAprotein involved in chlorophyll biosynthesis. Plant Cell 18: 3606-3616

Axelsson E, Lundqvist J, Sawicki A, Nilsson S, Schroder I, Al-Karadaghi S, Willows RD, Hansson M (2006b) Recessiveness and dominance in barley mutants deficient in Mg-chelatase subunit D, an AAA protein involved in chlorophyll biosynthesis. Plant Cell 18: 3606-3616

Balsera M, Soll J, Buchanan (2010) Redox extend its regulatory reach to chloroplast protein import. Trends in Plant Science 15: 515-521

Batschauer A, Apel K (1984) An inverse control by phytochrome of the expression of two nuclear genes in barley. European Journal of Biochemistry 143: 593-597

Bauer J, Chen K, Hiltbrunner A, Wehrli E, Eugster M, Schnell D, Kessler F (2000) The major protein import receptor of plastids is essential for chloroplast biogenesis. Nature 403: 203-207

Beale S (1999) Enzymes of chlorophyll biosynthesis. Photosynthesis Research 60: 43-73

Beale S, Weinstein J (1990) Tetrapyrrole metabolism in photosynthetic organisms. In: Dailey H (ed) Biosynthesis of Heme and Chlorophylls. McGraw-Hill, New York, pp 287-391

Beale S, Weinstein J (1991) Biochemistry and regulation of photosynthetic pigment formation in plants and algae. In: Jordan P (ed) Biosynthesis of Tetrapyrroles. Elsevier, pp 155-235

Becker T, Hritz J, Vogel M, Caliebe A, Bukau B, Soll J, Schlieff E (2004a) Toc12, a Novel Subunit of the Intermembrane Space Preprotein Translocon of Chloroplasts. Molecular Biology of the Cell 15: 5130-5144

Becker T, Jelic M, Vojita A, Radunz A, Soll J, Schleiff E (2004b) Preprotein recognition by the Toc complex. EMBO Journal 23: 520-530

Bedard J, Jarvis P (2005) Recognition and envelope translocation of chloroplast preproteins. Journal of Experimental Botany 56: 2287-2320

Bedard J, Kubis S, Bimanadham S, Jarvis P (2007) Functional similarity between the chloroplast translocon component, Tic40, and the human co-chaperone, Hsp70-interacting protein (Hip). Journal of Biological Chemistry 282: 21404-21414

Bendahmane A, Farnham G, Moffett P, Baulcombe D (2002) Constitutive gain-of-function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. The Plant Journal 32: 195-204

Bendahmane A, Querci M, Kanyuka K, Baulcombe D (2000) *Agrobacterium* transient expression system as a tool for the isolation of disease resistance genes: application to the *Rx2* locus in potato. The Plant Journal 21: 73-81

Bergdoll M, Remy M, Cagnon C, Masson J, Dumas P (1997) Proline-dependent oligomerization with arm exchange. Structure 5: 391-401

Bhushan S, Lefebvre B, Stahl A, Wright S, Bruce B, Boutry M, Glaser E (2003) Dual targeting and function of a protease in mitochondria and chloroplasts. EMBO Reports 4: 1073-1077

Bolter B, May T, Soll J (1998) A protein import receptor in pea chloroplasts, Toc86, is only a proteolytic fragment of a larger polypeptide. FEBS Letters 441: 59-62

Boodley J, Sheldrake R, Jr (1982) Cornell Peat-Lite Mixes for Commercial Plant Growing. In: Extension CC (ed). Cornell Cooperative Extension, pp 1-8

Bougri O, Grimm B (1996) Members of a low-copy number gene family encoding glutamyl-tRNA reductase are differentially expressed in barley. Plant Journal 9: 867-878

Bradbeer J, Atkinson Y, Borner T, Hagemann R (1979) Cytoplasmic synthesis of plastid polypeptide may be controlled by plastid-synthesized RNA. Nature 279: 816-817

Bradford M (1976) A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. Analytical Biochemistry 72: 248-254

Breitenbach J, Zhu C, Sandmann G (2001) Bleaching herbicide norflurazon inhibits phytoene desaturase by competition with the cofactors. Journal of Agricultural Food Chemistry 49: 5270-5272

Bukau B, Deuerling E, Pfund C, Craig E (2000) Getting newly synthesized proteins into shape. Cell 101: 119-122

Casado-Vela J, Martinez-Esteso M, Rodriguez E, Borras E, Elortza F, Bru-Martinez R (2010) iTRAQ-based quantitative analysis of protein mixtures with large fold change and dynamic range. Proteomics 10: 343-347

Castelfranco P, Weinstein J, Schwarcz S, Pardo A, Wezelman B (1979) The Mg insertion step in chlorophyll biosynthesis. Archives of Biochemistry and Biophysics 192: 592-598

Castelfranco PA, Zeng X (1991) Regulation of 5-Aminolevulinic Acid Synthesis in Developing Chloroplasts. Plant Physiology 97: 1-6

Castells E, Portoles S, Huang W, Mas P (2010) A functional connection between the clock component TOC1 and abscisic acid signaling pathways. Plant Signaling & Behavior 5: 409-411

Chaves I, Pokorny R, Byrdin M, Hoang N, Ritz T, Brettel K, Essen L, van der Horst G, Batschauer A, Ahmad M (2011) The cryptochromes: blue light photoreceptors in plants and animals. Annual Review of Plant Biology 62: 335-364

Chen K, Li H (2007) Precursor binding to an 880-kDa Toc complex as an early step during active import of protein into chloroplasts. Plant Journal 49: 149-158

Chen X, Smith M, Fitzpatrick L, Schnell D (2002) *In vivo* analysis of the role of atTic20 in protein import into chloroplasts. The Plant Cell 14: 641-654

Chevreux B, Pfisterer T, Drescher B, Driesel A, Muller W, Wetter T, Suhai S (2004) Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Research 14: 1147-1159

Chory J, Chatterjee M, Cook R, Elich T, Fankhauser C, Li J, Nagpal P, Neff M, Pepper A, Poole D, Reed J, Vitart V (1996) From seed germination to flowering, light controls plant development via the pigment phytochrome. Proceedings of the National Academy of Sciences 93: 12066-12071

Chou M, Chu C, Chen L, Akita M, Li H (2006) Stimulation of transit-peptide release and ATP hydrolysis by a cochaperone during protein import into chloroplasts. Journal of Cell Biology 175: 893-900

Cline K (2000) Gateway to the chloroplast. Nature 403: 148-149

Conesa A, Gotz S (2007) Blast2GO: A Comprehensive Suite for Functional Analysis in Plant Genomics. International Journal of Plant Genomics 2008: 619832

Conesa A, Gotz S, Garcia-Gomex J, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 21: 3675-3676

Conti A, Pancaldi S, Fambrini M, Michelotti V, Bonora A, Salvini M, Pugliesi C (2004) A Deficiency at the Gene Coding for ζ-Carotene Desaturase Characterizes the Sunflower *non dormant-1* Mutant. Plant Cell Physiology 45: 445-455

Cornah JE, Terry MJ, Smith AG (2003) Green or red: what stops the traffic in the tetrapyrrole pathway? Trends in Plant Science 8: 224-230

Davis S, Bhoo S, Durski A, Walker J, Vierstra R (2001) The heme-oxygenase family required for phytochrome chromophore biosynthesis is necessary for proper photomorphogenesis in higher plants. Plant Physiology 126: 656-669

Davis S, Kurepa J, Vierstra R (1999) The *Arabidopsis thaliana HY1* locus, required for phytochrome-chromophore biosynthesis, encodes a protein related to heme oxygenases. Proceedings of the National Academy of Sciences 96: 6541-6546

Davison P, Schubert H, Reid J, Iorg C, Heroux A, Hill C, Hunter C (2005) Structural and biochemical characterization of Gun4 suggests a mechanism for its role in chlorophyll biosynthesis. Biochemistry 44: 7603-7612

Della-Cioppa G, Bauer S, Klein B, Shah D, Fraley R, Kishore G (1986) Translocation of the precursor of 5-enolpyruvylshikimate-3-phosphate synthase into chloroplasts of higher plants *in vitro*. Proceedings of the National Academy of Sciences 83: 6893-6877

Duggan J, Gassman M (1974) Induction of porphyrin synthesis in etiolated bean leaves by chelators of iron. Plant Physiology 53: 206-215

Duy D, Wanner G, Meda A, von Wiren N, Soll J, Philippar K (2007) PIC1, an Ancient Permease in *Arabidopsis* Chloroplasts, Mediates Iron Transport. The Plant Cell 19: 986-1006

Emanuelsson O, Nielsen H, von Heijne G (1999) ChloroP, a neural network-based method for predicting chlorplast trnasit peptides and their cleavage sites. Protein Science 8: 978-984

Falciatore A, Merendino L, Barneche F, Ceol M, Meskauskiene R, Apel K, Rochaix J-D (2005) The FLP proteins act as regulators of chlorophyll synthesis in response to light and plastid signals in Chlamydomonas. Genes and Development 19: 176-187

Ferreira G, Franco R, Lloyd S, Moura I, Moura J, Huynh B (1995) Structure and function of ferrochelatase. Journal of Bioenergetics and Biomembranes 27: 221-229

Figueiredo J, Romer P, Lahaye T, Graham J, White F, Jones J (2011) *Agrobacterium*-mediated transient expression in citrus leaves: a rapid tool for gene expression and functional gene assay. Plant Cell Reports 30: 1339-1345

Filichkin S, Gelvin S (1992) Effect of Dimethylsulfoxide Concentration on Specificity of Primer Matching in PCR. BioTechniques 12: 828-830

Fodje M, Hansson A, Hansson M, Olsen J, Gough S, Willows R, Al-Karadaghi S (2001) Interplay between an AAA module and an integrin I domain may regulate the function of magnesium chelatase. Journal of Molecular Biology 311: 111-122

Forreiter C, van Cleve B, Schmidt A, Apel K (1990) Evidence for a general light-dependent negative control of NADPH-protochlorophyllide oxidoreductase in angiosperms. Planta 183: 126-132

Frackman S, Kobs G, Simpson D, Storts D (1998) Betaine and DMSO: Enhancing Agents for PCR. Promega Notes 65: 27

Freire M (2005) Translation initiation factor (iso) 4E interacts with BTF3, the beta subunit of the nascent polypeptide-associated complex. Gene 345: 271-277

Fryer M, Ball L, Oxborough K, Karpinski S, Mullineaux P, Baker N (2003) Control of ascorbate peroxidase 2 expression by hydrogen peroxide and leaf water status during excess light stress reveals a functional organisation of *Arabidopsis* leaves. Plant Journal 33: 691-705

Fuesler T, Castelfranco P, Wong Y (1984a) Formation of Mg-containing chlorophyll precursors from protoporphyrin IX, ∂-aminolevulinic acid, and

glutamate in isolated, photosynthetically competent, developing chloroplasts. Plant Physiology 74: 928-933

Fuesler T, Wright Jr. L, Castelfranco P (1981) Properties of the magnesium chelatase in greening etioplasts. Metal ion specificity and effect of substrate concentration. Plant Physiology 67: 246-249

Fuesler TP, Wong Y-S, Castelfranco PA (1984b) Localization of Mg-Chelatase and Mg-Protoporphyrin IX Monomethyl Ester (Oxidative) Cyclase Activities within Isolated, Developing Cucumber Chloroplasts'. Plant Physiology 75: 662-664

Fulgosi H, Soll J (2002) The chloroplast protein import receptors Toc34 and Toc159 are phosphorylated by distinct protein kinases. Journal of Biological Chemistry 277

Gadjieva R, Axelsson E, Olsson U, Hansson M (2005) Analysis of gun phenotype in barley magnesium chelatase and Mg-protoporphyrin IX monomethyl ester cyclase mutants. Plant Physiology and Biochemistry 43: 901-908

Gibson L, Marrison J, Leech R, Jensen P, Bassham D, Cibson M, Hunter C (1996a) A putative Mg-chelatase subunit from *Arabidopsis thaliana* cv C24. Plant Physiology 111: 61-71

Gibson L, Marrison J, Leech R, Jensen P, Bassham D, Gibson M, Hunter C (1996b) A putative Mg chelatase subunit from Arabidopsis thaliana cv C24. Sequence and transcript analysis of the gene, import of the protein into chloroplasts, and in situ localization of teh transcript and protein. Plant Physiology 111: 61-71

Gisk B, Yasui Y, Kohchi T, Frankenberg-Dinkel N (2010) Characterization f the haem oxygenase protein family in *Arabidopsis thaliana* reveals a diversity of functions. Biochem. J. 425: 425-434

Goldschmidt-Clermont M (1998) Coordination of nuclear and chloroplast gene expression in plant cells. International Review of Cytology 177: 155-180

Goslings D, Meskauskiene R, Kim C, Lee KP, Nater M, Apel K (2004) Concurrent interactions of heme and FLU with Glu tRNA reductase (HEMA1), the target of metabolic feedback inhibition of tetrapyrrole biosynthesis, in dark- and light-grown Arabidopsis plants. The Plant Journal 40: 957-967

Gray J, Sornarajah R, Zabron A, Duckett C, Khan M (1995) Chloroplast control of nuclear gene expression. In: Mathis P (ed) Photosynthesis: From Light to Biosphere. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp 543-550

Griffin T, Xie H, Bandhakavi S, Popko J, Mohan A, Carlis J, Higgins L (2007) iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. Journal of Proteome Research 6: 4200-4209

Guo R, Luo M, Weinstein (1998) Magnesium-Chelatase from Developing Pea Leaves. Plant Physiology 116: 605-615

Gutensohn M, Fan E, Frielingsdorf S, Hanner P, Hou B, Hust B, Klosgen R (2006) Toc, Tic, Tat *et al.*: structure and function of protein transport machineries in chloroplasts. Journal of Plant Physiology 163: 333-347

Hageman J, Robinson C, Smeekens S, Weisbeek P (1986) A thylakoid processing protease is required for complete maturation of the lumen protein plastocyanin. Nature 324: 567-569

Hanson P, Whiteheart S (2005) AAA+ proteins: have engine, will work. Nature Reviews Molecular Cell Biology 6: 519-529

Hansson A, Willows R, Roberts T, Hansson M (2002) Three semidominant barley mutants with single amino acid substitutios in the smallest magnesium chelatase subunit form defective AAA+ hexamers. Proceedings of the National Academy of Sciences 99: 13944-13949

Hansson M, Kannangara C (1997) ATPases and phosphate exchange activities in magnesium chelatase subunits of Rhodobacter sphaeroides. Proceedings of the National Academy of Sciences 94: 13351-13356

Heim R, Prasher D, Tsien R (1994) Wavelength mutations and posttranslational autoxidation of green fluorescent protein. Proceedings of the National Academy of Sciences 91: 12501-12504

Heins L, Mehrle A, Hemmler R, Wagner R, Kuchler M, Hormann F, Sveshnikov D, Soll J (2002) The preprotein conducting channel at the inner envelope membrane of plastids. The EMBO Journal 21: 2616-2625

Henningsen K, Boynton J, von Wettstein D (1993) Mutants at *xantha* and *albina* Loci in Relation to Chloroplast Biogenesis in Barley (*Hordeum vulgare* L.) The Royal Danish Academy of Sciences and Letters, Copenhagen, Denmark

Hiltbrunner P, Bauer J, Vidi P, Infanger S, Weibel P, Hohwy M, Kessler F (2001) Targeting of an abundant cytosolic form of the protein import receptor at Toc159 to the outer chloroplast membrane. Journal of Cell Biology 154: 309-316

Hirohashi T, Hase T, Nakai M (2001) Maize non-photosynthetic ferredoxin precursor is mis-sorted to the intermembrane space of chloroplasts in the presence of light. Plant Physiology 125: 2154-2163

Hirsch S, Muckel E, Heemeyer F, von Heijne G, Soll J (1994) A receptor component of the chloroplast protein translocation machinery. Science 266: 1989-1992

Holm M, Ma L, Qu L, Deng X (2002) Two interacting bZIP proteins are direct targets of COP1-mediated controll of light-dependent gene expression in *Arabidopsis*. Genes and Development 16: 1247-1259

Hoober J, Kahn A, Ash D, Hough S, Kannangara C (1988) Biosynthesis of ∂-aminolevulinate in greening barley leaves. IX. Structure of the substrate, mode of gabaculine inhibition, and the catalytic mechanism of glutamate-1-semialdehyde aminotransferase. Carlsberg Research Communications 53: 11-25

Hsu D, Zhao X, Zhao S, Kazantsev A, Wang R, Todo T, Wei Y, Sancar A (1996) Putative human blue-light photoreceptors hCRY1 and hCRY2 are flavoproteins. Biochemistry 35: 13871-13877

Hurwitz j, Weissbach A, Horecker B, Smyrniotis P (1956) Spinach Phosphoribulokinase. The Journal of Biological Chemistry 218: 769-783

Hynes R (2002) Integrins: bidirectional, allosteric signaling machines. Cell 110: 673-687

Ilag L, Kumar A, Soll D (1994) Light regulation of chlorophyll biosynthesis at the level of 5-aminolevulinate formation in Arabidopsis. Plant Cell 6: 265-275

Inaba T, Li M, Alvarez-Huerta M, Kessler F, Schnell D (2003) atTic110 functions as a scaffold for coordinating the stromal events of protein import into chloroplasts. Journal of Biological Chemistry 278: 38617-38627

Ivanova Y, Smith M, Chen K, Schnell D (2004) Members of the Toc159 import receptor family represent distinct pathways for protein targeting to plastids. Molecular Biology of the Cell 15: 3379-3392

Jackson-Constan D, Akita M, Keegstra K (2001) Molecular chaperones involved in chloroplast protein import. Biochimica et Biophysica Acta 1541: 102-113

Jaffe E (1993) Predicting the Zn(II) ligands in metalloproteins: Case study, porphobilogen synthase. . Comments Inorg Chem 15: 67-92

Janssen B, Gardner R (1990) Localized transient expression of GUS in leaf discs following cocultivation with Agrobacterium. Plant Molecular Biology 14: 61-72

Jarvis P (2008) Targeting of nucleus-encoded proteins to chloroplasts in plants. New Phytologist 179: 257-285

Jarvis P, Robinson C (2004) Mechanisms of protein import and routing in chloroplasts. Current Biology 14: R1064-R1077

Jeffrey S, Mantoura R, Bjornland T (1997) Data for the identification of 47 key phytoplankton pigments. In: Wright S, Jeffrey S, Mantoura R (eds) Phytoplankton pigments in oceanography: Guidelines to modern methods. UNESCO Publishing, Paris, France

Jensen P, Gibson L, Hunter C (1998) Determinants of catalytic activity with the use of purified I, D, and H subunits of the magnesium protoporphyrin IX chelatase from Synechocystis PCC6803. Biochemical Journal 334: 335-344

Jensen P, Gibson L, Hunter C (1999) ATPase activity associated witht the magnesium-protoporphyrin IX chelatase enzyme of *Synechocystis* PCC6803: evidence for ATP hydrolysis during Mg2+ insertion, and the MgATP-dependent interaction of the ChlI and ChlD subunits. Biochemical Journal 339: 127-134

Jia H-F, Chai Y-M, Li C-L, Lu D, Luo J-J, Qin L, Shen Y-Y (2011) Abscisic Acid Plays an Important Role in the Regulation of Strawberry Fruit Ripening. Plant Physiology 157: 188-199

Johansen L, Carrington J (2001) Silencing on the spot. Induction and suppression of RNA silencing in the *Agrobacterium*-mediated transient expression system. Plant Physiology 126: 930-938

Joyard J, Ferro M, Masselon C, Seigneurin-Berny D, Salvi D, Garin J, Rolland N (2009) Chloroplast Proteomics and the Compartmentation of Plastidial Isoprenoid Biosynthetic Pathways. Molecular Plant 2: 1154-1180

Jung H, Chory J (2010) Signaling between Chloroplasts and the Nucleus: Can a Systems Biology Approach Bring Clarity to a Complex and Highly Regulated Pathway? Plant Physiology 152: 453-459

Kajiwara S, Fraser P, Kondo K, Misawa N (1997) Expression of an exogenous isopentenyl diphosphate isomerase gene enhances isoprenoid biosynthesis in *Escherichia coli*. Biochem. J. 324: 421-426

Kapila J, de Rycke R, van Montagu M, Angenon G (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. Plant Science 122: 101-108

Karpinski S, Escobar C, Karpinska B, Creissen G, Mullineaux P (1997) Photosynthetic electron transport regulates the expression of cytosolic ascorbate peroxidase genes in *Arabidopsis* during excess light stress. Plant Cell 9: 627-640

Keegstra K, Cline K (1999) Protein import and routing systems of chloroplasts. Plant Cell 11: 557-570

Kessler F, Blobel G, Patel H, Schnell D (1994) Identification of two GTP-binding proteins in the chloroplast protein import machinery. Science 266: 1035-1039

Kessler F, Schnell D (2006) The function and diversity of plastid protein import pathways: a multilane GTPase highway into plastids. Traffic 7: 248-257

Kobayashi K, Mochizuki N, Yoshimura N, Motohashi K, Hisabori T, Masuda T (2008) Functional analysis of Arabidopsis thaliana isoforms of the Mg-chelatase CHLI subunit. Photochemical & Photobiological Sciences 7: 1188-1195

Kouranov A, Chen X, Fuks B, Schnell D (1998) Tic20 and Tic22 are new components of the protein import apparatus at the chloroplast inner envelope membrane. Journal of Cell Biology 143: 991-1002

Kouranov A, Schnell D (1997) Analysis of the interactions of preproteins with the import machinery over the course of protein import into chloroplasts. Journal of Cell Biology 139: 1677-1685

Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Lim J, Mittler R, Chory J (2007a) Multiple signals from damaged chloroplasts converge on a common pathway to regulate nuclear gene expression.". Science 316: 715-719

Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Lim J, Mittler R, Chory J (2007b) Signals from Chloroplasts Converge to Regulate Nuclear Gene Expression. Science 316: 715-719

Kovacs-Bogdan E, Soll J, Bolter B (2010) Protein import into chloroplasts: The Tic complex and its regulation. Biochimica et Biophysica Acta 1803: 740-747

Kubis S, Baldwin A, Patel R, Razzaq A, Dupree P, Lilley K, Kurth J, Leister D, Jarvis P (2003) The *Arabidopsis* ppi1 mutant is specifically defective in the expression, chloroplast import, and accumulation of photosynthetic proteins. Plant Cell 15: 1859-1871

Kumar A, Chaturvedi S, Soll D (1999) Selective inhibition of *HEMA* gene expression by photooxidation in *Arabidopsis thaliana*. Phytochemistry 51: 847-851

Kumar A, Csankovszki G, Soll D (1996) A second and differentially expressed glutamyl-tRNA reductase gene from *Arabidopsis thaliana*. Plant Molecular Biology 30: 419-426

Labbe-Bois R, Labbe P (1990) Tetrapyrrole and Heme biosynthesis in the yeast *Saccharomyces cerevisiae*. In: Dailey H (ed) Biosynthesis of Heme and Chlorophylls, New York, pp 235-285

Lake V, Olsson U, Willows R, Hansson M (2004) ATPase activity of magnesium chelatase subunit I is required to maintain subunit D in vivo. European Journal of Biochemistry 271: 2182-2188

Larkin M, Blackshields G, Brown N, Chenna R, McGettigan P, McWilliam H, Valentin F, Wallace I, Wilm A, Lopez R, Thompson J, Gibson T, Higgins D (2007) ClustalW and ClustalX version 2. Bioinformatics 23: 2947-2948

Larkin RM, Alonso J, M, Ecker JR, Chory J (2003) GUN4, a Regulator of Chlorophyll Synthesis and Intracellular Signaling. Science 299: 902-906

Lauring B, Kreibich G, Wiedmann M (1995) The intrinsic ability of ribosomes to bind to endoplasmic reticulum membranes is regulated by signal recognition particle and nascent-polypeptide-associated complex. Proceedings of the National Academy of Sciences 92: 9435-9439

Lee D, S L, Oh Y, Hwang I (2009) Multiple sequence motifs in the Rubisco small subunit transit peptide independently contribute to Toc159-dependent import of proteins into chloroplasts. Plant Physiology 151: 129-141

Lee J, Bankston L, Arnaout M, Liddington R (1995) Two conformations of the integrin A-domain (I-domain): a pathway for activation? Structure 3: 1333-1340

Lee K, Kim C, Lee D, Apel K (2003) *TIGRINA d*, required for regulating the biosynthesis of tetrapyrroles in barley, is an ortholog of the *FLU* gene of *Arabidopsis thaliana*. FEBS

Legnaioli T, Cuevas J, Mas P (2009) TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. The EMBO Journal 28: 3745-3757

Leister D (2003) Chloroplast research in the genomic age. Trends in Genetics 19: 47-56

Lermontova I, Kruse E, Mock H, Grimm B (1997) Cloning and characterization of a plastidal and a mitochondrial isoform of tobacco protoporphyrinogen IX oxidase. Proceedings of the National Academy of Sciences 94: 8895-8900

Li X, Henry R, Yuan J, Cline K, Hoffman N (1995) A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes. Proceedings of the National Academy of Sciences 92: 3789-3793

Liu Z, Butow R (2006) Mitochondrial retrograde signaling. Annual Review of Genetics 40: 159-185

Lundqvist J, Elmlund H, Peterson Wulff R, Berglund L, Elmlund D, Emanuelsson C, Hebert H, Willows RD, Hansson M, Lindahl M, Al-Karadaghi S (2010) ATP-Induced Conformational Dynamics in the AAA+ Motor Unit of Magnesium Chelatase. Structure 18: 354-365

MacArthur M, Thornton J (1991) Influence of proline residues on protein conformation. Journal of Molecular Biology 218: 397-412

Mano S, Hayashi M, Nishimura M (1999) Light regulates alternative splicing of hydroxypyruvate reductase in pumpkin. Plant Journal 17: 309-320

Martin T, Sharma R, Sippel C, Waegemann K, Soll J, Vothknecht U (2006) A protein kinase family in *Arabidopsis* phosphorylates chloroplast precursor proteins. Journal of Biological Chemistry 281

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, Leister D, Stoebe B, Hasegawa M, Penny D (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. Proceedings of the National Academy of Sciences 99: 12246-12251

Masoud S, Johnson L, White F (1992) The sequence within two primers influences the optimum concentration of dimethyl sulfoxide in the PCR. PCR Methods and Applications 2: 89-90

Mateo A, Muhlenbock P, Rusterucci C, Chang C, Miszalski Z, Karpinska B, Parker J, Mullineaux P, Karpinski S (2004) *LESION SIMULATING DISEASE 1* is required for acclimation to conditions that promote excess excitation energy. Plant Physiology 135: 2818-2830

May T, Soll J (2000) 14-3-3 proteins form a guidance complex with chloroplast precursor proteins in plants. Plant Cell 12: 53-64

McCormac A, Fischer A, Kumar A, Soll D, Terry M (2001) Regulation of *HEMA1* expression by phytochrome and a plastid signal during de-etiolation in *Arabidopsis thaliana*. Plant Journal 25: 549-561

McCormac A, Terry M (2002) Light-signalling pathways leading to the co-ordinated expression of *HEMA1* and *Lhcb* during chloroplast development in *Arabidopsis thaliana*. Plant Journal 32

Meskauskiene R, Apel K (2002) Interaction of FLU, a negative regulator of tetrapyrrole biosynthesis, with the glutamyl-tRNA reductase requires the tetratricopeptide repeat domain of FLU. FEBS 532: 27-30

Meskauskiene R, Nater M, Goslings D, Kessler F, op den Camp R, Apel K (2001) FLU: A negative regulator of chlorophyll biosynthesis in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences 98: 12826-12831

Michishita M, Videm V, Arnaout M (1993) A novel divalent cation-binding site in the A domain of the beta 2 integrin CR3 (CD11b/CD18) is essential for ligand binding. Cell 72: 857-867

Milborrow B (2001) The pathway of biosynthesis of abscisic acid in vascular plants: A review of the present state of knowledge of ABA biosynthesis. Journal of Experimental Botany 52: 1145-1164

Miller L (2007) Quantifying western blots without expensive commercial quantification software.

Mittler R, Vanderauwera S, Gollery M, Van Breusegem F (2004) Reactive oxygen gene network of plants. Trends in Plant Science 9: 490-498

Mochizuki N, Brusslan JA, Larkin R, Nagatani A, Chory J (2001) Arabidopsis genomes uncoupled 5 (GUN5) mutant reveals the involvement of Mg-chelatase H subunit in plastid-to-nucleus signal transduction. Proceedings of the National Academy of Sciences 98: 2053-2058

Mochizuki N, Tanaka R, Tanaka A, Masuda T, Nagatani A (2008) The steady-state level of Mg-protoporphyrin IX is not a determinant of plastid-to-nucleus signaling in Arabidopsis. Proceedings of the National Academy of Sciences 105: 15184-15189

Mock H, Keetman U, Kruse E, Rank B, Grimm B (1998) Defense Responses to Tetrapyrrole-Induced Oxidative Stress in Transgenic Plants with Reduced Uroporphyrinogen Decarboxylase or Coproporphyrinogen Oxidase Activity. Plant Physiology 116: 107-116

Montgomery B, Yeh K, Crepeau M, Lagarias J (1999) Modification of distinct aspects of photomorphogenesis via targeted expression of mammalian biliverdin reductase in transgenic Arabidopsis plants. Plant Physiology 121: 629-639

Mosinger E, Batschauer A, Schafer E, Apel K (1985) Phytochrome control of *in vitro* transcription of specific genes in Isolated nuclei from barley (*Hordeum vulgare*). European Journal of Biochemistry 147: 137-142

Moulin M, C. MA, Terry MJ, Smith AG (2008) Tetrapyrrole profiling in Arabidopsis seedlings reveals that retrograde plastid nuclear signaling is not due to Mg-protoporphyrin IX accumulation. Proceedings of the National Academy of Sciences 105: 15178-15183

Muller A, Hansson M (2009) The barley magnesium chelatase 150-kD subunit is not an abscisic acid receptor. Plant Physiology 150: 157-166

Muramoto T, Kohchi T, Yokota A, Hwang I, Goodman H (1999) The Arabidopsis Photomorphogenic Mutant *hy1* Is Deficient in Phytochrome Chromophore Biosynthesis as a Result of a Mutation in a Plastid Heme Oxygenase. The Plant Cell 11: 335-347

Murchie EH, Niyogi KK (2010) Manipulation of photoprotection to improve plant photosynthesis. Plant Physiology

Nakamura A, Shimada H, Masuda T, Ohta H, Takamiya K-i (2001) Two distinct isopentenyl diphosphate isomerases in cytosol and plastid are di¡erentially induced by environmental stresses in tobacco. FEBS letters 506: 61-64

Nakayama M, Masuda T, Sato N, Yamagata H, Bowler C, Ohta H, Shioi Y, Takamiya K (1995) Cloning, subcellular localization and expression of CHL1, a subunit of magnesium-chelatase in soybean. Biochemical and Biophysical Research Communications 215: 422-428

Natale P, Bruser T, Driessen A (2008) Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. Biochimica et Biophysica Acta 1778: 1735-1756

Nefissi R, Natsui Y, Miyata K, Oda A, Hase Y, Nakagawa M, Ghorbel A, Mizoguchi T (2011) Double loss-of-function mutation in EARLY FLOWERING 3 and CRYPTOCHROME 2 genes delays flowering under continuous light but accelerates it under long days and short days: an important role for *Arabidopsis CRY2* to accelerate flowering time in continuous light. Journal of Experimental Botany 62: 2731-2744

Nielsen H, Engelbrecht J, Brunak S, vonHeijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Engineering 10: 1-6

Nielsen O (1974) Photoconversion and Regeneration of Active Protochlorophyll(ide) in Mutants Defective in the Regulation of Chlorophyll Synthesis. Archives of Biochemistry and Biophysics 160: 430-439

Nott A, Jung H-S, Koussevitzky S, Chory J (2006) Plastid-to-Nucleus Retrograde Signaling. Annual Review of Plant Biology 57: 739-759

Nover L, Bharti K, Doring P, Mishra S, Ganguli A, Scharf K (2001) Arabidopsis and the heat stress transcription factor world: how many heat stress transcription factors do we need? Cell Stress Chaperones 6: 177-189

Oelmuller R, Levitan I, Bergfeld R, Rajasekhar VK, Mohr H (1986) Expression of nuclear genes as affected by treatments acting on the plastids. Planta 168: 482-492

Oelmuller R, Mohr H (1986) Photooxidative destruction of chloroplasts and its consequences for expression of nuclear genes. Planta 167: 106-113

Oreb M, Hofle A, Mirus O, Schleiff E (2008) Phosphorylation regulates the assembly of chloroplast import machinery. Journal of Experimental Botany 59: 2309-2316

Osanai T, Imashimizu M, Seki A, Sato S, Tabata S, Imamura S, Asayama M, Ikeuchi M, Tanaka K (2009) ChlH, the H subunit of the Mg-chelatase, is an anti-sigma factor for SigE in Synechocystis sp. PCC 6803. Proceedings of the National Academy of Sciences 106: 6860-6865

Osterlund MT, Hardtke CS, Wei N, Deng XW (2000) Targeted destabilization of HY5 during light-regulated development of Arabidopsis. Nature 405: 462-466

Pain D, Blobel G (1987) Protein import into chloroplast requires a chloroplast ATPase. Proceedings of the National Academy of Sciences 84: 3288-3292

Papenbrock J, Grafe S, Kruse E, Hanel F, Grimm B (1997) Mg-chelatase of tobacco: identification of a ChlD cDNA sequence encoding a third subunit, analysis of the interaction of the three subunits with the yeast two-hybrid system, and reconstitution of the enzyme activity by co-expression of recombinant CHL D, CHL H and CHL I. The Plant Journal 12: 981-990

Papenbrock J, Mock H-P, Kruse E, Grimm B (1999) Expression studies in tetrapyrrole biosynthesis: inverse maxima of magnesium chelatase and ferrochelatase activity during cyclic photoperiods. Planta 208: 264-273

Papenbrock J, Pfundel E, Mock H, Grimm B (2001) Decreased and increased expression of the subunit CHL I diminishes Mg chelatase activity and reduces chlorophyll synthesis in transgenic tobacco plants. The Plant Journal 22: 155-164

Pardo A, Chereskin B, Castelfranco P, Franceschi V, Wezelman B (1980) ATP requirement for Mg-chelatase in developing chloroplasts. Plant Physiology 65: 956-960

Parks B, Quail B (1991) Phytochrome-Deficient hy1 and hy2 Long Hypocotyl Mutants of Arabidopsis Are Defective in Phytochrome Chromophore Biosynthesis. The Plant Cell 3: 1177-1186

Perry S, Keegstra K (1994) Envelope membrane proteins that interact with chloroplastic precursor proteins. Plant Cell 6: 93-105

Peter E, Grimm B (2009) GUN4 is Required for Posttranslational Control of Plant Tetrapyrrole Biosynthesis. Molecular Plant 2: 1198-1210

Peteranderl R, Rabenstein M, Shin Y, Liu C, Wemmer D, King D, Nelson H (1999) Biochemical and biopysical characterization of the trimerization domain from the heat shock transcriptions factor. Biochemistry 38: 3559-3569

Petersen B, Kannangara C, Henningsen K (1999) Distribution of ATPase and ATP-to-ADP phosphate exchange activities in magnesium chelatase subunits of *Chlorobium vibrioforme* and *Synechocystis* PCC6803. Archives of Microbiology 171: 146-150

Pineau B, Gerard-Hirne C, Douce R, Joyard J (1993) Identification of the Main Species of Tetapyrrolic Pigments in Envelope Membranes from Spinach Chloroplasts. Plant Physiology 102: 821-828

Pizschke A, Hirt H (2010) New insights into an old story: *Agrobacterium*-induced tumor formation in plants by plant transformation. EMBO Journal 29: 1021-1032

Pomp D, Medrano J (1991) Organic Solvents as Facilitators of Polymerase Chain Reaction. BioTechniques 10: 58-59

Pontoppidan B, Kannangara C (1994) Purification and partial characterisation of barley glutamyl-tRNA(Glu) reductase, the enzyme that directs glutamate to chlorophyll biosynthesis. European Journal of Biochemistry 225: 529-537

Przybyla D, Gobel C, Imboden A, Hamberg M, Feussner I, Apel K (2008) Enzymatic, but not non-enzymatic, $^1O_2$-mediated peroxidation of polyunsaturated fatty acids forms part of the EXECUTER1-dependent stress response program in the *flu* mutant of *Arabidopsis thaliana*. Plant Journal 54: 236-248

Qbadou S, Becker T, Bionda T, Reger K, Ruprecht M, Soll J, Schleiff E (2007) Toc64 - A Preprotein-receptor at the outer membrane with bipartide function. Journal of Molecular Biology 367: 1330-1346

Qbadou S, Becker T, Mirus O, Tews I, Soll J, Schleiff E (2006) The molecular chaperone Hsp90 delivers precursor proteins to the chloroplast import receptor Toc64. The EMBO Journal 25

Qian P, Marklew C, Viney J, Davison P, Brindley A, Soderberg C, Al-Karadaghi S, Bullough P, Grossmann J, Hunter C (2012) Structure of the cyanobacterial Magnesium Chelatase H subunit determined by single particle reconstruction and small-angle X-ray scattering. Journal of Biological Chemistry 287: 4946-4956

Reid J, Hunter C (2002) Current understanding of the function of magnesium chelatse. Biochem Soc Trans 30: 643-645

Reid J, Hunter C (2004) Magnesim-dependent ATPase Activity and Cooperativity of Magnesium Chelatase from *Synechocystis* sp. PCC6803. The Journal of Biological Chemistry 279: 26893-26899

Reid J, Siebert C, Bullough P, Hunter C (2003) The ATPase activity of the ChlI subunit of magnesium chelatase and formation of a heptameric AAA+ ring. Biochemistry 42: 6912-6920

Reinbothe S, Reinbothe C, Lebedev N, Apel K (1996) PORA and PORB, Two Light-Dependent Protochlorophyllide-Reducing Enzymes of Angiosperm Chlorophyll Biosynthesis. The Plant Cell 8: 763-769

Richly E, Leister D (2004) An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of *Arabidopsis* and rice. Gene 329: 11-16

Robinson C, Bolhuis A (2004) Tat-dependent protein targeting in prokaryotes and chloroplasts. Biochimica et Biophysica Acta 1694: 135-147

Ron D, Walter P (2007) Signal integration in the endoplasmic reticulum unfolded protein response. Nature Reviews Molecular Cell Biology 8: 519-529

Rook F, Hadingham S, Li Y, Bevan M (2006) Sugar and ABA response pathways and the control of gene expression. Plant Cell Environment 29: 425-434

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology 132: 365-386

Ruckle ME, DeMarco SM, Larkin RM (2007) Plastid signals remodel light signaling networks and are essential for efficient chloroplast biogenesis in Arabidopsis. Plant Cell 19: 3944-3960

Ruprecht M, Bionda T, Sato T, Sommer MS, Endo T, Schleiff E (2010) On the Impact of Precursor Unfolding during Protein Import into Chloroplasts. Molecular Plant 3: 499-508

Sagar A, Horwitz B, Elliott R, Thompson W, Briggs W (1988) Light effects on several chloroplast components in Norflurazon-treated pea seedlings. Plant Physiology 88: 340-347

Salomon M, Fischer K, Flugge U, Soll J (1990) Sequence analysis and protein import studies of an outer chloroplast envelope polypeptide. Proceedings of the National Academy of Sciences 87: 5778-5782

Sargent F, Berks B, Palmer T (2006) Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins. Federation of European Microbiological Societies 254: 198-207

Sattler S, Mene-Saffrane L, Farmer E, Krischke M, Mueller M, DellaPenna D (2006) Nonenzymatic lipid peroxidation reprograms gene expression and activates defense markers in *Arabidopsis* tocopherol-deficient mutants. Plant Cell 18: 3706-3720

Sawers RJ, Farmer PR, Moffett P, Brutnell T (2006a) In planta transient expression as a system for genetic and biochemical analyses of chlorophyll biosynthesis. Plant Methods 2

Sawers RJ, Viney J, Farmer PR, Bussey RR, Olsefski G, Anufrikova K, Hunter CN, Brutnell T (2006b) The maize *Oil Yellow1 (Oy1)* gene encodes the I subunit of magnesium chelatase. Plant Molecular Biology 60: 95-106

Sawicki A, Willows R (2008) Kinetic Analyses of the Magnesium Chelatase Provide Insights into the Mechanism, Structure, and Formation of the Complex. The Journal of Biological Chemistry 283: 31294-31302

Sawicki A, Willows RD (2007) S-Adenosyl-L-methionine:magnesium-protoporphyrin IX O-methyltransferase from Rhodobacter capsulatus: mechanistic insights and stimulation with phospholipids. The Biochemical Journal 406: 469-478

Scharf K, Heider H, Hohfeld I, Lyck R, Schmidt E, Nover L (1998) The Tomato Hsf System: HsfA2 Needs Interaction with HsfA1 for Efficient Nuclear Import and May Be Loacalized in Cytoplasmic Heat Stress Granules. Molecular and Cellular Biology 18: 2240-2251

Schippers J, Nguyen H, Lu D, Schmidt R, Mueller-Roeber B (2012) ROS homeostasis during development: an evolutionary conserved strategy. Cellular and Molecular Life Sciences 26: 3245-3257

Schleiff E, Jelic M, Soll J (2003) A GTP-driven motor moves proteins across the outer envelope of chloroplasts. Proceedings of the National Academy of Sciences 100: 4604-4609

Schleiff E, Soll J, Kuchler M, Kuhlbrandt W, Harrer R (2003b) Characterization of the translocon of the outer envelope of chloroplasts. Journal of Cell Biology 160: 541-551

Schleiff E, Soll J, Sveshnikova N, Tien R, Wright S, Dabney-Smith C, Subramanian C, Bruce B (2002) Structural and guanosine triphosphate/diphosphate requirements for transit peptide recognition by the cytosolic domain of the chloroplast outer envelope receptor, Toc34. Biochemistry 41: 1934-1946

Schnell D, Blobel G (1993) Identification of intermediates in the pathway of protein import into chloroplasts and their localization to envelope contact sites. Journal of Cell Biology 120: 103-115

Schnell D, Kessler F, Blobel G (1994) Isolation of components of the chloroplast protein import machinery. Science 266: 1007-1012

Schon A, Krupp G, Gough S, Berry-Lowe S, Kannangara C, Soll D (1986) The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. Nature 322: 281-284

Schunemann D (2007) Mechanisms of protein import into thylakoids of chloroplasts. Biological Chemistry 388: 907-915

Seedorf M, Waegemann K, Soll J (1995) A constituent of the chloroplast import complex represents a new type of GTP-binding protein. Plant Journal 7: 401-411

Seo M, Koshiba T (2002) Complex regulation of ABA biosynthesis in plants. Trends in Plant Science 7: 41-48

Shang Y, Yan L, Liu Z, Cao Z, Mei C, Xin Q, Wu F, Wang X, Du S, Jiang T, Zhang X, Zhao R, Sun H, Liu R, Yu Y, Zhang D (2010) The Mg-Chelatase H Subunit of *Arabidopsis* Antagonizes a Group of WRKY Transcription Repressors to Relieve ABA-Responsive Genes of Inhibition. The Plant Cell 22: 1909-1935

Shao F, Golstein C, Ade J, Stoutemyer M, Dixon J, Innes R (2003) Cleavage of *Arabidopsis* PBS1 by Bacterial Type III Effector. Science 301: 1230-1233

Shen Y, Wang X, Wu F, Du S, Cao Z, Shang Y, Wang X, Peng C, Yu X, Zhu S, Fan R, Xu Y, Zhang D (2006a) The Mg-chelatase H subunit is an abscisic acid receptor. Nature 443: 823-826

Shen Y-Y, Wang X-F, Wu F-Q, Du S-Y, Cao Z, Shang Y, Wang X-L, Peng C-C, Yu X-C, Zhu S-Y, Fan R-C, Xu Y-H, Zhang D-P (2006b) The Mg-chelatase H subunit is an abscisic acid receptor. Nature 443: 823-826

Shih M, Heinrich P, Goodman H (1991) Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate-dehydrogenase from *Arabidopsis thaliana*. Gene 104: 133-138

Shilov I, Seymour S, Patel A, Loboda A, Tang W, Keating S, Hunter C, Nuwaysir L, Schaeffer D (2007) The Paragon Algorithm: A next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. Molecular and Cellular Proteomics 6: 1638-1655

Shkolnik-Inbar D, Bar-Zvi D (2010) *ABI4* Mediates Abscisic Acid and Cytokinin Inhibition of Lateral Root Formation by Reducing Polar Auxin Transport in *Arabidopsis*. The Plant Cell 22: 3560-3573

Sirijovski N, Lundqvist J, Rosenback M, Elmlund H, Al-Karadaghi S, Willows RD, Hansson M (2008) Substrate-binding Model of the Chlorophyll Biosynthetic Magnesium Chelatase BchH Subunit. The Journal of Biological Chemistry 283: 11652-11660

Smeekens S, Bauerle C, Hageman J, Keegstra K, Weisbeek P (1986) The role of the transit peptide in the routing of precursors toward different chloroplast compartments. Cell 46: 365-375

Smith E, Townsend C (1907) A plant-tumor of bacterial origin. Science 25: 671-673

Smith M (2006) Protein import into chloroplasts: an ever-evolving story. Canadian Journal of Botany 84: 531-542

Smith-Johannsen H, Gibbs S (1972) Effects of chloramphenicol on chloroplast and mitochondrial ultrastructure in Ochromonas danica. Journal of Cell Biology 52: 598-614

Sobotka R, Duhring U, Komenda J, Peter E, Gardian Z, Tichy M, Grimm B, Wilde A (2008) IMPORTANCE OF THE CYANOBACTERIAL GUN4 PROTEIN FOR CHLOROPHYLL METABOLISM AND ASSEMBLY OF PHOTOSYNTHETIC COMPLEXES. The Journal of Biological Chemistry 283: 25794-25802

Soll J, Schleiff E (2004) Protein import into chloroplasts. Nature Reviews Molecular Cell Biology 5: 198-208

Song X, Bandow J, Sherman J, Baker J, Brown P, McDowell M, Molloy M (2008) iTRAQ Experimental Design for Plasma Biomarker Discovery. Journal of Proteome Research 7: 2952-2958

Stahl T, Glockmann C, Soll J, Heins L (1999) Tic40, a new 'old' subunit of the chloroplast protein import translocon. Journal of Biological Chemistry 274: 37467-37472

Strand A, Asami T, Alonso J, Ecker J, Chory J (2003) Chloroplast to nucleus communication triggered by accumulation of Mg-protoporphyrinIX. Nature 421: 79-83

Strittmatter P, Soll J, Bolter B (2010) The Chloroplast Protein Import Machinery: A Review. In: Economou A (ed) Protein Secretion: Methods and Protocols. Springer-Verlag, New York, pp 307-321

Sullivan J, Gray J (1999) Plastid translation is required for the expression of nuclear photosynthesis genes in the dark and in roots of the pea *lip*1 mutant. Plant Cell 11: 901-910

Sun X, Feng P, Xu X, Guo H, Ma J, Chi W, Lin R, Lu C, Zhang L (2011) A chloroplast envelope-bound PHD transcription factor mediates chloroplast signals to the nucleus. Nature Communications 2: 477

Sun Z, Cunningham FJ, Gantt E (1998) Differential expression of two isopentenyl pyrophosphate isomerases and enhanced carotenoid accumulation in a unicelular chlorophyte. Proceedings of the National Academy of Sciences 95: 11482-11488

Sun Z, Gantt E, Cunningham FJ (1996) Cloning and functional analysis of the beta-carotene hydroxylase of *Arabidopsis thaliana* Journal of Biological Chemistry 271: 24349-24352

Susek RE, Ausubel FM, Chory J (1993) Signal Transduction Mutants of Arabidopsis Uncouple Nuclear CAB and RBCS Gene Expression from Chloroplast Development. Cell 74: 787-799

Takagi J (2007) Structural basis for ligand recognition by integrins. Current Opinion in Cell Biology 19: 557-564

Tanaka R, Tanaka A (2007) Tetrapyrrole biosynthesis in higher plants. Annual Review of Plant Biology 58: 321-346

Tanaka R, Yoshida K, Nakayashiki T, Masuda T, Tsuji H, Inokuchi H, Tanaka A (1996) Differential expression of two *hemA* mRNAs encoding glutamy-tRNA reductase proteins in greening cucumber seedlings. Plant Physiology 110: 1223-1230

Tanaka R, Yoshida K, Nakayashiki T, Tsuji H, Inokuchi H, Okada K, Tanaka A (1997) The third member of the *hemA* gene family encoding glutamyl-tRNA

reductase is primarily expressed in roots in *Hordeum vulgare*. Photosynthesis Research 53: 161-171

Teng Y, Su Y, Chen L, Lee Y, Hwang I, Li H (2006) Tic21 is an essential translocon component for protein translocation across the chloroplast inner envelope membrane. Plant Cell 18: 2247-2257

Terry M, Kendrick R (1999) Feedback inhibition of chlorophyll synthesis in the phytochrome chromophore deficient *aurea* and *yellow-green-2* mutants of tomato. Plant Physiology 119: 143-152

Terry M, Wahleithner J, Lagarias J (1993) Biosynthesis of the plant photoreceptor phytochrome. Archives of Biochemistry and Biophysics 306: 1-15

Theg S, Bauerle C, Olsen L, Selman B, Keegstra K (1989) Internal ATP is the only energy requirement for the translocation of precursor proteins across chloroplastic membranes. Journal of Biological Chemistry 264: 6730-6736

Thomas J, Weinstein J (1992) Free heme in isolated chloroplasts - an improved method of assay and its physiological importance. Plant Physiology and Biochemistry 30: 285-292

Thompson SJ, Robinson C, Mant A (1999) Dual Signal Peptides Mediate the Signal Recognition Particle/Sec-independent Insertion of a Thylakoid Membrane Polyprotein, PsbY. The Journal of Biological Chemistry 274: 4059-4066

Thompson W, White M (1991) Physiological and molecular studies of light-regulated nuclear genes in higher plants. Annual Review of Plant Physiology and Plant Molecular Biology 42: 423-466

Timmis J, Ayliffe M, Huang C, Martin W (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Natrure Reviews Genetics 5: 123-135

Triantaphylides C, Krischke M, Hoeberichts FA, Ksas B, Gresser G, Havaux M, Van Breusegem F, Mueller MJ (2008) Singlet oxygen is the major reactive oxygen species involved in photooxidative damage to plants. Plant Physiology 148: 960-968

Tsien R (1998) The Green Fluorescent Protein. The Annual Review of Biochemistry 67: 509-544

Tsuzuki T, Takahashi K, Inoue S, Okigaki Y, Tomiyama M, Hossain M, Shimazaki K, Murata Y, Kinoshita T (2011) Mg-chelatase H subunit affects ABA signaling in stomatal guard cells, but is not an ABA receptor in *Arabidopsis thaliana*. Journal of Plant Reseach 124: 527-538

Ulm R, Nagy F (2005) Signalling and gene regulation in response to ultraviolet light. Current Opinion in Plant Biology 8: 477-482

Vale R (2000) AAA proteins. Lords of the ring. Journal of Cell Biology 150: F13-F19

van den Wijngaard P, Vredenberg W (1999) The envelope anion channel involved in chloroplast protein import is associated with Tic110. Journal of Biological Chemistry 274: 25201-25204

Van der Hoorn J, Laurent F, Roth R, De Wit P (2000) Agroinfiltration is a versatile tool that facilitates comparative analyses of *Avr9/cf-9*-induced and *Avr4/Cf-4*-induced necrosis. Molecular Plant-Microbe Interactions 13: 439-446

van Grinsven M, Kool A (1988) Plastid gene regulation during development: An intriguing complexity of mechanisms. Plant Molecular Biology Reporter 6: 213-239

Vaquero C, Sack M, Chandler J, Drossard J, Schuster F, Monecke M, Schillberg S, Fischer R (1999) Transient expression of a tumor-specific single-chain fragment and a chimeric antibody in tobacco leaves. Proceedings of the National Academy of Sciences 96: 11128-11133

Vaquero C, Sack M, Schuster F, Finnern R, Drossard J, Schumann A, Fischer R (2002) A carcinoembryonic antigen-specific diabody produced in tobacco. The FASEB Journal 16: 408-410

Viitanen PV, Doran ER, Dunsmuir P (1988) What is the Role of the Transit Peptide in Thylakoid Integration of the Light-harvesting Chlorphyll *a/b* Protein? The Journal of Biological Chemistry 263: 15000-15007

Voigt C, Oster U, Vornke F, Jahns P, Dietz K-J, Leister D, Kleine T (2009) In-depth analysis of the distinctive effects of norflurazon implies that tetrapyrrole biosynthesis, organellar gene expression and ABA cooperate in the GUN-type of plastid signalling. Physiologia Plantarum 138: 503-519

Voinnet O, Rivas S, Mestre P, Baulcombe D (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. The Plant Journal 33: 949-956

von Wettstein D, Gough S, Kannangara C (1995) Chlorophyll Biosynthesis. The Plant Cell 7: 1039-1057

Vothknecht U, Kannangara C, von Wettstein D (1998) Barley glutamyl tRNA Glu reductase: mutations affeting haem inhibition and enzyme activity. Phytochemistry 47: 513-519

Waegemann K, Soll J (1996) Phosphorylation of the transit sequence of chloroplast precursor proteins. Journal of Biological Chemistry 271

Wagner D, Przybyla D, op den Camp R, Kim C, Landgraf F, Lee K, Wursch M, Laloi C, Nater M, Hideg E, Apel K (2004) The genetic basis of singlet oxygen-induced stress responses of *Arabidopsis thaliana*. Science 306: 1183-1185

Walker C, Weinstein J (1994) The magnesium-insertion step of chlorophyll biosynthesis is a two-stage reaction. Biochemical Journal 299(Pt 1): 277-284

Walker C, Weinstein J (1995) Re-examination of the localization of Mg-chelatase within the chloroplast. Physiologia Plantarum 94: 419-424

Walker CJ, Weinstein JD (1991) Further Characterization of the Magnesium Chelatase in Isolated Developing Cucumber Chloroplasts. Plant Physiology 95: 1189-1196

Walker CJ, Willows RD (1997) Mechanism and regulation of Mg-chelatase. Biochem. J. 327: 321-333

Wang L, Elliott M, Elliott T (1999) Conditional stability of the HemA protein (glutamyl-tRNA reductase) regulates heme biosynthesis in Salmonella typhimurium. Journal of Bacteriology 181: 1211-1219

Wang W, Huang D, Stachon D, Gough S, Kannangara C (1984) Purification, characterization and fractionation of the ∂-aminolevulinic acid synthesizing enzymes from light-grown *Chlamydomonas reinhardtii* cells. Plant Physiology 74: 569-575

Wang X-F, Xin Q, Shen Y-Y, Zhang D-P (2011) Approaches to the Identification of ABAR as an Abscisic Acid Receptor. Methods in Molecular Biology 773: 83-97

Waters MT, Wang P, Korkaric M, Capper RG, Saunders NJ, Langdale JA (2009) GLK Transcription Factors Coordinate Expression of the Photosynthetic Apparatus in Arabidopsis. The Plant Cell 21: 1109-1128

Weinstein J, Beale S (1985) Enzymatic conversion of glutamate to delta-aminolevulinate in soluble extracts of the unicellular green alga, *Chlorella vulgaris*. Archives of Biochemistry and Biophysics 237: 454-464

Whittaker C, Hynes R (2002) Distribution and Evolution of von Willebrand/Integrin A domains: Widely Dispersed Domains with Roles in Cell Adhesion and Elsewhere. Molecular Biology of the Cell 13: 3369-3387

Wiedmann B, Sakai H, Davis T, Wiedmann M (1994) A protein complex required for signal-sequence -specific sorting and translocation. Nature 370: 434-440

Willows R (2003) Biosynthesis of chlorophylls from protoporphyrin IX. Nat Prod Rep 20: 327-341

Willows R, Beale S (1998) Heterologous expression of the Rhodobacter capsulatus BchI, -D, and -H genes that encode magnesium chelatase subunits and characterization of the reconstituted enzyme. Journal of Biological Chemistry 273: 32406-32413

Willows R, Gibson L, Kanangara C, Hunter C, von Wettstein D (1996) Three separate proteins constitute the magnesium chelatase of *Rhodobacter sphaeroides*. European Journal of Biochemistry 235: 438-443

Willows R, Hansson A, Birch D, Al-Karadaghi S, Hansson M (2004) EM single particle analysis of the ATP-dependent BchI complex of magnesium chelatase: an AAA+ hexamer. Journal of Structural Biology 146: 227-233

Willows R, Hansson M, Beale S, Laurberg M, Al-Karadaghi S (1999) Crystallization and preliminary X-ray analysis of the *Rhodobacter capsulatus* magnesium chelatase BchI subunit. Acta Crystallographica Section D: Biological Crystallography 55: 689-690

Wu F, Xin Q, Cao Z, Liu Z, Du S, Mei C, Zhao C, Wang X, Shang Y, Jiang T, Zhang X, Yan L, Zhao R, Cui Z, Liu R, Sun H, Yang X, Su Z, Zhang D (2009) The magnesim-chelatase H subunit binds abscisic acid and functions in abscisic acid signaling: new evidence in Arabidopsis. Plant Physiology 150: 1940-1954

Yang K, Kim H, Jin U, Lee S, Park J, Lim Y, Pai H (2007) Silencing of NbBTF3 results in develpmental defects and disturbed gene expression in chloroplasts and mitochondria of higher plants. Planta 225: 1459-1469

Yaronskaya E, Ziemann V, Walter G, Averina N, Borner T, Grimm B (2003) Metabolic control of the tetrapyrrole biosynthetic pathway for porphyrin distribution in the barley mutant *albostrians*. Plant Journal 35: 512-522

Young J, Agashe V, Siegers K, Hartl F (2004) Pathways of chaperone-mediated protein folding in the cytosol. Nature Reviews Molecular Cell Biology 5: 781-791

Zurawski G, Clegg MT (1987) Evolution of Higher-plant Chloroplast DNA-encoded Genes: Implications for Structure -Function and Phylogenetic Studies. Annual Review of Plant Physiology 38: 391-418

Zygadlo A, Robinson C, Scheller HV, Mant A, Jensen PE (2006) The Properties of the Positively Charged Loop Region in PSI-G are Essential for its "spontaneous" Insertion into Thylakoids and Rapid Assembly into Photosystem I Complex. The Journal of Biological Chemistry 281: 10548-10554