Analyses on the Formation of Regulatory Systems for Expression and Maturation of Photosynthetic Complexes in *Arabidopsis thaliana*

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Dagmar Anna Lyska

aus Gleiwitz

Düsseldorf, Mai 2011

aus dem Institut für Entwicklungs- und Molekularbiologie der Pflanzen
der Heinrich-Heine Universität Düsseldorf

Gedruckt mit der Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. P. Westhoff
Koreferent: Prof. Dr. P. Jahns

Tag der mündlichen Prüfung:

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 23.05.2011

Contents

1) Dagmar Lyska, Kerstin Schult, Karin Meierhoff and Peter Westhoff (2011). **pAUL: A Gateway-based vector system for adaptive expression and flexible tagging of proteins in *Arabidopsis*.** Submitted to Journal of Experimental Botany for publication.
2) Dagmar Lyska, Susanne Paradies, Karin Meierhoff and Peter Westhoff (2007). **HCF208, a homolog of *Chlamydomonas* CCB2, is required for accumulation of native cytochrome b_6 in *Arabidopsis thaliana*.** *Plant Cell Physiol,* **48**, 1737-1746.
3) **Molecular characterization of HCF208 localization and interactions.**
4) **Analyses on protein function and complex formation of the PsbH synthesis factor HCF107.**

I. General Introduction

I.1 Chloroplast structure and function

Chloroplasts are the characteristic organelles of plants and green algae. They are the sites of photosynthesis converting light to chemical energy and atmospheric CO_2 to carbohydrates. Furthermore, other essential processes like lipid metabolism and biosynthesis of amino acids, purine and pyrimidine bases, chlorophyll and other tetrapyrroles take place in chloroplasts (Neuhaus and Emes, 2000, Finkemeier and Leister, 2010). In seed plants, chloroplasts develop from non-photosynthetic proplastids, which are also the initial form of chromoplasts (pigment accumulation), amyloplasts (starch storage), and leucoplasts or elaioplasts (lipid storage). Differentiation of chloroplasts from proplastids in leaf cells is initiated by light. Absence or insufficiency of light leads to the development of unpigmented etioplasts, which eventually can convert to chloroplasts upon illumination (Waters and Langdale, 2009).

Chloroplasts consist of several compartments. Two layers of membranes, the outer and inner envelopes, delimit it from the cytosol of the surrounding cell. The membranes are the sites of lipid biosynthesis and exchange of molecules between the cytosol and the soluble compartment of chloroplasts referred to as stroma. The stroma harbors the transcription- and translation-machinery of the chloroplast, as well as most metabolic enzymes and enzymes of the “dark reactions” of photosynthesis, the Calvin cycle (Calvin, 1962; Wolusiuk et al, 1993). The stroma also harbors the complex thylakoid membrane, which surrounds the lumenal compartment. Thylakoid membranes are differentiated into two domains: stacked structures called grana lamellae and the interconnecting single membrane regions called stroma lamellae. Inside the thylakoid membranes the four large protein complexes photosystem II (PSII; Ferreira et al., 2004, Nelson and Yocum, 2006), the cytochrome b_6f complex (Kurisu et al., 2003; Stroebel et al., 2003), photosystem I (PSI; Nelson and Yocum, 2006), and ATP-synthase (Seelert et al., 2000) are embedded (Figure 1). They carry out the “light reactions” of photosynthesis, where light energy is fixed and converted to ATP and NADPH, which are forwarded to the Calvin cycle reactions. PSII is predominantly located in grana lamellae, whereas the majority of PSI and ATP-synthase can be found in stroma lamellae. The cytochrome b_6f complex is distributed equally between the two membrane types (Anderson, 2002; Kim et al., 2005). The heterologous distribution of complexes is tightly coupled to the photosynthetic events (Albertsson, 2001).

Linear electron transport through the thylakoid membrane begins with the excitation of a chlorophyll pair P680 at PSII generating $P680^+$ and leading to charge separation. Electrons are transferred to plastoquinone, which is reduced to form plastoquinol. $P680^+$ is de-excited

by electrons generated from water oxidation via the manganese (Mn_4Ca) cluster at the oxygen evolving complex (OEC) at the lumenal side of PSII. Electrons from plastoquinol are then forwarded to the cytochrome b_6f complex, which routes them to the soluble electron carrier protein plastocyanin. Subsequently, PSI, which is oxidized due to charge separation of its associated chlorophyll pair P700, adsorbs electrons from plastocyanin. Electrons released from PSI are passed to ferredoxin, which then reduces $NADP^+$ to NADPH with help of ferredoxin-NADPH oxidoreductase (FNR). The linear electron transport is coupled to the translocation of electrons from the stroma to the lumen generating the proton motive force driving ATP synthesis from ADP and inorganic phosphate by the ATP-sythase (Figure 1; Nelson and Ben-Shem, 2004).

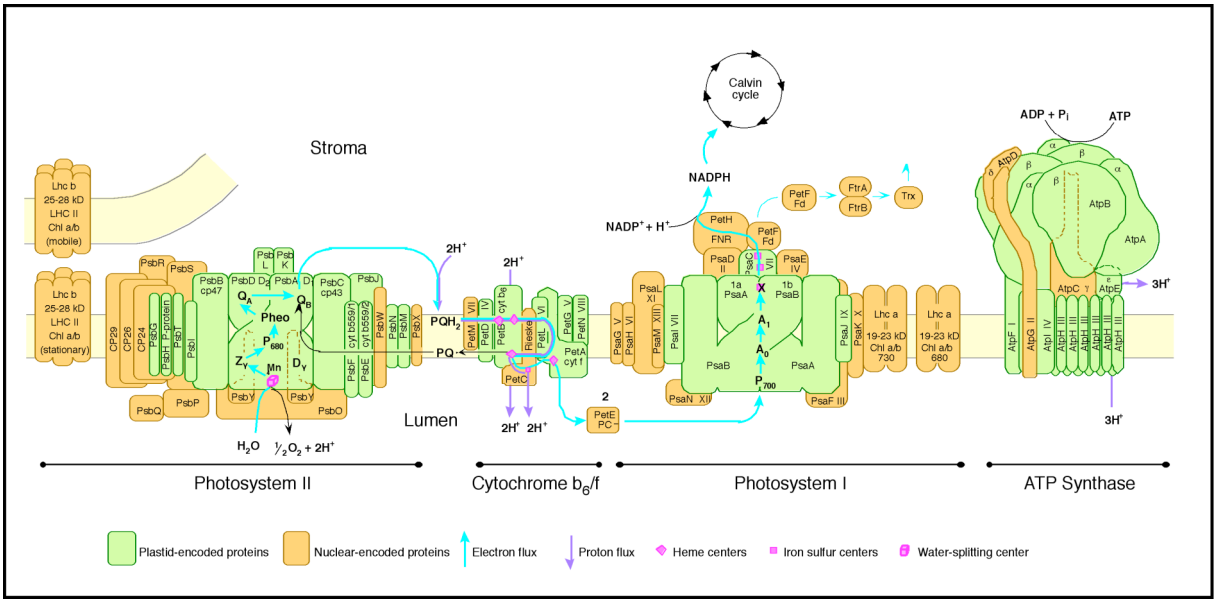


Figure 1: Scheme of the thylakoid membrane complexes of higher plant chloroplasts.
The linear electron transport flux is indicated by blue arrows. Contribution of the chloroplast and nucleus to subunit composition is demonstrated by different coloration of the subunits (Modified from Race et al., 1999).

An alternative pathway of electrons through the thylakoid membrane referred to as cyclic electron transport was described in 1954 by Arnon and coworkers. However, only in the past years genetic and biochemical methods have gained insight into the components involved in this alternative electron transport and its mechanism. Cyclic electron transport is supposed to be required for balancing NADPH and ATP ratio (Kramer et al., 2004) and photoprotection of the photosynthetic apparatus (Shikanai, 2007) and solely depends on PSI reactions (Munekage and Shikanai, 2005). Electrons generated on the reducing side of PSI are re-injected to the plastoquinone pool, thus generating an additional proton motive force and boosting ATP synthesis. Recently, a supercomplex that drives cyclic electron transport has been purified from the green alga *Chlamydomonas reinhardtii* (Iwai et al., 2010). It includes PSI with light harvesting complexes (LHC) I and II, the cytochrome b_6f complex, ferredoxin-

NADPH oxidoreductase (FNR), and the integral membrane protein PGRL1 (DalCorso et al., 2008).

I.2 Plastid evolution

The organelles of eukaryotic cells, namely mitochondria and in terms of plants also plastids, descended from bacterial ancestors (Dyall et al., 2004). It is suggested that mitochondria evolved from an α proteobacterium-like ancestor taken up by an archae-type host >1.5 billion years ago (Gray et al., 1999). Phylogenetic, structural, and biochemical analyses show that primary plastids developed 1.2 to 1.5 billion years ago from a cyanobacterium-like ancestor engulfed by a mitochondrium-possessing eucaryot (Martin and Russell, 2003). Three major autotrophic lineages evolved from primary endosymbiosis: glaucophytes, red algae, and green algae, which are the ancestors of higher plants (Hjorth et al., 2004). Plastids from primary endosymbionts were transferred laterally to other heterotrophic eukaryotes, a process termed secondary endosymbiosis (Cavalier-Smith, 1982). Lineages derived from secondary endosymbiosis include euglenophytes, chlorarachniophytes, and chromalveolates (Gould et al., 2008). Comparison of present cyanobacterial genomes with plastid genomes (plastomes) points to massive loss of genetic information in consequence of endosymbiotic events. Whereas cyanobacteria like *Anabena sp*. PCC 7120 and *Synechocistis sp*. PCC3168 have 5366 and 3268 protein-encoding genes, respectively (Kaneko et al., 1996; Kaneko et al., 2001), plastomes encode for significantly less proteins (*Arabidopsis thaliana*: 87 (Sato et al., 1999), *C. reinhardtii*: 99 (Maul et al., 2002)). However, the total number of plastid proteins is estimated to be between 2000 and 3600 (Leister, 2003; Richly and Leister, 2004) and thus overall correlates to the protein content of cyanobacteria. These numbers imply that the majority of the genes were transferred to the nucleus or lost in the course of co-evolution of the symbiont and its host cell. Genes that “remained” in the plastid genome mainly encode for its transcription- and translation-machinery, as well as for subunits of the thylakoid membrane complexes (Figure 1) and the large subunit of ribulose-1,5-bisphosphate carboxylase/ oxygenase (Martin et al, 2002; Timmis et al., 2004). Maintenance of certain genes in the plastome is supposed to be due to two reasons: (i) proteins involved in photosynthesis cannot be transported into the plastid since they are too hydrophobic or toxic if allowed to accumulate in the cytoplasm (Bogorad, 1975; Allen, 2003) and/ or (ii) rapid regulatory control of organellar proteins involved in redox reactions by the redox state of the organelle can only be achieved if the proteins are encoded by the organelle (Allen, 2003).
Plastid proteins encoded by the nucleus are synthesized in the cytosol and subsequently transported into the organelle. The majority of the proteins are synthesized as pre-proteins carrying N-terminal signal peptides (cTP, chloroplast Transit Peptide), which are necessary

and sufficient for protein targeting into the plastid stroma (Jarvis, 2008). Pre-proteins are transported through the outer and inner envelope membranes in an energy-consuming process by the multiprotein complexes TOC (Translocon at the Outer envelope membrane of Chloroplasts) and TIC (Translocon at the Inner envelope membrane of Chloroplasts; Jarvis, 2008; Li and Chiu, 2010). Upon translocation of the pre-protein through the inner envelope membrane the signal peptide is cleaved off by the stromal processing peptidase (SPP) (Richter et al., 2005). Some proteins are transported into the chloroplast by alternative, non-canonical pathways via the ER and Golgi (Villarejo et al., 2005; Nanjo et al., 2006; Radhamony and Theg, 2006; Kitajima et al., 2009) or other, so far unidentified pathways (Miras et al., 2002, 2007; Nada and Soll, 2004).

Once (pre-) proteins have entered the plastid stroma they eventually need to be redirected to the thylakoid lumen or membranes. Four distinct pathways are required for the transport of proteins across or into thylakoid membranes: the secretory (Sec) and twin-arginine-translocase (Tat) pathways for transport of proteins across the membrane into the lumen and the signal recognition particle (SRP)-dependent and spontaneous pathways for transport into the thylakoid membrane (Schünemann 2007; Aldridge et al., 2009). Substrates for Sec- and Tat-pathways exhibit bipartite transit peptides with two transport signals in tandem. After the first signal peptide, the cTP, is cleaved off upon arrival of the protein in the stroma, the signal peptide for lumenal localization is exposed and again is removed by a peptidase after translocation to the lumen (Thylakoid Processing Peptidase, TPP, Halpin et al., 1989). The Sec pathway is related to the export mechanism of proteins into the periplasm of bacteria (Mitra et al., 2006; Robson and Collinson, 2006). It requires ATP and protein substrates in an unfolded state for transport (Hynds et al., 1998; Marques et al., 2004). In contrast, the ΔpH-dependent Tat pathway is able to transport folded substrates (Hynds et al., 1998; Marques et al., 2004). Like the Sec pathway, it has a bacterial counterpart (Müller and Klösgen, 2005). In contrast to the bacterial SRP pathway, which targets most inner membrane proteins co-translationally into membranes (Luirink and Sinning, 2004), the plastid SRP pathway seems to be specific to the members of the abundant LHCP family (Light-Harvesting Chlorophyll a/b binding Protein; Schünemann, 2004; Luirink et al., 2005). The thylakoid targeting sequence of SRP-substrates is given by the amino acid sequence of the mature protein (Viitanen et al., 1988). As the name implies, the spontaneous or unassisted pathway of protein insertion does not require any known protein transport machinery. Proteins that are inserted to the membrane spontaneously harbor an N-terminal hydrophobic region, which is exposed after removal of the cTP (Robinson and Mant, 1997).

I.3 Regulation of plastid gene expression

The lateral transfer of major parts of the plastid ancestral genome to the nucleus and the emergence of novel nuclear-encoded plastid-localized proteins demands for coordination of nuclear and plastid gene expression (Figure 2). Retrograde (plastid-to-nucleus) and anterograde (nucleus-to-plastid) signaling mechanisms evolved to permit communication between the two compartments (Woodson and Chory, 2008). Plastids transmit their developmental or functional state by signals, which originate from (i) plastid gene expression, (ii) pigment biosynthesis (e.g. tetrapyrroles), (iii) reactive oxygen species, (iv) redox states of the components of the photosynthetic electron transport, and (v) metabolite pool changes (Pogson et al., 2008; Kleine et al., 2009; Pfannschmidt, 2010). On the other hand, the nucleus controls (i) all steps of plastid gene expression (transcription, RNA-processing, -editing, -stability, translation), (ii) complex assembly, (iii) protein import, and (iv) enzyme activity in response to plastid, developmental and environmental signals (Woodson and Chory, 2008). Additionally, it provides structural components of the photosynthetic thylakoid membrane complexes and other complexes (Figure 1; Herrmann et al., 1985; Wollman et al., 1999).

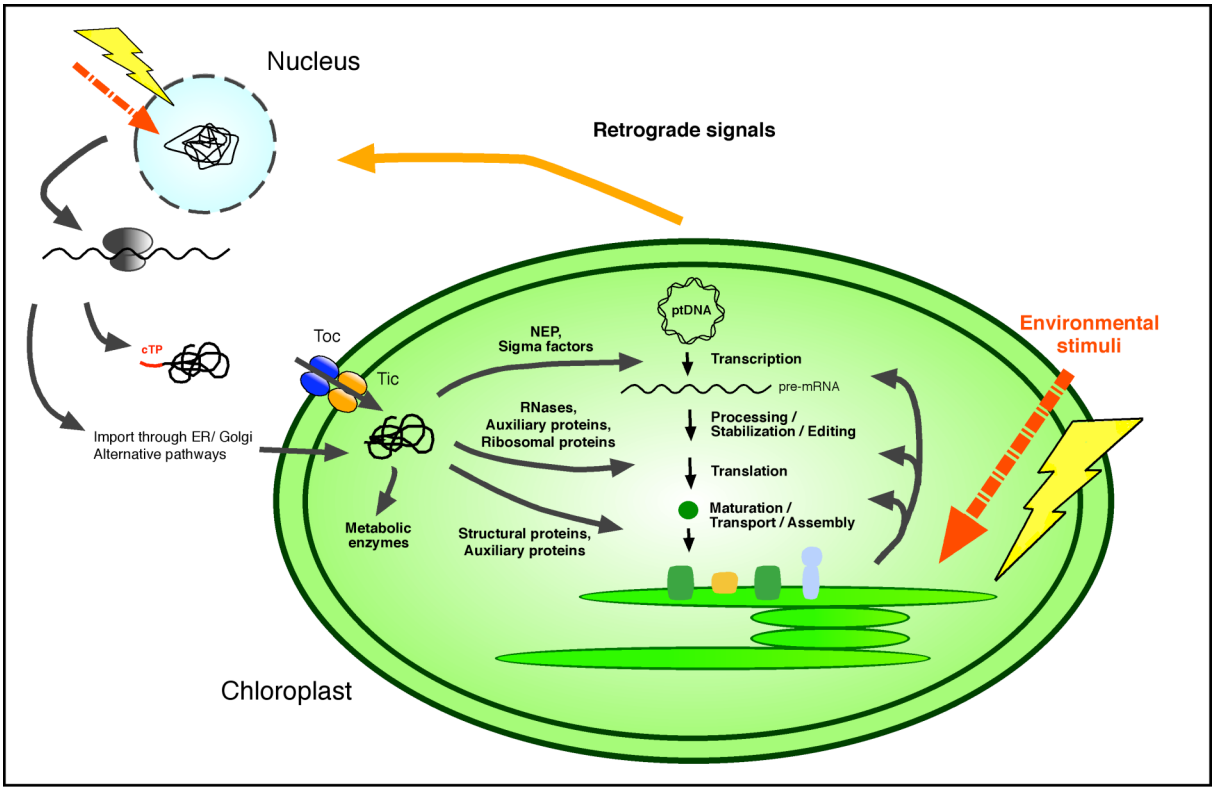


Figure 2: Overview of regulation levels of plastid gene expression.
All levels of plastid gene expression are regulated by retrograde (plastid-to-nucleus) and anterograde (nucleus-to-plastid) signaling mechanisms as well as environmental stimuli (e.g. light).

I.3.1 Regulation of transcription

In higher plants, transcription of plastid genes is performed by distinct RNA polymerases, one plastid-encoded RNA polymerase (PEP) and two nuclear-encoded RNA polymerases (NEP) named RPOTp and RPOTmp (Maliga, 1998; Liere and Börner, 2007). The nuclear-encoded polymerases are monomeric phage-type enzymes (Lerbs-Mache, 1993; Courtois et a., 2007; Swiatecka-Hagenbruch et al., 2007), which have evolved only in seed plants and mosses by duplication of the gene encoding a mitochondrial phage-type enzyme (Hedtke et al., 1997, 2000). These polymerases could not be identified in green algae. Most NEP-dependent promoters exhibit a YRTA sequence motif similar to plant mitochondria promoters (Kühn et al., 2005) and show no similarity to PEP-dependent promoters. NEP is supposed to drive expression of housekeeping genes during early phases of plastid and plant development (Lerbs-Mache, 1993; Mullet, 1993; Hajdukiewicz et al., 1997). However, NEP is also present in mature chloroplasts driving transcription of genes encoding ClpP (a proteolytic subunit of ATP-dependent protease), ribosomal proteins and ribosomal RNA (Bligny et al., 2000; Cahoon et al., 2004; Azevedo et al., 2006; Swiatecka-Hagenbruch et al., 2008).

The plastid-encoded RNA polymerase is a multimeric eubacteria-type enzyme, which recognizes “consensus” promoters having conserved sequences, TTGACA and TATAAT, centered at 35 and 10 bp upstream from a transcription initiation site, respectively (Harley and Reynolds, 1987; Ishihama, 1988; Lonetto et al., 1992). The PEP core enzyme consists of subunits α, β, $\beta'$and β'', which are encoded by the genes *rpoA*, *rpoB, rpoC1*, and *rpoC2* in the plastid genome (Shiina et al., 2005). It is regulated by nuclear-encoded transcription initiators of the sigma type, which bind to the core enzyme to form the holoenzyme and initiate transcription. Most higher plant genomes encode for six sigma factors (SIG1-6; except poplar, which has nine), whereas in green algae only one sigma factor gene has been identified, which is not an ortholog to any plant gene (Carter et al., 2004; Lysenko, 2006). The sigma factors are supposed to have distinct functions in regulation of plastid gene expression (Lysenko, 2007; Lerbs-Mache, 2010). Analyses of sigma factor mutants, mostly carried out in *Arabidopsis*, revealed that at least two sigma factors are essential (SIG2 and SIG6; Shirano et al., 2000; Ishizaki et al., 2005; Loschelder et al, 2006) and that each sigma factor is specific for one or several plastid genes, which may be redundant (Lerbs-Mache, 2010). Transcription of sigma factors SIG1 and SIG5 is induced by red and blue light, thus indicating a function under specific light conditions (Onda et al., 2008). Also, activity of sigma factors is supposed to be modulated by redox reactions, phosphorylation, interaction with other proteins, and eventually proteolytic cleavage (Lerbs-Mache, 2010). Originally, PEP-dependent transcription was assumed to play a role in later developmental stages of plants

by transcription of photosynthesis-related genes (Mullet, 1993), but recent studies revealed that PEP is already active in seeds and is required for efficient germination (Demarsy et al., 2006). Also, both NEP and PEP were found to be active throughout the entire leaf development (Zoschke et al., 2007) and almost all plastid genes are transcribed from several independent promoters allowing access to NEP as well as PEP (Swiatecka-Hagenbruch et al., 2007). Together with the redundancy of some sigma factors, and the lack of NEP and diversity of sigma factors in green algae, these results raise the question whether all these factors are truly required for regulation of gene expression (Lerbs-Mache, 2010). Maier et al. (2008) provide a contrary hypothesis indicating that the large number of regulatory nuclear-encoded factors evolved in order to suppress point mutations occurring in plastid promoter regions, i.e. to assure constant plastid genome transcription by continually creating new promoter/ polymerase pairs.

However, transcription plays a minor role in regulation of plastid gene expression, as transcriptional activity mostly does not correspond to chloroplast transcript levels, in particular during developmental processes (Deng and Gruissem, 1987). In fact, RNA-processing, -splicing, -editing, and -stability are supposed to be the major regulators of plastid gene expression.

I.3.2 Regulation of transcript maturation

I.3.2.1 Transcript processing and stability

In plastids, newly synthesized transcripts undergo several maturation steps prior to translation (Figure 3). Most genes are arranged in polycistronic gene clusters and transcribed from a single promoter (Herrmann, 1992; Sugiura, 1992). As translation of monocistronic RNAs is often more effective than translation of polycistronic forms (Barkan et al., 1994; Sturm et al, 1994; Hirose et al, 1997), intercistronic cleavage by endonucleases is required. Subsequently, 5´ and 3´ ends of cleaved products and of a priori monocistronic transcripts are modified by exonucleolytic and/or endonucleolytic trimming and polyadenylation. All these processes are dependent on nuclear-encoded proteins (Figure 3). Intercistronic cleavage is performed by endoribonucleases, which usually exhibit low sequence specificities, thus being supposed to be guided to cleavage sites by sequence-specific factors. RNase E (Schein et al., 2008) and RNase J (de la Sierra-Gallay et al., 2008) are candidates for endonucleases involved in intercistronic cleavage of plastid transcripts. A PPR (PentatricoPeptide Repeat) family protein, CRR2 from *Arabidopsis*, also exhibits endoribonuclease activity (Okuda et al, 2009) but is specific to the *rps7-ndhB* intergenic region (Hashimoto et al., 2003). Also, sequence-specific proteins without nuclease activity

were found to be required for intercistronic cleavage: The PPR family protein HCF152 from *Arabidopsis* is specific for the processing of the *psbH-petB* intergenic region (Meierhoff et al., 2003) and/ or stabilization of *petB* transcripts (Nakamura et al., 2003). The maize CRP1 protein is involved in processing of the *petB-petD* intergenic region, but it also is associated to translation of the *psaC* and *petA* transcripts (Fisk et al., 1999; Schmitz-Linneweber et al., 2005). Finally, the PPR38 protein from *Physcomitrella patens* was identified as a factor specific for *clpP-rps12* intergenic cleavage and splicing and stability of the *clpP* transcript (Hattori et al., 2007; Hattori and Sugita, 2009).

So far, the mechanism for 5´end formation of transcripts was believed to be accomplished by site-specific intercistronic cleavage by an endonuclease since no protein with 5´ → 3´ exoribonuclease activity was known to be located in plastids. On the contrary, maturation of 3´ ends is supposed to be accomplished by the 3´ → 5´ exoribonuclease activity of polynucleotide phosphorylase (PNPase), which is sensitive to secondary structures and thus arrests at 3´ terminal stem-loop structures (Yehudai-Resheff et al., 2001). However, recently a 5´ → 3´ exoribonuclease function was found for RNase J (Mathy et al., 2007) implying similar mechanisms for 3´ and 5´ end maturation (Figure 3). Additionally, the finding that some termini located in intergenic regions, e.g. *atpI-atpH* and *petB-petD* overlap after processing (Barkan et al., 1994; Pfalz et al., 2009), points to a role of 5´ → 3´ exonucleolytic degradation of 5´ ends. Pfalz et al. (2009) and Prikryl et al. (2010) postulate that endonucleolytic cleavages in intergenic regions occur stochastically and sequences are trimmed exonucleolytically until a prescribed position is reached. This position is either given by the presence of a stem-loop structure as mentioned above or it is marked by specific RNA-binding proteins. In maize, the PPR10 protein was found to define 5´ and 3´ ends of processed *psaJ-rpl33* or *atpI-atpH* intergenic regions, respectively (Pfalz et al., 2009). During binding to the 5´end of *atpH*, PPR10 also activates translation by releasing a ribosome-binding region from an RNA-duplex (Prikryl et al., 2010). A similar function for CRP1, HCF152 and PPR38 rather than their participation in mediating sequence-specific endonucleolytic cleavage is discussed.

Stabilization of transcripts, i.e. protection from exonucleolytic degradation, is accomplished by similar mechanisms. The general pathway of RNA decay in chloroplasts is a three-step mechanism involving 3´ ends, which are usually protected by secondary structures (Stern and Gruissem, 1987; Drager et al, 1996). In the first step, an endonucleolytic cleavage removes 3´ protecting secondary structures or RNA-binding proteins. It is followed by polyadenylation and exonucleolytic degradation (Schuster and Stern, 2009), which is accomplished by PNPase. PNPase exhibits 3´ → 5´ exoribonuclease activity and high

substrate affinity for polyA sequences but it can also function as a polyA polymerase when the ADP:P_i ratio is suitably high (Yehudai-Resheff et al., 2001). Whether other exoribonucleases are also involved in RNA decay is unclear but there is evidence for a role of plastid-localized RNase R in this process (Kishine et al., 2004; Bollenbach et al., 2005).

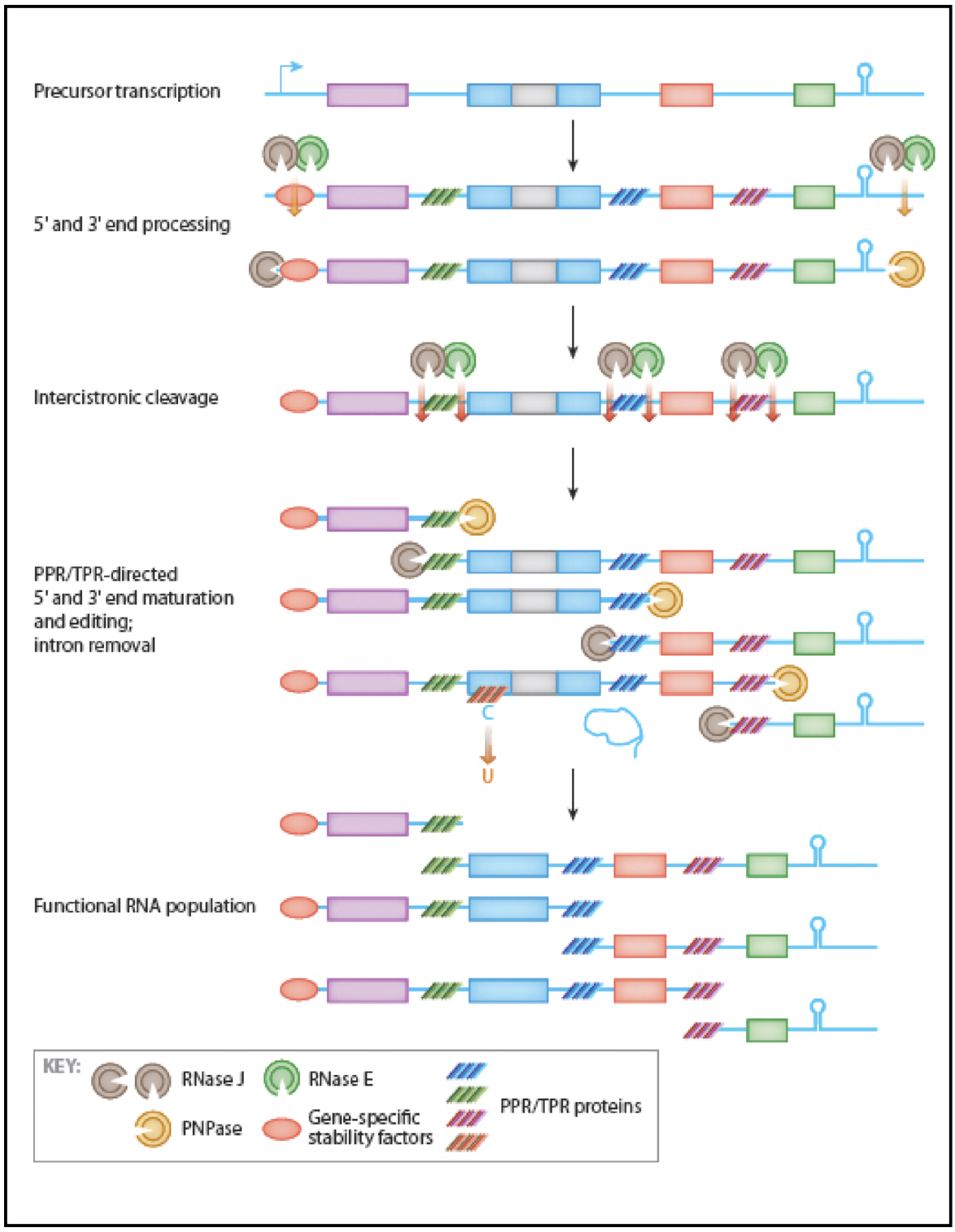


Figure 3: Chloroplast transcript maturation.
Primary, polycistronic transcripts are processed by intercistronic cleavage, intron removal and editing. 5´ and 3´ ends are protected from exonucleolytic degradation by specific factors (Stern et al., 2010).

5´ UTRs from *Chlamydomonas psbD* and *rbcL* transcripts and the tobacco *psbA* transcript were also found to form secondary structures (Nickelsen et al., 1999; Zou et al., 2003; Suay et al., 2005). However, protection from 5´ → 3´ degradation by RNA-binding proteins seems a more prevalent mechanism as a large number of factors has been identified, preferentially in *Chlamydomonas*. The TPR (TetratricoPeptide Repeat) family protein NAC2 is required for stability of the *psbD* transcript (Kuchka et al., 1989; Boudreau et al., 2000). It is part of a high molecular weight complex that binds to the *psbD* 5´end and protects it from 5´ → 3´ degradation (Nickelsen et al., 1999). Recently, NAC2 was found in a 550kDa complex with the translational activator RBP40 indicating tight coupling of 5´ end-mediated RNA stabilization and translation initiation (Schwarz et al., 2007). Accordingly, the 5´ end of *psbB* in *Chlamydomonas* is stabilized by MBB1, which also is a TPR familiy protein and forms a complex of 300kDa (Monod et al., 1992; Vaistij et al., 2000a, 2000b). Stability of the *petA* transcript also is determined by its 5´ end. The PPR family protein MCA1 acts on the first 21 nucleotides of the *petA* transcript and protects it from 5´ → 3´ degradation (Loiselay et al., 2008). Furthermore, the abundance of MCA1 was found to be the rate-limiting factor of *petA* transcript accumulation (Raynaud et al., 2007). Other 5´ end stabilizing proteins from *Chlamydomonas*, which should be mentioned, are the *petD*-specific MCD1 protein (Murakami et al., 2005) and the PPR family protein MRL1, which targets *rbcL* transcripts and was found in *Arabidopsis*, too (Johnson et al., 2010). Compared to the large number of 5´ end stabilizing factors from *Chlamydomonas*, only few factors were identified in higher plants. Next to the abovementioned MRL1 protein and PPR10, which was discussed in conjunction with transcript maturation, only HCF145 from *Arabidopsis* presumed to be involved in *psaA*, *psaB*, and *rps14* transcript stability (Lezhneva and Meurer, 2004) can be listed in this context.

To summarize, transcript processing and stability are tightly coupled processes. Nuclear-encoded proteins are supposed to bind to predetermined sequences of both 5´ and 3´ ends thus protecting transcripts from 5´ → 3´ and 3´ → 5´ exonucleolytic cleavage and defining their termini.

I.3.2.2 RNA splicing

Many plastid transcripts (mRNAs and tRNAs) harbor introns, which must be removed prior to translation (Figure 3). Depending on splicing mechanisms and primary and secondary structures, organellar introns are classified into group I and group II introns, respectively (Houghland et al., 2006; Pyle and Lambowitz, 2006; Fedorova and Zingler, 2007). Group II introns are sub-divided into group IIA to group IID, with only group IIA and IIB occurring in flowering plants (Michel et al., 1989).

Higher plant plastids exhibit only one group I intron located in the *trnL* transcript, whereas there are five in *Chlamydomonas reinhardtii* and even more in other *Chlamydomonas* species (Turmel et al., 1993). However, there is a set of about 20 group II introns conserved among higher plant organellar genomes (Schmitz-Linneweber and Barkan, 2007). Originally, introns were mobile genetic elements encoding a maturase, which facilitates splicing of the host intron (Eickbush, 1999; Wank et al., 1999). This ability has been widely lost in higher plants. The only remaining intron-encoded maturase, MatK, is encoded by the *trnK* transcript and was found to co-immunopurify with all group IIA introns except *clpP* intron 2 (Zoschke et al., 2010). In fact, splicing of plastid introns is rather mediated by at least 12 nuclear-encoded proteins (de Longevialle et al, 2010).

Two main protein families were found in this context: CRM (Chloroplast RNA splicing and ribosome Maturation; Barkan et al., 2007) domain proteins and PPR proteins (Schmitz-Linneweber and Small, 2008). Additionally, some proteins harboring different sequence motifs are involved in splicing.

Five CRM domain proteins have been identified: CRS1 (Chloroplast RNA Splicing 1), CFM2 and CFM3 (CRM Family Member 2 and 3), and CAF1 and CAF2 (CRS2-Associated Factors 1 and 2; Stern et al., 2010). CRS2, a peptidyl-tRNA hydrolase (PTH) related protein, which lost PTH activity, is supposed to participate directly in splicing (Jenkins and Barkan, 2001). The CAF1 and CAF2 proteins interact with CRS2 and target overall nine group II introns in maize (Ostheimer et al., 2003; Asakura and Barkan, 2006). Also, the other CRM domain proteins are involved in splicing of several targets (Stern et al., 2010). In contrast, PPR proteins involved in splicing seem to be generally specific to one transcript. Hence, the maize PPR4 and PPR5 proteins are involved in *rps12*- and *trnG*-intron splicing, respectively, whereas in *Arabidopsis*, splicing of *ycf3* intron 2 is dependent on the PPR protein OPT51 (Organellar Transcript Processing 51; Schmitz-Linneweber et al., 2006; Beick et al., 2008; de Longevialle et al., 2008).

Despite the large number of splicing factors identified, there is no concrete mechanistic model for their function. It is hypothesized that one or more factors interact with introns in a sequence-specific manner and change its folding, thus enabling recognition by general splicing factors (de Longevialle et al., 2010). Whether this is the general mechanism of plastid intron splicing, needs to be further elucidated.

I.3.2.3 RNA editing

Editing is the process of modification of transcripts by altering individual nucleotides. It is a common mechanism in all plants, except liverworts (Freyer et al., 1997) and it is absent in algae and cyanobacteria. In general, cytidine (C) is exchanged to uridine (U) in chloroplasts,

but in some hornworts and ferns U to C conversion also occurs (Kugita et al., 2003; Wolf et al., 2004). Typically, there are about 30 editing sites in plastids of vascular plants, e.g. *Arabidopsis* has 34 editing sites in 18 different genes (Tillich et al., 2005; Chateigner-Boutin and Small, 2007). Editing often restores conserved amino acids required for protein function (Zito et al., 1997; Sugita et al., 2006) or creates new translation initiation codons (Hoch et al., 1991; Kudla et al., 1992; Neckermann et al, 1994).

The mechanism of editing is so far unknown. However, both, *cis*- and *trans*-acting elements have been proposed to play a role. Analyses of sequences surrounding editing sites hint at the requirement of elements immediately upstream of the respective site and it is speculated that these sites serve as binding sites for *trans*-acting elements (Chaudhuri and Maliga, 1996; Bock et al., 1996; Sasaki et al., 2006). Indeed, several exclusively nuclear-encoded *trans*-acting factors have been identified and all, except one, belong to the family of PPR proteins. CRR4, the first *trans*-acting factor to be identified, is required for editing of a C at position 2 in *ndhD* transcripts and thus restoration of an AUG start codon (Kotera et al., 2005). It is the only factor that has been demonstrated to directly bind to its target editing site (Okuda et al., 2006; 2007). Like CCR4, the majority of *trans*-acting PPR family proteins are specific to one transcript (Okuda et al., 2007; Zhou et al., 2008; Cai et al., 2009). However, five PPR family proteins have been reported to target up to three editing sites, e.g. OTP84 targets *ndhF* C290, *psbZ* C50, and *ndhB* C1481 (Hammani et al., 2009) suggesting that editing factors must be able to distinguish pyrimidines from purines and sometimes even recognize specific bases (Hammani et al., 2009).

The only non-PPR family protein involved in RNA editing is CP31, which belongs to the family of plastid ribonucleoproteins characterized by a twin RNA recognition motif (RRM) and has two paralogs in *Arabidopsis*, CP31A and CP31B (Tillich et al., 2009). Since *cp31a* and *cp31b* mutants are affected in editing of several, but not all, sites, CP31A and CP31B proteins are supposed to act supportive to PPR proteins in recognition of editing-site *cis*-elements (Tillich et al., 2009).

I.3.3 Regulation of translation

Translation of plastid transcripts is performed by bacterial-type 70S ribosomes according to the prokaryotic origin of chloroplasts. Although there is a certain degree of conservation of translation factors and most ribosomal proteins, bacterial and plastid translation machineries differ (Beligni et al., 2004). Differences include lack or divergence of some ribosomal proteins and gain of protein components for translational regulation of plastid gene expression (Zerges, 2000; Beligni et al., 2004). Thus, both the large and small ribosomal subunit also include nuclear-encoded plastid specific ribosomal proteins (PSRP), which were identified

and characterized in spinach and *Chlamydomonas* plastids (Subramanian, 1993; Yamaguchi et al., 2002, 2003). Also, several nuclear-encoded proteins from *Chlamydomonas* and higher plants have been identified as transcript-specific translation regulators. The function of some of these factors will be discussed later on in the text.

Besides varying *trans*-acting factors, the organization of *cis*-elements displays a major difference between plastid and bacterial translation. In bacteria, translation is initiated by binding of the 3´ terminus of the 16S rRNA to a Shine-Dalgarno (SD) sequence (typically GGAGG) in the 5´ UTR of the transcript. The SD sequence is located between 5 and 9 nucleotides upstream from the start codon and mediates correct positioning of the ribosome for translation initiation (Yusupova et al., 2001; Laursen et al., 2005). SD-like sequences can also be found in 5´ UTRs of plastid transcripts, but they differ in position and sequence (Sugiura et al., 1998). While translation of some transcripts is dependent on SD-like sequences (*rbcL*, *atpE* and *rps14* from tobacco; *psbA* from *Chlamydomonas*), other transcripts are only partially dependent (*rps12* and *petB* from tobacco) or independent of them (*psbA* and *atpB* from tobacco; *petD*, *atpB*, *atpE*, *rps4* and *rps7* from *Chlamydomonas*); Mayfield et al., 1994; Sakamoto et al., 1994; Hirose and Sugiura, 1996; Fargo et al., 1998; Hirose and Sugiura, 2004a, 2004b). Interestingly, the *rps2* SD-like sequence from tobacco, was even shown to be a negative regulatory element for translation (Plader and Sugiura, 2003). The dispensability of SD-like sequences in 5´ UTRs of plastid transcripts points to the existence of distinct regulatory *cis*-elements. As an example, the 5´ UTR of the *psbA* transcript in tobacco possesses a SD-like sequence located at -33 nt from the start codon, but it is dispensable for translation (Hirose and Sugiura, 1996), which stands in contrast to the requirement of a SD-like sequence in *Chlamydomonas psbA* translation (Mayfield et al., 1994). Indeed, three other elements within the 5´ UTR of tobacco *psbA* have been identified: two sequences complementary to the 3´ terminus of chloroplast 16S rRNA (RBS1 and RBS2) located at -9 and -22 nt upstream of the start codon, respectively and an AU rich sequence (UAAAUAAA) located between RBS1 and RBS2 (Hirose and Sugiura, 1996). It is hypothesized that RBS1 and RBS2 interact cooperatively with the 3´ end of 16S rRNA resulting in looping out of the AU rich sequence, which facilitates the interaction of *trans*-acting factors (Hirose and Sugiura, 1996). Similar sequence elements have been described for the *psbD* 5´ UTR of *Chlamydomonas*, with a SD-like motif (PRB1) and the PRB2 site at -13 nt and -30 nt upstream of the start codon, respectively and a polyU rich element located in between (Nickelsen et al., 1999). Like the sequence elements from *psbA* 5´ UTR, those of *psbD* are supposed to serve as binding sites for *trans*-acting factors (Nickelsen et al., 1999; Ossenbühl and Nickelsen, 2000). Other *cis*-elements affecting protein synthesis were found in *psbC*, *petD* and *rps*7 5´ UTRs in *Chlamydomonas* (Zerges et al., 1997; Fargo et al., 1999;

Higgs et al., 1999; Zerges et al., 2003) and the *atpB* mRNA in tobacco (Hirose and Suigura, 2004b) mainly in the context of targets for *trans*-acting factors.

Trans-acting factors of translation are encoded by the nucleus and generally are specific to single transcripts. They have been implicated in general regulation of plastid translation, but also light- and assembly-dependent regulation of translation (Marin-Navaro et al., 2007). The most striking example of light-regulated translation is the *psbA* transcript in *Chlamydomonas*. A stem-loop structure in the 5´ UTR (Mayfield et al., 1994) is bound by a complex of several proteins in a light-dependent manner: RB47 is a poly A-binding protein (cPAB1) whose binding to the *psbA* 5´UTR is enhanced under reducing conditions (Kim and Mayfield, 1997; Yohn et al., 1998). Modulation of cPAB1 RNA binding properties in response to the redox status of the cell is supposed to be regulated by RB60, a protein disulfide isomerase homolog (cPDI), and TBA1, an oxidoreductases homolog (Kim and Mayfield, 1997; Somanchi et al., 2005; Algerand et al., 2006). Additionally, ADP-dependent phosphorylation of cPDI is presumed to control the RNA-binding capacity of cPAB1 (Danon and Mayfield, 1994). Another protein, RB55, was identified as part of the complex binding to the *psbA* 5´ UTR, but it has not been cloned or characterized so far (Danon and Mayfield, 1991). Finally, RB38, another RNA-binding protein, which is independent of the redox status, was described to bind specifically to uridine-rich sequences of the *psbA* 5´UTR (Barnes et al., 2004). However, RB38 was now found to be identical to RBP40 from *Chlamydomonas*, which is involved in *psbD*- rather than *psbA*-translation (Schwarz et al., 2007). RBP40 is supposed to bind to the polyU-rich element of the *psbD* 5´UTR (Ossenbühl and Nickelsen, 2000) and destabilize the stem-loop structure thereby enabling binding of the small ribosomal subunit to the initiation codon (Klinkert et al., 2006). Guidance of RBP40 to its target sequence is predicted to be mediated by the stability factor NAC2 (see 3.2.1), which, together with RBP40, is part of a high molecular weight complex of approximately 550 kDa (Schwarz et al., 2007). As for *psbA*, translation of *psbD* is induced by light (Malnoe et al., 1988), but it is not resolved whether the RBP40/NAC2 complex participates in light regulation.

Besides light regulation, plastid translation can also be modulated by the assembly status of proteins into their respective complexes. This regulatory process is referred to as “Control by Epistasy of Synthesis” (CES) and has been reported for all four thylakoid membrane complexes (Choquet et al., 1998; Wostrikoff et al., 2004; Minai et al., 2006; Drapier et al., 2007). Regulation of the cytochrome *f* subunit of the cytochrome b_6f complex from *Chlamydomonas* is the best-studied process so far. When not assembled in the complex, cytochrome *f* interacts with MCA1, which usually binds and stabilizes the *petA* transcript and interacts with the translational factor TCA1 (see I.3.2.1). Interaction with cytochrome *f* leads to degradation of MCA1 and thus destabilization of the *petA* transcript and prevention of its

translation, and therefore autoinhibition of cytochrome *f* synthesis (Raynaud et al., 2007; Boulouis et al., 2011).
A number of other nuclear-encoded *trans*-acting factors regulating translation have been identified in *Chlamydomonas* and higher plants (Marin-Navaro et al., 2007). These include CRP1 for *petA*, *psaC*, and *petD* in maize (Barkan et al., 1994; Schmitz-Linneweber et al., 2005), TBC1, TBC2, and TBC3 for *psbC* in *Chlamydomonas* (Auchincloss et al., 2002; Zerges et al., 2003), HCF173 for *psbA* in *Arabidopsis* (Schult et al., 2007), ATAB2 for multiple targets from PSI and PSII in *Arabidopsis* (Barneche et al., 2006), and PPR10 for *atpH* in maize (Prikryl et al., 2010). However, the precise function of the majority of these factors has not been elucidated so far.

Regulation of initiation is definitely the most important step of plastid translation. However, elongation and termination also display targets for regulatory processes. Elongation of the D1 protein is regulated by light (Muhlbauer and Eichacker, 1998; Zhang et al., 2000), presence of co-factors like chlorophyll (Mullet et al., 1990) and the availability of assembly partners (de Vitry et al., 1989). The nuclear-encoded factors, VIR-115 from barley and AC115 from *Chlamydomonas*, are assumed to stabilize translation intermediates of D1 or D2, respectively and facilitate their folding, co-factor binding, and/or co-translational membrane insertion (Kim et al., 1994; Rattanachaikunsopon et al., 1999).
Recognition of stop codons and translation termination is dependent on release factors in both eukaryotes and prokaryotes. In bacteria, release factors RF1 (prfA) decoding UAA and UAG and RF2 (prfB) decoding UAA and UGA have been identified (Scolnick et al, 1968; Nakamura and Ito, 1998). In *Arabidopsis*, a homolog to prfB, AtPRFB or HCF109, has been cloned and shown to be required for termination of transcripts with UGA and UAA stop codons and regulate transcript stability and protein synthesis (Meurer et al., 2002).

I.3.4 Posttranslational regulation: protein maturation and complex assembly

The last important steps of plastid gene expression include protein maturation, i.e. co-factor binding and proteolytic processing, and the assembly of all subunits to form an active complex. Additionally, thylakoid membrane complexes have to be inserted into the lipid bilayer. All of these processes can occur post- or co-translationally and are dependent on both general and complex-specific auxiliary factors.

I.3.4.1 Membrane insertion

Insertion of complex subunits into the thylakoid membrane is facilitated by the “spontaneous“ or the SRP pathway (see I.2), the latter being specific to LHCP family proteins (Schünemann, 2004; Luirink et al., 2005). Moreover, LHCP membrane insertion is dependent on the ALB3 translocase, which is the plastid homolog to the membrane insertion factors Oxa1p and YidC from mitochondria and *Escherichia coli*, respectively (Moore et al., 2000; Bellafiore et al., 2002; Kuhn et al, 2003). The SRP pathway is believed to recruit LHCPs from the stroma and target them to the membrane and ALB3 (Aldridge et al., 2009). ALB3 also interacts with the chloroplast SECY translocase (Klostermann et al., 2002; Pasch et al., 2005), which is supposed to be involved in co-translational membrane insertion of the D1 protein (Zhang et al., 2001). Yeast split-ubiquitin and co-immunoprecipitation experiments also revealed a direct interaction between ALB3 and subunits from PSII and PSI and the CF_0III subunit from ATP-synthase (Pasch et al., 2005; Göhre et al., 2006), indicating a general role of ALB3 in membrane insertion of proteins. The maturation and assembly of thylakoid membrane complexes is mediated by more specific factors since each complex and each subunit exhibit different properties.

I.3.4.2 Posttranslational modifications / subunit maturation

Proteolytic truncation of proteins to generate their mature status is one of the maturation events occurring in plastids. In such a way, the PSII reaction center protein D1 is synthesized as a precursor (pD1) with a C-terminal extension of 9 amino acid residues in higher plants (Marder et al., 1984; Takahashi et al., 1988). Towards its integration into the thylakoid membrane, the D1 precursor is processed to its mature form by the nuclear-encoded lumenal Carboxy-Terminal Processing protease (CtpA; Anbudurai et al., 1994; Fujita et al., 1995). Recently, additional components facilitating D1 precursor processing have been identified in *Synechocystis* (PratA; Klinkert et al., 2004) and *Arabidopsis* (LPA19; Wei et al., 2010), although their precise function is not clear yet. C-terminal processing was reported to be a prerequisite for binding of the manganese (Mn_4Ca) cluster to PSII (Nixon et al., 1992; Roose and Pakrasi, 2004).

Another aspect of protein maturation is binding of co-factors. The photosynthetic complexes are associated with pigments for light harvesting and charge separation, and redox-active factors for electron transfer. PSII possesses several chlorophylls and xanthophylls, two pheophytins, two c-type haems, one manganese (Mn_4Ca) cluster, one bicarbonate, two Ca^{2+} ions, one Fe^{2+} ion, and one Cl^- ion (Guskov et al., 2009). Similarly, PSI is associated with a large number of chlorophylls and xanthophylls, according to the function of both

photosystems in light harvesting. Redox reactions are utilized by two phylloquinones and three [4Fe-4S] clusers (Amunts et al., 2007). Finally, the cytochrome b_6f complex exhibits four hemes (two b- and c-type hemes each), one [2Fe-2S] cluster and one chlorophyll and xanthophyll each (see also Figure 1; Stroebel et al., 2003; Kurisu et al., 2003).

Plastid Fe-S clusters are assembled by a machinery related to the eubacterial SUF (Sulfur Utilization Factor) system (Balk and Pilon, 2011). Attachment of assembled clusters to their target proteins seems to require further components, such as scaffold and carrier/ transfer proteins. So far, two proteins were found to be involved in [4Fe-4S] cluster binding to PSI: HCF101 and RubA (Schöttler et al, 2011). HCF101 was found to transiently bind [4Fe-4S] clusters and transfer them onto apoproteins (Schwenkert et al., 2010). The cyanobacterial RubA protein is specifically required for incorporation of the [4Fe-4S] cluster associated with PsaAB (F_x; Shen et al., 2002). RubA-homologs are encoded by higher plant genomes but their function has not been elucidated so far.

As mentioned above, the cytochrome b_6f complex is associated with four hemes. The two non-covalently bound b-type hemes are bound by the cytochrome b_6 polypeptide and are supposed to bind spontaneously (Robertson et al., 1994). On the contrary, the two c-type hemes bind covalently to cytochrome b_6 and cytochrome *f*, respectively.

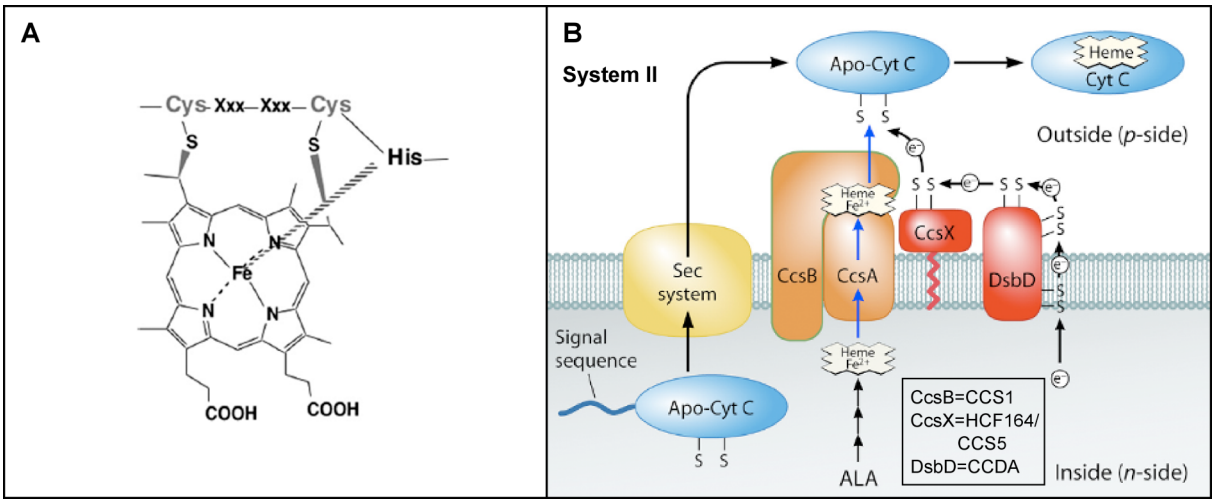


Figure 4: C-type cytochrome maturation in plastids
(A) C-type heme ligated to a CXXCH heme binding motif of cytochrome. Cysteines form thioether bonds with vinyl groups of the heme; the histidine acts as a ligand to the heme iron atom (Giegé et al., 2008). (B) System II maturation pathway operating at the electropositive, lumenal side of the thylakoid membrane. The apocytochrome is transported to the lumen by the Sec-pathway; HCF164/ CCS5/CcsX and CCDA/ DsbD are required for thioreduction of apocytochrome; CCSA and CCS1/ CcsB transport and provide heme and chaperone the apocytochrome (Kranz et al., 2009).

For biogenesis of c-type cytochromes, three distinct pathways (system I-III) have been characterized so far, although there is evidence for additional mechanisms (Allen et al., 2004). Whereas system III, which occurs in mitochondria of fungi, vertebrates and

invertebrates, consists of a heme lyase as the single component (Dumont et al., 1987; Giegé et al., 2008), systems I and II are more complex. System I (or Ccm system) occurs in α- and γ-proteobacteria and mitochondria of plants and red algae, and archaea and consists of up to ten membrane-bound proteins (Allen et al., 2008; Sanders et al., 2010). System II operates in some gram-positive bacteria, cyanobacteria, some β-, δ- and ε-proteobacteria and plastids of plants and algae (Kranz et al., 2002). In plastids, it is required for biogenesis of cytochromes *f* and c_6 (Rurek, 2008). Cytochrome *f* possesses a classical CXXCH heme-binding motif whose two cysteines form thioether linkages with the vinyl groups of the heme (Figure 4 A). The apoprotein is synthesized in the plastid stroma but targeted to the electropositive thylakoid lumen in a Sec-dependent way (Rohl and van Wijk, 2001), where the conversion of apo- to holo-cytochrome *f* occurs (Figure 4 B). Ligation of heme to cytochrome *f* requires several processes: heme, which is synthesized in the stroma, needs to be transported to the lumen and must be transferred to or kept in its reduced state to assure that the vinyl groups are chemically active; also, the cysteine residues of the heme binding motif must be reduced to form thioether bonds to the heme vinyl groups (Kranz et al., 2009). From genetic studies in *Chlamydomonas* at least six factors are supposed to be involved in cytochrome *f* (and c_6) maturation: the nuclear-encoded proteins CCS1-CCS5, and the plastid-encoded CcsA (Howe and Merchant, 1992; Xie and Merchant, 1996; Inoue et al., 1997; Xie et al., 1998). CcsA and CCS1 (also: CCSB) are supposed to form a complex and function in heme delivery and its ligation to the apoprotein (Figure 4 B; Dreyfuss et al., 2003; Hamel et al., 2003). CcsA contains a “WWD” motif (WGXϕWXWD, where ϕ is an aromatic residue) exposed to the lumen, which is also found in two system I components and is thought to transport heme into the thylakoid lumen and present it to the apocytochrome for ligation (Goldman et al., 1998; Xie and Merchant, 1998; Rurek, 2008; Kranz et al, 2009). CCS1, on the other hand, is supposed to chaperone reduced apocytochromes (Rurek, 2008). HCF164, a thioredoxin-like protein from *Arabidopsis* was also found to be required for cytochrome *f* maturation (Lennartz et al., 2001) and recently, CCS5 was described as its homolog in *Chlamydomonas* (Gabily et al., 2010). Also, the thiol disulfide transporter homolog CCDA is required for cytochrome *f* maturation in *Arabidopsis* but it has not yet been characterized in *Chlamydomonas* (Page et al., 2004). Both proteins seem to be involved in a thiooxidation/ thioreduction pathway in which reducing equivalents are transferred to the electropositive side of the thylakoid membrane (lumen) and used for the reduction of apocytochromes (Figure 4 B; Rurek, 2008). Recently, the CCS4 protein was identified and shown to also play a role in the reducing pathway (Gabilly et al., 2011).

In contrast to cytochrome *f*, cytochrome b_6 binds its c-type heme (heme c_n) on the electronegative side of the thylakoid membrane, the stroma, by only a single cysteine residue

(Cys35; Kurisu et al., 2003; Stroebel et al., 2003). The maturation pathway mediating this atypical heme binding is a part of this thesis.

I.3.4.3 Assembly of complexes

Assembly of thylakoid membrane complexes is accomplished in a step-by-step manner. Usually, one protein serves as an anchor or scaffold for subsequent assembly steps. For PSII, the D2 protein, together with the cytochrome b_{559} α and β subunits seems to be the initiation point of assembly (Figure 5; Komenda et al., 2004; Rokka et al., 2005; Minai et al., 2006). Subsequently, the D1 and PsbI proteins are attached to form the reaction center (RCII; Figure 5; Rokka et al., 2005; Minai et al., 2006; Dobrakova et al., 2007). Assembly of the inner antenna protein CP47 forms the RC47 complex and facilitates binding of several small subunits like the phosphoprotein PsbH (Figure 5; Komenda et al., 2004; Rokka et al., 2005; Minai et al., 2006). After incorporation of CP43 the monomeric core complex is complete and assembly of the manganese (Mn_4Ca) cluster and oxygen evolving complex can occur, as well as dimerization and supercomplex formation (Figure 5; Rokka et al., 2005). A multitude of assembly-assisting proteins has been identified for PSII, although their mode of action is largely unknown. LPA1 from *Arabidopsis* is an intrinsic membrane protein and is supposed to function as a chaperone for PSII assembly by interaction with D1 (Figure 5; Peng et al., 2006). Another protein, PAM68 is assumed to associate with LPA1 and D1 in an early intermediate complex (Figure 5; Armbruster et al., 2010). Folding and/or assembly of CP43 into PSII are dependent on the ALB3 interacting proteins LPA2 and LPA3 (Figure 5; Ma et al., 2007; Cai et al., 2010). Similarly, HCF136 is involved in an early assembly step of PSII assembly, as in absence of this factor, PSII subunits are synthesized but no stable complexes can be detected (Figure 5; Meurer et al., 1998; Plücken et al., 2002). At the level of supercomplex formation AtFKBP20-2, a lumenal immunophilin, and Psb29 were reported to play a role (Figure 5; Keren et al., 2005; Lima et al., 2006).

PSII is not only synthesized *de novo*, but it is constantly renewed due to photooxidative damage (Aro et al., 1993). The D1 protein is the major target of the PSII repair cycle (Mattoo et al., 1981; Ohad et al., 1990). Damaged PSII complexes are phosphorylated and migrate from the grana to the stoma lamelae, where D1 is degraded and replaced by a newly synthesized protein (Mulo et al., 2008). Degradation is supposed to be predominately performed by FtsH and Deg proteases (Kato and Sakamoto, 2009) and in *Arabidopsis*, the lumenal AtTLP18.3 protein was shown to be involved in degradation of D1, but also in dimerization of PSII reaction centers (Sirpiö et al., 2007). Chlorophyll released during repair may be bound by ELIP proteins in higher plants, which belong to the LIL (Light-harvesting-Like) protein family (Hutin et al., 2003). Reassembly of PSII is assumed to require the

abovementioned general assembly factors plus repair-specific factors, like Hsp70 proteins and PPL1 (Ishihara et al., 2007; Mulo et al., 2008).

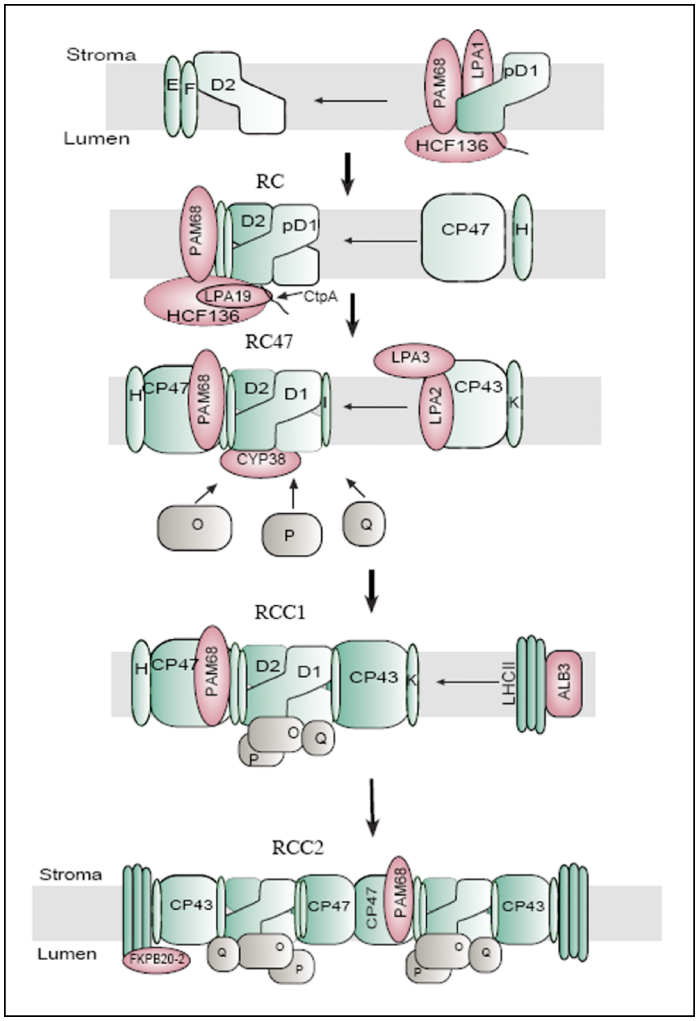


Figure 5: Roles of known auxiliary proteins in regulation of PSII assembly (Meierhoff and Westhoff, 2011).

In contrast to the large number of auxiliary factors for assembly of PSII, only a few proteins have been identified in conjunction with PSI and the cytochrome b_6f complex. The starting point for cytochrome b_6f complex assembly is cytochrome b_6, which rapidly interacts with subunit IV. This dimer is a prerequisite for cytochrome *f* binding and subsequent assembly of the small subunits PetG, PetL, PetM, and PetN. The Rieske protein appears to be bound only loosely to the complex and is likely to contribute to dimer stability, which is the active form of the cytochrome b_6f complex (Choquet and Vallon, 2000; Kurisu et al., 2003; Stroebel et al., 2003). Biogenesis factors have only been identified in the context of (i) processing and/ or stabilization of transcripts (see 3.2.1 and 3.2.2), (ii) stabilization of transcripts and CES autoregulation of cytochrome *f* (see section/chapter 3.3), and (iii) co-factor binding (see 3.4.3), which is important for protein stability (Kuras et al., 1995; de Vitry et al., 2004). Probably, HCF153 from *Arabidopsis* is involved in assembly of cytochrome b_6f, because

mutants exhibit drastically reduced amounts of the major subunits and their translation is not impaired (Lennartz et al., 2006). However, the molecular function has not been revealed, yet. PSI assembly is initiated by the co-translational insertion of PsaA and PsaB, which then bind to their co-factors to form the PsaAB reaction center. Subsequently, PsaC with its two [4Fe-4S] clusters is attached, followed by PsaD, PsaE and other subunits (Schöttler et al., 2011). So far, the only nuclear-encoded auxiliary protein involved in PSI assembly is Pyg7-1 but its function has not been defined (Stöckel et al., 2006).

II. Aims of this PhD thesis

Many nuclear-encoded factors that are involved in biogenesis of the photosynthetic complexes act in concert with other proteins. Thereby, they may either interact transiently or form stable complexes. The focus of this thesis was to provide tools for interaction analyses and to characterize complex formation of two distinct nuclear-encoded auxiliary factors for biogenesis of the cytochrome b_6f complex and photosystem II (PSII).

(i) Manuscript 1 (Lyska et al., 2011) describes the development, cloning and characterization of binary vectors for in-frame translational fusions of genes encoding chloroplast biogenesis factors with different affinity tags. Those tags can be used for detection of proteins and affinity purification, and therefore analysis of protein-protein interactions and complexes *in planta*. To ensure proper expression of fusion proteins, three endogenous promoter sequences from genes encoding chloroplast biogenesis factors were cloned and characterized using β-glucoronidase (GUS) reporter fusions. Different tags and promoters were tested for their interference with protein function of the nuclear-encoded chloroplast biogenesis factors High Chloroplhyll Fluorescence 208 (HCF208) and HCF107.

(ii) Manuscripts 2 (Lyska et al., 2007) and 3 describe the identification and characterization of the cytochrome b_6f complex biogenesis factor HCF208. The *hcf208* mutant was characterized biochemically and the *HCF208* locus was determined by positional cloning and confirmed by complementation of the mutant phenotype with the wild-type cDNA. Interactions of HCF208 with other proteins were assayed using 2D-Blue Native-PAGE, the Yeast Split Ubiquitin system and affinity purification.

(iii) Manuscript 4 describes analyis of the PSII biogenesis factor HCF107. HCF107 was previously identified as a nuclear-encoded factor required for the expression of the small PsbH subunit of PSII and it was found to act in a membrane-associated protein complex. Using sucrose density gradient centrifugation and an HCF107-specific antibody, the dynamics of complex formation and localization in dependence on light were analyzed.

III. Theses

Nuclear-encoded auxiliary proteins regulate the expression of plastid-encoded subunits of the photosynthetic complexes as well as their maturation and assembly.

Work presented here demonstrates:

1) The Gateway-based pAUL-vector series represents an adequate tool for moderate expression of nuclear-encoded proteins required for chloroplast biogenesis. Further, the pAUL-vectors allow differential tagging of proteins and their detection and purification.
2) The nuclear-encoded factor HCF208/ AtCCB2 s required for the maturation of the cytochrome b_6 subunit of the cytochrome b_6f complex by facilitating attachment of the covalent heme c_n at the stromal side of the cytochrome b_6f complex. Mutants affected in this factor accumulate less cytochrome b_6 amounts; however, the remaining protein is capable of complex formation.
3) HCF208/ AtCCB2 acts in concert with three other factors AtCCB1, AtCCB3 and AtCCB4. HCF208/ AtCCB2 and AtCCB4 form stable heterodimers, which transiently interact with AtCCB3 via HCF208/ AtCCB2 to facilitate heme attachment to cytochrome b_6.
4) The nuclear-encoded factor HCF107 (Sane et al., 2005) is required for the stability and translation of transcripts encoding for the photosystem II subunit PsbH. HCF107 forms a membrane-associated high molecular weight complex, which presumably is involved in translation. Formation of this complex and the association with HCF107 with chloroplast membranes are light-regulated processes.

IV.1 Summary

Chloroplast biogenesis is controlled by nuclear-encoded factors that are synthesized in the cytosol and subsequently imported into the plastid. In higher plants and green algae, a large number of auxiliary proteins regulating the expression (transcription, RNA-processing, -editing, -stability, translation), maturation and assembly of the photosynthetic complexes have been identified. Since many of these factors act in concert with other proteins and/ or as part of protein complexes, analyses of interactions and complex formation can help shading light on their precise function.

In the present PhD thesis a set of plant binary vectors for differential tagging of proteins in *Arabidopsis thaliana* was developed. Diverse mono-, di- and triple-tags enable detection and flexible purification of proteins and their interaction partners. To prevent overexpression of tagged proteins, which is disadvantageous for co-purification of interaction partners, promoter sequences from three chloroplast biogenesis factors (HCF107, HCF136 and HCF173) were chosen to drive expression. Promoters were analyzed histochemically and quantatively by β-glucoronidase (GUS) fusions. Different promoter/ tag combinations were tested for their influence on protein function using the chloroplast biogenesis factors HCF208 and HCF107. Single-step and tandem affinity purification of HCF208 via different tags confirmed the integrity of the cloned tags

In the second part of this thesis, the factor HCF208 from *Arabidopsis* was identified by screening of a mutant collection carrying ethylmethane sulfunate (EMS)-induced point mutations for lines affected in photosynthetic electron transport. By spectroscopic measurements and immunological analyses the *hcf208* mutant was shown to be deficient in biogenesis of the cytochrome b_6f complex, primarily by absence of the covalent heme c_n of cytochrome b_6. *HCF208* was found to encode for the *Arabidopsis* homolog of *Chlamydomonas reinhardtii* CCB2, which acts in concert at least three other proteins (CCB1, CCB3 and CCB4) displaying a new pathway of heme attachment to cytochrome b_6, referred to as “system IV”. As analyzed in this study *Arabidopsis* T-DNA mutants affected in loci encoding for the AtCCB1, AtCCB2 and AtCCB4 proteins and the *hcf208* mutant exhibited similar phenotypes. HCF208 was shown to be an integral membrane protein. Interaction analyses using two-dimensional Blue Native-PAGE, yeast split ubiquitin assays and affinity purification of tagged HCF208 indicated a stable dimerization between HCF208 and AtCCB4, and a transient HCF208-AtCCB3 interaction. The protein encoded by the *hcf208* mutant allele, which carries a point mutation in the first of two transmembrane domains, was affected in the AtCCB3- but not AtCCB4-interaction indicating distinct interaction sites.

Summary

The second chloroplast biogenesis factor examined in this PhD thesis is HCF107 that is required for expression of the PSII subunit PsbH. The function of HCF107 was supposed to be processing and/ or stabilization of *psbH* transcripts and their translation. In the present study, the HCF107 protein was found to form large membrane-associated complexes, which contain RNA and presumably also ribosomes, pointing to a function of HCF107 in translation. Formation of HCF107 complexes and the localization of HCF107 inside the chloroplast were found to vary in light and in darkness pointing to a light-dependent regulation of PsbH synthesis. A model of HCF107 in light-dependent *psbH* transcript stabilization, recruitment to chloroplast membranes and co-translational membrane insertion of PsbH was provided.

IV.2 Zusammenfassung

Die Biogenese von Chloroplasten wird durch nuclear-kodierte Faktoren kontrolliert, die im Zytosol synthetisiert und anschliessend in die Plastide importiert werden. In höheren Pflanzen und Grünalgen wurde eine grosse Anzahl von Hilfsfaktoren identifiziert, welche die Expression (Transkription, RNA-Prozessierung, -Edierung, -Stabilität, Translation), Reifung und Assemblierung der photosynthetischen Komplexe regulieren. Da viele dieser Faktoren in Verbindung mit anderen Proteinen und/ oder als Teil eines Proteinkomplexes wirken, kann durch Untersuchungen von Interaktionen und Komplexformation die genaue Funktion dieser Proteine aufgedeckt werden.

In der vorliegenden Arbeit wurde ein Set von binären Vektoren für differentielle Etikettierung von Proteinen in *Arabidopsis thaliana* entwickelt. Verschiedene Mono-, Di- und Triple-Proteinetikette erlauben die Detektion und die flexible Aufreinigung von Proteinen und ihrer Interaktionspartner. Um Überexpression von etikettierten Proteinen zu vermeiden, was sich negativ auf die Aufreinigung von interagierenden Proteinen auswirken würde, wurden Promotorsequenzen von drei Chloroplasten-Biogenesefaktoren (HCF107, HCF136 und HCF173) auserwählt die Expression zu steuern. Die Promotoren wurden histochemisch und quantitativ durch β-Glucoronidase (GUS)-Fusionen analysiert. Verschiedene Promotor/ Etikett-Kombinationen wurden auf ihren Einfluss auf die Funktion der Chloroplasten-Biogenesefaktoren HCF208 und HCF107 getestet. Affinitätsaufreinigungen von HCF208 über einen und mehrere Aufreinigungsschritte unter Anwendung verschiedener Etikette bestätigten die Integrität der klonierten Etikette.

Im zweiten Teil dieser Arbeit wurde durch die Untersuchung einer EMS (Ethylmethan-sulfonat)-induzierten *Arabidopsis*-Mutantenkollektion, die im photosynthetischen Elektronen-transport gestört ist, der Faktor HCF208 identifiziert. Spektroskopische Messungen und immunologische Analysen der *hcf208*-Mutante deckten einen Defekt in der Biogenese des Cytochrom b_6f-Komplexes auf, welcher aus dem Fehlen des kovalent gebundenen Häms c_n an Cytochrom b_6 resultierte. *HCF208* kodiert für das *Arabidopsis*-Homolog des CCB2-Proteins aus *Chlamydomonas reinhardtii*, welches in Verbindung mit mindestens drei anderen Proteinen (CCB1, CCB3 und CCB4) einen neuen Häm-Assemblierungsweg für Cytochrom b_6 bildet, welcher als „System IV“ bezeichnet wird. In der vorliegenden Arbeit analysierte T-DNA-Insertionsmutanten aus *Arabidopsis*, welche in den AtCCB1, AtCCB2 und AtCCB4 Loci betroffen sind, und die *hcf208*-Mutante wiesen vergleichbare Phänotypen auf. HCF208 wurde als integrals Membranprotein identifiziert. Interaktionsanalysen unter zu Hilfenahme von zweidimensionaler “Blue Native”-Gelelektrophorese, Hefe Split Ubiquitin

Analysen und Affinitätsaufreinigungen des etikettiertem HCF208-Proteins, wiesen auf ein stabiles HCF208-AtCCB4-Dimer hin, welches über HCF208 transient mit AtCCB3 interagiert. Das durch das *hcf208*-Mutantenallel kodierte Protein, welches eine Punktmutation in der ersten von zwei Transmembrandomänen aufweist, war in seiner Interaktion mit AtCCB3, aber nicht AtCCB4, eingeschränkt, was auf unterschiedliche Interaktionsdomänen hinwies.

Der zweite im Rahmen dieser Arbeit untersuchte Chloroplasten-Biogenesefaktor ist das Protein HCF107, welches für die Expression der PSII-Untereinheit PsbH benötigt wird. HCF107 wurde eine Funktion in der Prozessierung und/ oder Stabilisierung des *psbH*-Transkripts und seiner Translation zugesprochen. In der vorliegenden Arbeit wurde HCF107 durch Saccharose-Dichtegradienten-Zentrifugationen als Teil eines hochmolekularen, membranassoziierten Komplexes identifiziert, welcher RNA und wahrscheinlich auch Ribosomen enhielt, was auf eine Funktion von HCF107 in der Translation hinwies. Ferner wurde herausgefunden, dass die Bildung von HCF107-Komplexen und die Lokalisierung des Proteins im Chloroplasten sich in Licht und in Dunkelheit unterscheiden, was auf eine lichtabhängige Regulation der PsbH-Synthese hinwies. Ein Modell für die lichtabhängige *psbH*-Transkript-Stabilisierung und -Rekrutierung an Chloroplastenmembranen und die kotranslationale Membraninsertion von PsbH wurde aufgestellt.

V. Literature

Albertsson, P. (2001) A quantitative model of the domain structure of the photosynthetic membrane. *Trends Plant Sci*, **6**, 349-358.

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

Alergand, T., Peled-Zehavi, H., Katz, Y. and Danon, A. (2006) The chloroplast protein disulfide isomerase RB60 reacts with a regulatory disulfide of the RNA-binding protein RB47. *Plant Cell Physiol*, **47**, 540-548.

Allen, J.F. (2003) The function of genomes in bioenergetic organelles. *Philos Trans R Soc Lond B Biol Sci*, **358**, 19-37; discussion 37-18.

Allen, J.W., Ginger, M.L. and Ferguson, S.J. (2004) Complexity and diversity in c-type cytochrome biogenesis systems. *Biochem Soc Trans*, **33**, 145-146.

Allen, J.W., Jackson, A.P., Rigden, D.J., Willis, A.C., Ferguson, S.J. and Ginger, M.L. (2008) Order within a mosaic distribution of mitochondrial c-type cytochrome biogenesis systems? *FEBS J*, **275**, 2385-2402.

Amunts, A., Drory, O. and Nelson, N. (2007) The structure of a plant photosystem I supercomplex at 3.4 A resolution. *Nature*, **447**, 58-63.

Anbudurai, P.R., Mor, T.S., Ohad, I., Shestakov, S.V. and Pakrasi, H.B. (1994) The ctpA gene encodes the C-terminal processing protease for the D1 protein of the photosystem II reaction center complex. *Proc Natl Acad Sci U S A*, **91**, 8082-8086.

Anderson, J.M. (2002) Changing concepts about the distribution of Photosystems I and II between grana-appressed and stroma-exposed thylakoid membranes. *Photosynth Res*, **73**, 157-164.

Armbruster, U., Zuhlke, J., Rengstl, B., Kreller, R., Makarenko, E., Ruhle, T., Schunemann, D., Jahns, P., Weisshaar, B., Nickelsen, J. and Leister, D. (2010) The Arabidopsis thylakoid protein PAM68 is required for efficient D1 biogenesis and photosystem II assembly. *Plant Cell*, **22**, 3439-3460.

Arnon, D.I., Allen, M.B. and Whatley, F.R. (1954) Photosynthesis by isolated chloroplasts. *Nature*, **174**, 394-396.

Aro, E.M., Virgin, I. and Andersson, B. (1993) Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. *Biochim Biophys Acta*, **1143**, 113-134.

Asakura, Y. and Barkan, A. (2006) Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant Physiol*, **142**, 1656-1663.

Auchincloss, A.H., Zerges, W., Perron, K., Girard-Bascou, J. and Rochaix, J.D. (2002) Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in Chlamydomonas reinhardtii. *J Cell Biol*, **157**, 953-962.

Azevedo, J., Courtois, F. and Lerbs-Mache, S. (2006) Sub-plastidial localization of two different phage-type RNA polymerases in spinach chloroplasts. *Nucleic Acids Res*, **34**, 436-444.

Balk, J. and Pilon, M. (2011) Ancient and essential: the assembly of iron-sulfur clusters in plants. *Trends Plant Sci*, **16**, 218-226.

Barkan, A., Walker, M., Nolasco, M. and Johnson, D. (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. *EMBO J*, **13**, 3170-3181.

Barkan, A., Klipcan, L., Ostersetzer, O., Kawamura, T., Asakura, Y. and Watkins, K.P. (2007) The CRM domain: an RNA binding module derived from an ancient ribosome-associated protein. *RNA*, **13**, 55-64.

Barneche, F., Winter, V., Crevecoeur, M. and Rochaix, J.D. (2006) ATAB2 is a novel factor in the signalling pathway of light-controlled synthesis of photosystem proteins. *EMBO J*, **25**, 5907-5918.

Barnes, D., Cohen, A., Bruick, R.K., Kantardjieff, K., Fowler, S., Efuet, E. and Mayfield, S.P. (2004) Identification and characterization of a novel RNA binding protein that associates with the 5'-untranslated region of the chloroplast psbA mRNA. *Biochemistry*, **43**, 8541-8550.

Beick, S., Schmitz-Linneweber, C., Williams-Carrier, R., Jensen, B. and Barkan, A. (2008) The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Mol Cell Biol*, **28**, 5337-5347.

Beligni, M.V., Yamaguchi, K. and Mayfield, S.P. (2004) The translational apparatus of Chlamydomonas reinhardtii chloroplast. *Photosynth Res*, **82**, 315-325.

Bellafiore, S., Ferris, P., Naver, H., Gohre, V. and Rochaix, J.D. (2002) Loss of Albino3 leads to the specific depletion of the light-harvesting system. *Plant Cell*, **14**, 2303-2314.

Bligny, M., Courtois, F., Thaminy, S., Chang, C.C., Lagrange, T., Baruah-Wolff, J., Stern, D. and Lerbs-Mache, S. (2000) Regulation of plastid rDNA transcription by interaction of CDF2 with two different RNA polymerases. *EMBO J*, **19**, 1851-1860.

Bock, R., Hermann, M. and Kossel, H. (1996) In vivo dissection of cis-acting determinants for plastid RNA editing. *EMBO J*, **15**, 5052-5059.

Bogorad, L. (1975) Evolution of organelles and eukaryotic genomes. *Science*, **188**, 891-898.

Bollenbach, T.J., Lange, H., Gutierrez, R., Erhardt, M., Stern, D.B. and Gagliardi, D. (2005) RNR1, a 3'-5' exoribonuclease belonging to the RNR superfamily, catalyzes 3' maturation of chloroplast ribosomal RNAs in Arabidopsis thaliana. *Nucleic Acids Res*, **33**, 2751-2763.

Boudreau, E., Nickelsen, J., Lemaire, S.D., Ossenbuhl, F. and Rochaix, J.D. (2000) The Nac2 gene of Chlamydomonas encodes a chloroplast TPR-like protein involved in psbD mRNA stability. *EMBO J*, **19**, 3366-3376.

Boulouis, A., Raynaud, C., Bujaldon, S., Aznar, A., Wollman, F.A. and Choquet, Y. (2011) The nucleus-encoded trans-acting factor MCA1 plays a critical role in the regulation of cytochrome f synthesis in Chlamydomonas chloroplasts. *Plant Cell*, **23**, 333-349.

Cahoon, A.B., Harris, F.M. and Stern, D.B. (2004) Analysis of developing maize plastids reveals two mRNA stability classes correlating with RNA polymerase type. *EMBO Rep*, **5**, 801-806.

Cai, W., Ji, D., Peng, L., Guo, J., Ma, J., Zou, M., Lu, C. and Zhang, L. (2009) LPA66 is required for editing psbF chloroplast transcripts in Arabidopsis. *Plant Physiol*, **150**, 1260-1271.

Cai, W., Ma, J., Chi, W., Zou, M., Guo, J., Lu, C. and Zhang, L. (2010) Cooperation of LPA3 and LPA2 is essential for photosystem II assembly in Arabidopsis. *Plant Physiol*, **154**, 109-120.

Calvin, M. (1962) The path of carbon in photosynthesis. *Science*, **135**, 879-889.

Carter, M.L., Smith, A.C., Kobayashi, H., Purton, S. and Herrin, D.L. (2004) Structure, circadian regulation and bioinformatic analysis of the unique sigma factor gene in Chlamydomonas reinhardtii. *Photosynth Res*, **82**, 339-349.

Cavalier-Smith, T. (1982) The evolutionary origin and phylogeny of eukaryote flagella. *Symp Soc Exp Biol*, **35**, 465-493.

Chateigner-Boutin, A.L. and Small, I. (2007) A rapid high-throughput method for the detection and quantification of RNA editing based on high-resolution melting of amplicons. *Nucleic Acids Res*, **35**, e114.

Chaudhuri, S. and Maliga, P. (1996) Sequences directing C to U editing of the plastid psbL mRNA are located within a 22 nucleotide segment spanning the editing site. *EMBO J*, **15**, 5958-5964.

Choquet, Y., Stern, D.B., Wostrikoff, K., Kuras, R., Girard-Bascou, J. and Wollman, F.A. (1998) Translation of cytochrome f is autoregulated through the 5' untranslated region of petA mRNA in Chlamydomonas chloroplasts. *Proc Natl Acad Sci U S A*, **95**, 4380-4385.

Choquet, Y. and Vallon, O. (2000) Synthesis, assembly and degradation of thylakoid membrane proteins. *Biochimie*, **82**, 615-634.

Courtois, F., Merendino, L., Demarsy, E., Mache, R. and Lerbs-Mache, S. (2007) Phage-type RNA polymerase RPOTmp transcribes the rrn operon from the PC promoter at early developmental stages in Arabidopsis. *Plant Physiol*, **145**, 712-721.

DalCorso, G., Pesaresi, P., Masiero, S., Aseeva, E., Schunemann, D., Finazzi, G., Joliot, P., Barbato, R. and Leister, D. (2008) A complex containing PGRL1 and PGR5 is involved in the switch between linear and cyclic electron flow in Arabidopsis. *Cell*, **132**, 273-285.

Danon, A. and Mayfield, S.P. (1991) Light regulated translational activators: identification of chloroplast gene specific mRNA binding proteins. *EMBO J*, **10**, 3993-4001.

Danon, A. and Mayfield, S.P. (1994) ADP-dependent phosphorylation regulates RNA-binding in vitro: implications in light-modulated translation. *EMBO J*, **13**, 2227-2235.

de la Sierra-Gallay, I.L., Zig, L., Jamalli, A. and Putzer, H. (2008) Structural insights into the dual activity of RNase J. *Nat Struct Mol Biol*, **15**, 206-212.

de Longevialle, A.F., Hendrickson, L., Taylor, N.L., Delannoy, E., Lurin, C., Badger, M., Millar, A.H. and Small, I. (2008) The pentatricopeptide repeat gene OTP51 with two LAGLIDADG motifs is required for the cis-splicing of plastid ycf3 intron 2 in Arabidopsis thaliana. *Plant J*, **56**, 157-168.

de Longevialle, A.F., Small, I.D. and Lurin, C. (2010) Nuclearly encoded splicing factors implicated in RNA splicing in higher plant organelles. *Mol Plant*, **3**, 691-705.

de Vitry, C., Olive, J., Drapier, D., Recouvreur, M. and Wollman, F.A. (1989) Posttranslational events leading to the assembly of photosystem II protein complex: a study using photosynthesis mutants from Chlamydomonas reinhardtii. *J Cell Biol*, **109**, 991-1006.

de Vitry, C., Desbois, A., Redeker, V., Zito, F. and Wollman, F.A. (2004) Biochemical and spectroscopic characterization of the covalent binding of heme to cytochrome b6. *Biochemistry*, **43**, 3956-3968.

Demarsy, E., Courtois, F., Azevedo, J., Buhot, L. and Lerbs-Mache, S. (2006) Building up of the plastid transcriptional machinery during germination and early plant development. *Plant Physiol*, **142**, 993-1003.

Deng, X.W. and Gruissem, W. (1987) Control of plastid gene expression during development: the limited role of transcriptional regulation. *Cell*, **49**, 379-387.

Dobakova, M., Tichy, M. and Komenda, J. (2007) Role of the PsbI protein in photosystem II assembly and repair in the cyanobacterium Synechocystis sp. PCC 6803. *Plant Physiol*, **145**, 1681-1691.

Drager, R.G., Zeidler, M., Simpson, C.L. and Stern, D.B. (1996) A chloroplast transcript lacking the 3' inverted repeat is degraded by 3'→5' exoribonuclease activity. *RNA*, **2**, 652-663.

Drapier, D., Rimbault, B., Vallon, O., Wollman, F.A. and Choquet, Y. (2007) Intertwined translational regulations set uneven stoichiometry of chloroplast ATP synthase subunits. *EMBO J*, **26**, 3581-3591.

Dreyfuss, B.W., Hamel, P.P., Nakamoto, S.S. and Merchant, S. (2003) Functional analysis of a divergent system II protein, Ccs1, involved in c-type cytochrome biogenesis. *J Biol Chem*, **278**, 2604-2613.

Dumont, M.E., Ernst, J.F., Hampsey, D.M. and Sherman, F. (1987) Identification and sequence of the gene encoding cytochrome c heme lyase in the yeast Saccharomyces cerevisiae. *EMBO J*, **6**, 235-241.

Dyall, S.D., Brown, M.T. and Johnson, P.J. (2004) Ancient invasions: from endosymbionts to organelles. *Science*, **304**, 253-257.

Eickbush, T.H. (1999) Mobile introns: retrohoming by complete reverse splicing. *Curr Biol*, **9**, R11-14.

Fargo, D.C., Zhang, M., Gillham, N.W. and Boynton, J.E. (1998) Shine-Dalgarno-like sequences are not required for translation of chloroplast mRNAs in Chlamydomonas reinhardtii chloroplasts or in Escherichia coli. *Mol Gen Genet*, **257**, 271-282.

Fargo, D.C., Boynton, J.E. and Gillham, N.W. (1999) Mutations altering the predicted secondary structure of a chloroplast 5' untranslated region affect its physical and biochemical properties as well as its ability to promote translation of reporter mRNAs

both in the Chlamydomonas reinhardtii chloroplast and in Escherichia coli. *Mol Cell Biol*, **19**, 6980-6990.

Fedorova, O. and Zingler, N. (2007) Group II introns: structure, folding and splicing mechanism. *Biol Chem*, **388**, 665-678.

Ferreira, K.N., Iverson, T.M., Maghlaoui, K., Barber, J. and Iwata, S. (2004) Architecture of the photosynthetic oxygen-evolving center. *Science*, **303**, 1831-1838.

Finkemeier, I. and Leister, D. (2010) Plant chloroplasts and other plastids. In *Encyclopedia of Life Sciences (ELS) (Chichester: John Wiley & Sons, Ltd).*

Fisk, D.G., Walker, M.B. and Barkan, A. (1999) Molecular cloning of the maize gene crp1 reveals similarity between regulators of mitochondrial and chloroplast gene expression. *EMBO J*, **18**, 2621-2630.

Freyer, R., Kiefer-Meyer, M.C. and Kossel, H. (1997) Occurrence of plastid RNA editing in all major lineages of land plants. *Proc Natl Acad Sci U S A*, **94**, 6285-6290.

Fujita, S., Inagaki, N., Yamamoto, Y., Taguchi, F., Matsumoto, A. and Satoh, K. (1995) Identification of the carboxy-terminal processing protease for the D1 precursor protein of the Photosystem II reaction center of spinach. *Plant Cell Physiol*, **36**, 1169-1177.

Gabilly, S.T., Dreyfuss, B.W., Karamoko, M., Corvest, V., Kropat, J., Page, M.D., Merchant, S.S. and Hamel, P.P. (2010) CCS5, a thioredoxin-like protein involved in the assembly of plastid c-type cytochromes. *J Biol Chem*, **285**, 29738-29749.

Gabilly, S.T., Kropat, J., Karamoko, M., Page, M.D., Nakamoto, S.S., Merchant, S.S. and Hamel, P.P. (2011) A novel component of the disulfide-reducing pathway required for cytochrome c assembly in plastids. *Genetics*, **187**, 793-802.

Giege, P., Grienenberger, J.M. and Bonnard, G. (2008) Cytochrome c biogenesis in mitochondria. *Mitochondrion*, **8**, 61-73.

Gohre, V., Ossenbuhl, F., Crevecoeur, M., Eichacker, L.A. and Rochaix, J.D. (2006) One of two alb3 proteins is essential for the assembly of the photosystems and for cell survival in Chlamydomonas. *Plant Cell*, **18**, 1454-1466.

Goldman, B.S., Beck, D.L., Monika, E.M. and Kranz, R.G. (1998) Transmembrane heme delivery systems. *Proc Natl Acad Sci U S A*, **95**, 5003-5008.

Gould, S.B., Waller, R.F. and McFadden, G.I. (2008) Plastid evolution. *Annu Rev Plant Biol*, **59**, 491-517.

Gray, M.W., Burger, G. and Lang, B.F. (1999) Mitochondrial evolution. *Science*, **283**, 1476-1481.

Guskov, A., Kern, J., Gabdulkhakov, A., Broser, M., Zouni, A. and Saenger, W. (2009) Cyanobacterial photosystem II at 2.9-A resolution and the role of quinones, lipids, channels and chloride. *Nat Struct Mol Biol*, **16**, 334-342.

Hajdukiewicz, P.T., Allison, L.A. and Maliga, P. (1997) The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO J*, **16**, 4041-4048.

Halpin, C., Elderfield, P.D., James, H.E., Zimmermann, R., Dunbar, B. and Robinson, C. (1989) The reaction specificities of the thylakoidal processing peptidase and Escherichia coli leader peptidase are identical. *EMBO J*, **8**, 3917-3921.

Hamel, P.P., Dreyfuss, B.W., Xie, Z., Gabilly, S.T. and Merchant, S. (2003) Essential histidine and tryptophan residues in CcsA, a system II polytopic cytochrome c biogenesis protein. *J Biol Chem*, **278**, 2593-2603.

Hammani, K., Okuda, K., Tanz, S.K., Chateigner-Boutin, A.L., Shikanai, T. and Small, I. (2009) A study of new Arabidopsis chloroplast RNA editing mutants reveals general features of editing factors and their target sites. *Plant Cell*, **21**, 3686-3699.

Harley, C.B. and Reynolds, R.P. (1987) Analysis of E. coli promoter sequences. *Nucleic Acids Res*, **15**, 2343-2361.

Hashimoto, M., Endo, T., Peltier, G., Tasaka, M. and Shikanai, T. (2003) A nucleus-encoded factor, CRR2, is essential for the expression of chloroplast ndhB in Arabidopsis. *Plant J*, **36**, 541-549.

Hattori, M., Miyake, H. and Sugita, M. (2007) A Pentatricopeptide repeat protein is required for RNA processing of clpP Pre-mRNA in moss chloroplasts. *J Biol Chem*, **282**, 10773-10782.

Hattori, M. and Sugita, M. (2009) A moss pentatricopeptide repeat protein binds to the 3' end of plastid clpP pre-mRNA and assists with mRNA maturation. *FEBS J*, **276**, 5860-5869.

Hedtke, B., Borner, T. and Weihe, A. (1997) Mitochondrial and chloroplast phage-type RNA polymerases in Arabidopsis. *Science*, **277**, 809-811.

Hedtke, B., Borner, T. and Weihe, A. (2000) One RNA polymerase serving two genomes. *EMBO Rep*, **1**, 435-440.

Herrmann, R.G., Westhoff, P., Alt, J., Tittgen, J. and Nelson, N. (1985) Thylakoid membrane proteins and their genes. In van Vloten-Doting, L., Groot, G.S.P and Hall, T.C. (EDs) Molecular form and function of the plant genome. *Plenum Publishing Corporation, New York*, 233-256.

Herrmann, R.G., Westhoff, P. and Link, G. (1992) Biogenesis of plastids in higher plants. In Plant Gene Research: Cell Organelles, R.G. Herrmann, ed (Vienna: Springer Verlag), 275-349.

Higgs, D.C., Shapiro, R.S., Kindle, K.L. and Stern, D.B. (1999) Small cis-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. *Mol Cell Biol*, **19**, 8479-8491.

Hirose, T. and Sugiura, M. (1996) Cis-acting elements and trans-acting factors for accurate translation of chloroplast psbA mRNAs: development of an in vitro translation system from tobacco chloroplasts. *EMBO J*, **15**, 1687-1695.

Hirose, T. and Sugiura, M. (1997) Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast ndhD mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. *EMBO J*, **16**, 6804-6811.

Hirose, T. and Sugiura, M. (2004a) Functional Shine-Dalgarno-like sequences for translational initiation of chloroplast mRNAs. *Plant Cell Physiol*, **45**, 114-117.

Hirose, T. and Sugiura, M. (2004b) Multiple elements required for translation of plastid atpB mRNA lacking the Shine-Dalgarno sequence. *Nucleic Acids Res*, **32**, 3503-3510.

Hjorth, E., Hadfi, K, Gould, S.B., Kawach, O., Sommer, M., Zauner, S. and Maier, U.G. (2004) Zero, one, two, three, and perhaps four. Endosymbiosis and the gain and loss of plastids. *Endocytobiol Cell Res*, **15**, 459-468.

Hoch, B., Maier, R.M., Appel, K., Igloi, G.L. and Kossel, H. (1991) Editing of a chloroplast mRNA by creation of an initiation codon. *Nature*, **353**, 178-180.

Houghland, J.L., Piccirilli, J.A., Forconi, M., Jihee, L. and Herschlag, D. (2006) How the group I intron works: a case study of RNA structure and function. In The RNA world. Third edition, ed. RF Gesteland, TR Cech, JF Atkins, 133-205. Cold Spring Harbor, NY, Cold Spring Harbor Lab.

Howe, G. and Merchant, S. (1992) The biosynthesis of membrane and soluble plastidic c-type cytochromes of Chlamydomonas reinhardtii is dependent on multiple common gene products. *EMBO J*, **11**, 2789-2801.

Hutin, C., Nussaume, L., Moise, N., Moya, I., Kloppstech, K. and Havaux, M. (2003) Early light-induced proteins protect Arabidopsis from photooxidative stress. *Proc Natl Acad Sci U S A*, **100**, 4921-4926.

Hynds, P.J., Robinson, D. and Robinson, C. (1998) The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane. *J Biol Chem*, **273**, 34868-34874.

Inoue, K., Dreyfuss, B.W., Kindle, K.L., Stern, D.B., Merchant, S. and Sodeinde, O.A. (1997) Ccs1, a nuclear gene required for the post-translational assembly of chloroplast c-type cytochromes. *J Biol Chem*, **272**, 31747-31754.

Ishihama, A. (1988) Promoter selectivity of prokaryotic RNA polymerases. *Trends Genet*, **4**, 282-286.

Ishihara, S., Takabayashi, A., Ido, K., Endo, T., Ifuku, K. and Sato, F. (2007) Distinct functions for the two PsbP-like proteins PPL1 and PPL2 in the chloroplast thylakoid lumen of Arabidopsis. *Plant Physiol*, **145**, 668-679.

Ishizaki, Y., Tsunoyama, Y., Hatano, K., Ando, K., Kato, K., Shinmyo, A., Kobori, M., Takeba, G., Nakahira, Y. and Shiina, T. (2005) A nuclear-encoded sigma factor,

Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *Plant J*, **42**, 133-144.

Iwai, M., Takizawa, K., Tokutsu, R., Okamuro, A., Takahashi, Y. and Minagawa, J. (2010) Isolation of the elusive supercomplex that drives cyclic electron flow in photosynthesis. *Nature*, **464**, 1210-1213.

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

Jenkins, B.D. and Barkan, A. (2001) Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. *EMBO J*, **20**, 872-879.

Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.A. and Vallon, O. (2010) MRL1, a conserved Pentatricopeptide repeat protein, is required for stabilization of rbcL mRNA in Chlamydomonas and Arabidopsis. *Plant Cell*, **22**, 234-248.

Kaneko, T., Nakamura, Y., Wolk, C.P., Kuritz, T., Sasamoto, S., Watanabe, A., Iriguchi, M., Ishikawa, A., Kawashima, K., Kimura, T., Kishida, Y., Kohara, M., Matsumoto, M., Matsuno, A., Muraki, A., Nakazaki, N., Shimpo, S., Sugimoto, M., Takazawa, M., Yamada, M., Yasuda, M. and Tabata, S. (2001) Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium Anabaena sp. strain PCC 7120. *DNA Res*, **8**, 205-213; 227-253.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M. and Tabata, S. (1996) Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions (supplement). *DNA Res*, **3**, 185-209.

Kato, Y. and Sakamoto, W (2009) Protein quality control in chloroplasts: a current model of D1 protein degradation in the photosystem II repair cycle. *J Biochem*, **146**, 463-469.

Keren, N., Ohkawa, H., Welsh, E.A., Liberton, M. and Pakrasi, H.B. (2005) Psb29, a conserved 22-kD protein, functions in the biogenesis of Photosystem II complexes in Synechocystis and Arabidopsis. *Plant Cell*, **17**, 2768-2781.

Kim, E.H., Chow, W.S., Horton, P. and Anderson, J.M. (2005) Entropy-assisted stacking of thylakoid membranes. *Biochim Biophys Acta*, **1708**, 187-195.

Kim, J., Klein, P.G. and Mullet, J.E. (1994) Vir-115 gene product is required to stabilize D1 translation intermediates in chloroplasts. *Plant Mol Biol*, **25**, 459-467.

Kim, J. and Mayfield, S.P. (1997) Protein disulfide isomerase as a regulator of chloroplast translational activation. *Science*, **278**, 1954-1957.

Kishine, M., Takabayashi, A., Munekage, Y., Shikanai, T., Endo, T. and Sato, F. (2004) Ribosomal RNA processing and an RNase R family member in chloroplasts of Arabidopsis. *Plant Mol Biol*, **55**, 595-606.

Kitajima, A., Asatsuma, S., Okada, H., Hamada, Y., Kaneko, K., Nanjo, Y., Kawagoe, Y., Toyooka, K., Matsuoka, K., Takeuchi, M., Nakano, A. and Mitsui, T. (2009) The rice alpha-amylase glycoprotein is targeted from the Golgi apparatus through the secretory pathway to the plastids. *Plant Cell*, **21**, 2844-2858.

Kleine, T., Voigt, C. and Leister, D. (2009) Plastid signalling to the nucleus: messengers still lost in the mists? *Trends Genet*, **25**, 185-192.

Klinkert, B., Ossenbuhl, F., Sikorski, M., Berry, S., Eichacker, L. and Nickelsen, J. (2004) PratA, a periplasmic tetratricopeptide repeat protein involved in biogenesis of photosystem II in Synechocystis sp. PCC 6803. *J Biol Chem*, **279**, 44639-44644.

Klinkert, B., Elles, I. and Nickelsen, J. (2006) Translation of chloroplast psbD mRNA in Chlamydomonas is controlled by a secondary RNA structure blocking the AUG start codon. *Nucleic Acids Res*, **34**, 386-394.

Klostermann, E., Droste Gen Helling, I., Carde, J.P. and Schunemann, D. (2002) The thylakoid membrane protein ALB3 associates with the cpSecY-translocase in Arabidopsis thaliana. *Biochem J*, **368**, 777-781.

Komenda, J., Reisinger, V., Muller, B.C., Dobakova, M., Granvogl, B. and Eichacker, L.A. (2004) Accumulation of the D2 protein is a key regulatory step for assembly of the photosystem II reaction center complex in Synechocystis PCC 6803. *J Biol Chem*, **279**, 48620-48629.

Kotera, E., Tasaka, M. and Shikanai, T. (2005) A pentatricopeptide repeat protein is essential for RNA editing in chloroplasts. *Nature*, **433**, 326-330.

Kramer, D.M., Avenson, T.J. and Edwards, G.E. (2004) Dynamic flexibility in the light reactions of photosynthesis governed by both electron and proton transfer reactions. *Trends Plant Sci*, **9**, 349-357.

Kranz, R.G., Beckett, C.S. and Goldman, B.S. (2002) Genomic analyses of bacterial respiratory and cytochrome c assembly systems: Bordetella as a model for the system II cytochrome c biogenesis pathway. *Res Microbiol*, **153**, 1-6.

Kranz, R.G., Richard-Fogal, C., Taylor, J.S. and Frawley, E.R. (2009) Cytochrome c biogenesis: mechanisms for covalent modifications and trafficking of heme and for heme-iron redox control. *Microbiol Mol Biol Rev*, **73**, 510-528, Table of Contents.

Kuchka, M.R., Goldschmidt-Clermont, M., van Dillewijn, J. and Rochaix, J.D. (1989) Mutation at the Chlamydomonas nuclear NAC2 locus specifically affects stability of the chloroplast psbD transcript encoding polypeptide D2 of PS II. *Cell*, **58**, 869-876.

Kudla, J., Igloi, G.L., Metzlaff, M., Hagemann, R. and Kossel, H. (1992) RNA editing in tobacco chloroplasts leads to the formation of a translatable psbL mRNA by a C to U substitution within the initiation codon. *EMBO J*, **11**, 1099-1103.

Kugita, M., Yamamoto, Y., Fujikawa, T., Matsumoto, T. and Yoshinaga, K. (2003) RNA editing in hornwort chloroplasts makes more than half the genes functional. *Nucleic Acids Res*, **31**, 2417-2423.

Kuhn, A., Stuart, R., Henry, R. and Dalbey, R.E. (2003) The Alb3/Oxa1/YidC protein family: membrane-localized chaperones facilitating membrane protein insertion? *Trends Cell Biol*, **13**, 510-516.

Kuhn, K., Weihe, A. and Borner, T. (2005) Multiple promoters are a common feature of mitochondrial genes in Arabidopsis. *Nucleic Acids Res*, **33**, 337-346.

Kuras, R., Buschlen, S. and Wollman, F.A. (1995) Maturation of pre-apocytochrome f in vivo. A site-directed mutagenesis study in Chlamydomonas reinhardtii. *J Biol Chem*, **270**, 27797-27803.

Kurisu, G., Zhang, H., Smith, J.L. and Cramer, W.A. (2003) Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science*, **302**, 1009-1014.

Laursen, B.S., Sorensen, H.P., Mortensen, K.K. and Sperling-Petersen, H.U. (2005) Initiation of protein synthesis in bacteria. *Microbiol Mol Biol Rev*, **69**, 101-123.

Leister, D. (2003) Chloroplast research in the genomic age. *Trends Genet*, **19**, 47-56.

Lennartz, K., Plucken, H., Seidler, A., Westhoff, P., Bechtold, N. and Meierhoff, K. (2001) HCF164 encodes a thioredoxin-like protein involved in the biogenesis of the cytochrome b(6)f complex in Arabidopsis. *Plant Cell*, **13**, 2539-2551.

Lennartz, K., Bossmann, S., Westhoff, P., Bechtold, N. and Meierhoff, K. (2006) HCF153, a novel nuclear-encoded factor necessary during a post-translational step in biogenesis of the cytochrome bf complex. *Plant J*, **45**, 101-112.

Lerbs-Mache, S. (1993) The 110-kDa polypeptide of spinach plastid DNA-dependent RNA polymerase: single-subunit enzyme or catalytic core of multimeric enzyme complexes? *Proc Natl Acad Sci U S A*, **90**, 5509-5513.

Lerbs-Mache, S. (2010) Function of plastid sigma factors in higher plants: regulation of gene expression or just preservation of constitutive transcription? *Plant Mol Biol.*, in press.

Lezhneva, L. and Meurer, J. (2004) The nuclear factor HCF145 affects chloroplast psaA-psaB-rps14 transcript abundance in Arabidopsis thaliana. *Plant J*, **38**, 740-753.

Li, H.M. and Chiu, C.C. (2010) Protein transport into chloroplasts. *Annu Rev Plant Biol*, **61**, 157-180.

Liere, K. and Börner, T. (2007) Transcription and transcriptional regulation in plastids. In Bock R (ed) Topics in current genetics, **32**, 121-174.

Lima, A., Lima, S., Wong, J.H., Phillips, R.S., Buchanan, B.B. and Luan, S. (2006) A redox-active FKBP-type immunophilin functions in accumulation of the photosystem II supercomplex in Arabidopsis thaliana. *Proc Natl Acad Sci U S A*, **103**, 12631-12636.

Loiselay, C., Gumpel, N.J., Girard-Bascou, J., Watson, A.T., Purton, S., Wollman, F.A. and Choquet, Y. (2008) Molecular identification and function of cis- and trans-acting determinants for petA transcript stability in Chlamydomonas reinhardtii chloroplasts. *Mol Cell Biol*, **28**, 5529-5542.

Lonetto, M., Gribskov, M. and Gross, C.A. (1992) The sigma 70 family: sequence conservation and evolutionary relationships. *J Bacteriol*, **174**, 3843-3849.

Loschelder, H., Schweer, J., Link, B. and Link, G. (2006) Dual temporal role of plastid sigma factor 6 in Arabidopsis development. *Plant Physiol*, **142**, 642-650.

Luirink, J. and Sinning, I. (2004) SRP-mediated protein targeting: structure and function revisited. *Biochim Biophys Acta*, **1694**, 17-35.

Luirink, J., von Heijne, G., Houben, E. and de Gier, J.W. (2005) Biogenesis of inner membrane proteins in Escherichia coli. *Annu Rev Microbiol*, **59**, 329-355.

Lysenko, E.A. (2003) Analysis of the evolution of the Sig gene family encoding plant sigma factors. *Russ J Plant Physiol*, **53**, 605-614.

Lysenko, E.A. (2007) Plant sigma factors and their role in plastid transcription. *Plant Cell Rep*, **26**, 845-859.

Ma, J., Peng, L., Guo, J., Lu, Q., Lu, C. and Zhang, L. (2007) LPA2 is required for efficient assembly of photosystem II in Arabidopsis thaliana. *Plant Cell*, **19**, 1980-1993.

Maier, U.G., Bozarth, A., Funk, H.T., Zauner, S., Rensing, S.A., Schmitz-Linneweber, C., Borner, T. and Tillich, M. (2008) Complex chloroplast RNA metabolism: just debugging the genetic programme? *BMC Biol*, **6**, 36.

Maliga, P. (1998) Two plastid RNA polymerases of higher plants: an evolving story. *Trends in Plant Science*, **3**, 4-6.

Malnoe, P., Mayfield, S.P. and Rochaix, J.D. (1988) Comparative analysis of the biogenesis of photosystem II in the wild-type and Y-1 mutant of Chlamydomonas reinhardtii. *J Cell Biol*, **106**, 609-616.

Marder, J.B., Goloubinoff, P. and Edelman, M. (1984) Molecular architecture of the rapidly metabolized 32-kilodalton protein of photosystem II. Indications for COOH-terminal processing of a chloroplast membrane polypeptide. *J Biol Chem*, **259**, 3900-3908.

Marin-Navarro, J., Manuell, A.L., Wu, J. and S, P.M. (2007) Chloroplast translation regulation. *Photosynth Res*, **94**, 359-374.

Marques, J.P., Schattat, M.H., Hause, G., Dudeck, I. and Klosgen, R.B. (2004) In vivo transport of folded EGFP by the DeltapH/TAT-dependent pathway in chloroplasts of Arabidopsis thaliana. *J Exp Bot*, **55**, 1697-1706.

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

Martin, W. and Russell, M.J. (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. *Philos Trans R Soc Lond B Biol Sci*, **358**, 59-83; discussion 83-55.

Mathy, N., Benard, L., Pellegrini, O., Daou, R., Wen, T. and Condon, C. (2007) 5'-to-3' exoribonuclease activity in bacteria: role of RNase J1 in rRNA maturation and 5' stability of mRNA. *Cell*, **129**, 681-692.

Mattoo, A.K., Pick, U., Hoffman-Falk, H. and Edelman, M. (1981) The rapidly metabolized 32,000-dalton polypeptide of the chloroplast is the "proteinaceous shield" regulating photosystem II electron transport and mediating diuron herbicide sensitivity. *Proc Natl Acad Sci U S A*, **78**, 1572-1576.

Maul, J.E., Lilly, J.W., Cui, L., dePamphilis, C.W., Miller, W., Harris, E.H. and Stern, D.B. (2002) The Chlamydomonas reinhardtii plastid chromosome: islands of genes in a sea of repeats. *Plant Cell*, **14**, 2659-2679.

Mayfield, S.P., Cohen, A., Danon, A. and Yohn, C.B. (1994) Translation of the psbA mRNA of Chlamydomonas reinhardtii requires a structured RNA element contained within the 5' untranslated region. *J Cell Biol*, **127**, 1537-1545.

Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N. and Schuster, G. (2003) HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. *Plant Cell*, **15**, 1480-1495.

Meierhoff, K. and Westhoff, P. (2011) Biogenesis of the thylakoid membrane – PSII a case study. In Chloroplast development during leaf growth and senescence (Advanced photosynthesis and respiration). In press.

Meurer, J., Plucken, H., Kowallik, K.V. and Westhoff, P. (1998) A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in Arabidopsis thaliana. *EMBO J*, **17**, 5286-5297.

Meurer, J., Lezhneva, L., Amann, K., Godel, M., Bezhani, S., Sherameti, I. and Oelmuller, R. (2002) A peptide chain release factor 2 affects the stability of UGA-containing transcripts in Arabidopsis chloroplasts. *Plant Cell*, **14**, 3255-3269.

Michel, F., Umesono, K. and Ozeki, H. (1989) Comparative and functional anatomy of group II catalytic introns--a review. *Gene*, **82**, 5-30.

Minai, L., Wostrikoff, K., Wollman, F.A. and Choquet, Y. (2006) Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. *Plant Cell*, **18**, 159-175.

Miras, S., Salvi, D., Ferro, M., Grunwald, D., Garin, J., Joyard, J. and Rolland, N. (2002) Non-canonical transit peptide for import into the chloroplast. *J Biol Chem*, **277**, 47770-47778.

Miras, S., Salvi, D., Piette, L., Seigneurin-Berny, D., Grunwald, D., Reinbothe, C., Joyard, J., Reinbothe, S. and Rolland, N. (2007) Toc159- and Toc75-independent import of a transit sequence-less precursor into the inner envelope of chloroplasts. *J Biol Chem*, **282**, 29482-29492.

Mitra, K., Frank, J. and Driessen, A. (2006) Co- and post-translational translocation through the protein-conducting channel: analogous mechanisms at work? *Nat Struct Mol Biol*, **13**, 957-964.

Monod, C., Goldschmidt-Clermont, M. and Rochaix, J.D. (1992) Accumulation of chloroplast psbB RNA requires a nuclear factor in Chlamydomonas reinhardtii. *Mol Gen Genet*, **231**, 449-459.

Moore, M., Harrison, M.S., Peterson, E.C. and Henry, R. (2000) Chloroplast Oxa1p homolog albino3 is required for post-translational integration of the light harvesting chlorophyll-binding protein into thylakoid membranes. *J Biol Chem*, **275**, 1529-1532.

Muhlbauer, S.K. and Eichacker, L.A. (1998) Light-dependent formation of the photosynthetic proton gradient regulates translation elongation in chloroplasts. *J Biol Chem*, **273**, 20935-20940.

Muller, M. and Klosgen, R.B. (2005) The Tat pathway in bacteria and chloroplasts (review). *Mol Membr Biol*, **22**, 113-121.

Mullet, J.E., Klein, P.G. and Klein, R.R. (1990) Chlorophyll regulates accumulation of the plastid-encoded chlorophyll apoproteins CP43 and D1 by increasing apoprotein stability. *Proc Natl Acad Sci U S A*, **87**, 4038-4042.

Mullet, J.E. (1993) Dynamic regulation of chloroplast transcription. *Plant Physiol*, **103**, 309-313.

Mulo, P., Sirpio, S., Suorsa, M. and Aro, E.M. (2008) Auxiliary proteins involved in the assembly and sustenance of photosystem II. *Photosynth Res*, **98**, 489-501.

Munekage, Y. and Shikanai, T. (2005) Cyclic electron transport through photosystem I. *Plant Biotechnol*, **22**, 361-369.

Murakami, S., Kuehnle, K. and Stern, D.B. (2005) A spontaneous tRNA suppressor of a mutation in the Chlamydomonas reinhardtii nuclear MCD1 gene required for stability of the chloroplast petD mRNA. *Nucleic Acids Res*, **33**, 3372-3380.

Nada, A. and Soll, J. (2004) Inner envelope protein 32 is imported into chloroplasts by a novel pathway. *J Cell Sci*, **117**, 3975-3982.

Nakamura, T., Meierhoff, K., Westhoff, P. and Schuster, G. (2003) RNA-binding properties of HCF152, an Arabidopsis PPR protein involved in the processing of chloroplast RNA. *Eur J Biochem*, **270**, 4070-4081.

Nakamura, Y. and Ito, K. (1998) How protein reads the stop codon and terminates translation. *Genes Cells*, **3**, 265-278.

Nanjo, Y., Oka, H., Ikarashi, N., Kaneko, K., Kitajima, A., Mitsui, T., Munoz, F.J., Rodriguez-Lopez, M., Baroja-Fernandez, E. and Pozueta-Romero, J. (2006) Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-golgi to the chloroplast through the secretory pathway. *Plant Cell*, **18**, 2582-2592.

Neckermann, K., Zeltz, P., Igloi, G.L., Kossel, H. and Maier, R.M. (1994) The role of RNA editing in conservation of start codons in chloroplast genomes. *Gene*, **146**, 177-182.

Nelson, N. and Ben-Shem, A. (2004) The complex architecture of oxygenic photosynthesis. *Nat Rev Mol Cell Biol*, **5**, 971-982.

Nelson, N. and Yocum, C.F. (2006) Structure and function of photosystems I and II. *Annu Rev Plant Biol*, **57**, 521-565.

Neuhaus, H.E. and Emes, M.J. (2000) Nonphotosynthetic Metabolism in Plastids. *Annu Rev Plant Physiol Plant Mol Biol*, **51**, 111-140.

Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M. and Rochaix, J.D. (1999) Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in Chlamydomonas. *Plant Cell*, **11**, 957-970.

Nixon, P.J., Trost, J.T. and Diner, B.A. (1992) Role of the carboxy terminus of polypeptide D1 in the assembly of a functional water-oxidizing manganese cluster in photosystem II of the cyanobacterium Synechocystis sp. PCC 6803: assembly requires a free carboxyl group at C-terminal position 344. *Biochemistry*, **31**, 10859-10871.

Ohad, I., Adir, N., Koike, H., Kyle, D.J. and Inoue, Y. (1990) Mechanism of photoinhibition in vivo. A reversible light-induced conformational change of reaction center II is related to an irreversible modification of the D1 protein. *J Biol Chem*, **265**, 1972-1979.

Okuda, K., Nakamura, T., Sugita, M., Shimizu, T. and Shikanai, T. (2006) A pentatricopeptide repeat protein is a site recognition factor in chloroplast RNA editing. *J Biol Chem*, **281**, 37661-37667.

Okuda, K., Myouga, F., Motohashi, R., Shinozaki, K. and Shikanai, T. (2007) Conserved domain structure of pentatricopeptide repeat proteins involved in chloroplast RNA editing. *Proc Natl Acad Sci U S A*, **104**, 8178-8183.

Okuda, K., Chateigner-Boutin, A.L., Nakamura, T., Delannoy, E., Sugita, M., Myouga, F., Motohashi, R., Shinozaki, K., Small, I. and Shikanai, T. (2009) Pentatricopeptide repeat proteins with the DYW motif have distinct molecular functions in RNA editing and RNA cleavage in Arabidopsis chloroplasts. *Plant Cell*, **21**, 146-156.

Onda, Y., Yagi, Y., Saito, Y., Takenaka, N. and Toyoshima, Y. (2008) Light induction of Arabidopsis SIG1 and SIG5 transcripts in mature leaves: differential roles of cryptochrome 1 and cryptochrome 2 and dual function of SIG5 in the recognition of plastid promoters. *Plant J*, **55**, 968-978.

Ossenbuhl, F. and Nickelsen, J. (2000) cis- and trans-Acting determinants for translation of psbD mRNA in Chlamydomonas reinhardtii. *Mol Cell Biol*, **20**, 8134-8142.

Ostheimer, G.J., Williams-Carrier, R., Belcher, S., Osborne, E., Gierke, J. and Barkan, A. (2003) Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO J*, **22**, 3919-3929.

Page, M.L., Hamel, P.P., Gabilly, S.T., Zegzouti, H., Perea, J.V., Alonso, J.M., Ecker, J.R., Theg, S.M., Christensen, S.K. and Merchant, S. (2004) A homolog of prokaryotic thiol disulfide transporter CcdA is required for the assembly of the cytochrome b6f complex in Arabidopsis chloroplasts. *J Biol Chem*, **279**, 32474-32482.

Pasch, J.C., Nickelsen, J. and Schunemann, D. (2005) The yeast split-ubiquitin system to study chloroplast membrane protein interactions. *Appl Microbiol Biotechnol*, **69**, 440-447.

Peng, L., Ma, J., Chi, W., Guo, J., Zhu, S., Lu, Q., Lu, C. and Zhang, L. (2006) LOW PSII ACCUMULATION1 is involved in efficient assembly of photosystem II in Arabidopsis thaliana. *Plant Cell*, **18**, 955-969.

Pfalz, J., Bayraktar, O.A., Prikryl, J. and Barkan, A. (2009) Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *EMBO J*, **28**, 2042-2052.

Pfannschmidt, T. (2010) Plastidial retrograde signalling--a true "plastid factor" or just metabolite signatures? *Trends Plant Sci*, **15**, 427-435.

Plader, W. and Sugiura, M. (2003) The Shine-Dalgarno-like sequence is a negative regulatory element for translation of tobacco chloroplast rps2 mRNA: an additional mechanism for translational control in chloroplasts. *Plant J*, **34**, 377-382.

Plucken, H., Muller, B., Grohmann, D., Westhoff, P. and Eichacker, L.A. (2002) The HCF136 protein is essential for assembly of the photosystem II reaction center in Arabidopsis thaliana. *FEBS Lett*, **532**, 85-90.

Pogson, B.J., Woo, N.S., Forster, B. and Small, I.D. (2008) Plastid signalling to the nucleus and beyond. *Trends Plant Sci*, **13**, 602-609.

Prikryl, J., Rojas, M., Schuster, G. and Barkan, A. (2010) Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proc Natl Acad Sci U S A*, **108**, 415-420.

Pyle, A.M. and Lambowitz (2006) Group II introns: ribozymes that splice RNA and invade DANN. In The RNA world. Third edition, ed. RF Gesteland, TR Cech, JF Atkins, 469-505. Cold Spring Harbor, NY: Cold Spring Harbor Lab.

Race, H.L., Herrmann, R.G. and Martin, W. (1999) Why have organelles retained genomes? *Trends Genet*, **15**, 364-370.

Radhamony, R.N. and Theg, S.M. (2006) Evidence for an ER to Golgi to chloroplast protein transport pathway. *Trends Cell Biol*, **16**, 385-387.

Rattanachaikunsopon, P., Rosch, C. and Kuchka, M.R. (1999) Cloning and characterization of the nuclear AC115 gene of Chlamydomonas reinhardtii. *Plant Mol Biol*, **39**, 1-10.

Raynaud, C., Loiselay, C., Wostrikoff, K., Kuras, R., Girard-Bascou, J., Wollman, F.A. and Choquet, Y. (2007) Evidence for regulatory function of nucleus-encoded factors on mRNA stabilization and translation in the chloroplast. *Proc Natl Acad Sci U S A*, **104**, 9093-9098.

Richly, E. and 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.

Richter, S., Zhong, R. and Lamppa, G. (2005) Function of the stromal processing peptidase in the chloroplast import. *Physiol Plant*, **123**, 362-368.

Robertson, D.E., Farid, R.S., Moser, C.C., Urbauer, J.L., Mulholland, S.E., Pidikiti, R., Lear, J.D., Wand, A.J., DeGrado, W.F. and Dutton, P.L. (1994) Design and synthesis of multi-haem proteins. *Nature*, **368**, 425-432.

Robinson, C. and Mant, A. (1997) Targeting of proteins into and accross the thylakoid membrane. *Trends Plant Sci*, **2**, 431-436.

Robson, A. and Collinson, I. (2006) The structure of the Sec complex and the problem of protein translocation. *EMBO Rep*, **7**, 1099-1103.

Rohl, T. and van Wijk, K.J. (2001) In vitro reconstitution of insertion and processing of cytochrome f in a homologous chloroplast translation system. *J Biol Chem*, **276**, 35465-35472.

Rokka, A., Suorsa, M., Saleem, A., Battchikova, N. and Aro, E.M. (2005) Synthesis and assembly of thylakoid protein complexes: multiple assembly steps of photosystem II. *Biochem J*, **388**, 159-168.

Roose, J.L. and Pakrasi, H.B. (2004) Evidence that D1 processing is required for manganese binding and extrinsic protein assembly into photosystem II. *J Biol Chem*, **279**, 45417-45422.

Rurek, M. (2008) Proteins involved in maturation pathways of plant mitochondrial and plastid c-type cytochromes. *Acta Biochim Pol*, **55**, 417-433.

Sakamoto, W., Chen, X., Kindle, K.L. and Stern, D.B. (1994) Function of the Chlamydomonas reinhardtii petd 5' untranslated region in regulating the accumulation of subunit IV of the cytochrome b6/f complex. *Plant J*, **6**, 503-512.

Sanders, C., Turkarslan, S., Lee, D.W. and Daldal, F. (2010) Cytochrome c biogenesis: the Ccm system. *Trends Microbiol*, **18**, 266-274.

Sasaki, T., Yukawa, Y., Wakasugi, T., Yamada, K. and Sugiura, M. (2006) A simple in vitro RNA editing assay for chloroplast transcripts using fluorescent dideoxynucleotides: distinct types of sequence elements required for editing of ndh transcripts. *Plant J*, **47**, 802-810.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E. and Tabata, S. (1999) Complete structure of the chloroplast genome of Arabidopsis thaliana. *DNA Res*, **6**, 283-290.

Schein, A., Sheffy-Levin, S., Glaser, F. and Schuster, G. (2008) The RNase E/G-type endoribonuclease of higher plants is located in the chloroplast and cleaves RNA similarly to the E. coli enzyme. *RNA*, **14**, 1057-1068.

Schmitz-Linneweber, C., Williams-Carrier, R. and Barkan, A. (2005) RNA immunoprecipitation and microarray analysis show a chloroplast Pentatricopeptide repeat protein to be associated with the 5' region of mRNAs whose translation it activates. *Plant Cell*, **17**, 2791-2804.

Schmitz-Linneweber, C., Williams-Carrier, R.E., Williams-Voelker, P.M., Kroeger, T.S., Vichas, A. and Barkan, A. (2006) A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. *Plant Cell*, **18**, 2650-2663.

Schmitz-Linneweber, C. and Barkan, A. (2007) RNA splicing and RNA editing in chloroplasts. In Cell and Molecular Biology of Plastids, ed. R Bock, 213-248. Heidelberg, Germany: Springer.

Schmitz-Linneweber, C. and Small, I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci*, **13**, 663-670.

Schottler, M.A., Albus, C.A. and Bock, R. (2011) Photosystem I: Its biogenesis and function in higher plants. *J Plant Physiol*.

Schult, K., Meierhoff, K., Paradies, S., Toller, T., Wolff, P. and Westhoff, P. (2007) The nuclear-encoded factor HCF173 is involved in the initiation of translation of the psbA mRNA in Arabidopsis thaliana. *Plant Cell*, **19**, 1329-1346.

Schunemann, D. (2004) Structure and function of the chloroplast signal recognition particle. *Curr Genet*, **44**, 295-304.

Schunemann, D. (2007) Mechanisms of protein import into thylakoids of chloroplasts. *Biol Chem*, **388**, 907-915.

Schuster, G. and Stern, D. (2009) RNA polyadenylation and decay in mitochondria and chloroplasts. *Prog Mol Biol Transl Sci*, **85**, 393-422.

Schwarz, C., Elles, I., Kortmann, J., Piotrowski, M. and Nickelsen, J. (2007) Synthesis of the D2 protein of photosystem II in Chlamydomonas is controlled by a high molecular mass complex containing the RNA stabilization factor Nac2 and the translational activator RBP40. *Plant Cell*, **19**, 3627-3639.

Schwenkert, S., Netz, D.J., Frazzon, J., Pierik, A.J., Bill, E., Gross, J., Lill, R. and Meurer, J. (2010) Chloroplast HCF101 is a scaffold protein for [4Fe-4S] cluster assembly. *Biochem J*, **425**, 207-214.

Scolnick, E., Tompkins, R., Caskey, T. and Nirenberg, M. (1968) Release factors differing in specificity for terminator codons. *Proc Natl Acad Sci U S A*, **61**, 768-774.

Seelert, H., Poetsch, A., Dencher, N.A., Engel, A., Stahlberg, H. and Muller, D.J. (2000) Structural biology. Proton-powered turbine of a plant motor. *Nature*, **405**, 418-419.

Shen, G., Antonkine, M.L., van der Est, A., Vassiliev, I.R., Brettel, K., Bittl, R., Zech, S.G., Zhao, J., Stehlik, D., Bryant, D.A. and Golbeck, J.H. (2002) Assembly of photosystem I. II. Rubredoxin is required for the in vivo assembly of F(X) in Synechococcus sp. PCC 7002 as shown by optical and EPR spectroscopy. *J Biol Chem*, **277**, 20355-20366.

Shiina, T., Tsunoyama, Y., Nakahira, Y. and Khan, M.S. (2005) Plastid RNA polymerases, promoters, and transcription regulators in higher plants. *Int Rev Cytol*, **244**, 1-68.

Shikanai, T. (2007) Cyclic electron transport around photosystem I: genetic approaches. *Annu Rev Plant Biol*, **58**, 199-217.

Shirano, Y., Shimada, H., Kanamaru, K., Fujiwara, M., Tanaka, K., Takahashi, H., Unno, K., Sato, S., Tabata, S., Hayashi, H., Miyake, C., Yokota, A. and Shibata, D. (2000) Chloroplast development in Arabidopsis thaliana requires the nuclear-encoded transcription factor sigma B. *FEBS Lett*, **485**, 178-182.

Sirpio, S., Allahverdiyeva, Y., Suorsa, M., Paakkarinen, V., Vainonen, J., Battchikova, N. and Aro, E.M. (2007) TLP18.3, a novel thylakoid lumen protein regulating photosystem II repair cycle. *Biochem J*, **406**, 415-425.

Somanchi, A., Barnes, D. and Mayfield, S.P. (2005) A nuclear gene of Chlamydomonas reinhardtii, Tba1, encodes a putative oxidoreductase required for translation of the chloroplast psbA mRNA. *Plant J*, **42**, 341-352.

Stern, D.B., Goldschmidt-Clermont, M. and Hanson, M.R. (2010) Chloroplast RNA metabolism. *Annu Rev Plant Biol*, **61**, 125-155.

Stern, D.B. and Gruissem, W. (1987) Control of plastid gene expression: 3' inverted repeats act as mRNA processing and stabilizing elements, but do not terminate transcription. *Cell*, **51**, 1145-1157.

Stockel, J., Bennewitz, S., Hein, P. and Oelmuller, R. (2006) The evolutionarily conserved tetratrico peptide repeat protein pale yellow green7 is required for photosystem I accumulation in Arabidopsis and copurifies with the complex. *Plant Physiol*, **141**, 870-878.

Stroebel, D., Choquet, Y., Popot, J.L. and Picot, D. (2003) An atypical haem in the cytochrome b(6)f complex. *Nature*, **426**, 413-418.

Sturm, N.R., Kuras, R., Buschlen, S., Sakamoto, W., Kindle, K.L., Stern, D.B. and Wollman, F.A. (1994) The petD gene is transcribed by functionally redundant promoters in Chlamydomonas reinhardtii chloroplasts. *Mol Cell Biol*, **14**, 6171-6179.

Suay, L., Salvador, M.L., Abesha, E. and Klein, U. (2005) Specific roles of 5' RNA secondary structures in stabilizing transcripts in chloroplasts. *Nucleic Acids Res*, **33**, 4754-4761.

Subramanian, A.R. (1993) Molecular genetics of chloroplast ribosomal proteins. *Trends Biochem Sci*, **18**, 177-181.

Sugita, M., Miyata, Y., Maruyama, K., Sugiura, C., Arikawa, T. and Higuchi, M. (2006) Extensive RNA editing in transcripts from the PsbB operon and RpoA gene of plastids from the enigmatic moss Takakia lepidozioides. *Biosci Biotechnol Biochem*, **70**, 2268-2274.

Sugiura, M. (1992) The chloroplast genome. *Plant Mol Biol*, **19**, 149-168.

Sugiura, M., Hirose, T. and Sugita, M. (1998) Evolution and mechanism of translation in chloroplasts. *Annu Rev Genet*, **32**, 437-459.

Swiatecka-Hagenbruch, M., Emanuel, C., Hedtke, B., Liere, K. and Borner, T. (2008) Impaired function of the phage-type RNA polymerase RpoTp in transcription of chloroplast genes is compensated by a second phage-type RNA polymerase. *Nucleic Acids Res*, **36**, 785-792.

Swiatecka-Hagenbruch, M., Liere, K. and Borner, T. (2007) High diversity of plastidial promoters in Arabidopsis thaliana. *Mol Genet Genomics*, **277**, 725-734.

Takahashi, M., Shiraishi, T. and Asada, K. (1988) COOH-terminal residues of D1 and the 44 kDa CPa-2 at spinach photosystem II core complex. *FEBS Lett*, **240**, 6-8.

Tillich, M., Funk, H.T., Schmitz-Linneweber, C., Poltnigg, P., Sabater, B., Martin, M. and Maier, R.M. (2005) Editing of plastid RNA in Arabidopsis thaliana ecotypes. *Plant J*, **43**, 708-715.

Tillich, M., Hardel, S.L., Kupsch, C., Armbruster, U., Delannoy, E., Gualberto, J.M., Lehwark, P., Leister, D., Small, I.D. and Schmitz-Linneweber, C. (2009) Chloroplast ribonucleoprotein CP31A is required for editing and stability of specific chloroplast mRNAs. *Proc Natl Acad Sci U S A*, **106**, 6002-6007.

Timmis, J.N., Ayliffe, M.A., Huang, C.Y. and Martin, W. (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet*, **5**, 123-135.

Turmel, M., Gutell, R.R., Mercier, J.P., Otis, C. and Lemieux, C. (1993) Analysis of the chloroplast large subunit ribosomal RNA gene from 17 Chlamydomonas taxa. Three internal transcribed spacers and 12 group I intron insertion sites. *J Mol Biol*, **232**, 446-467.

Vaistij, F.E., Boudreau, E., Lemaire, S.D., Goldschmidt-Clermont, M. and Rochaix, J.D. (2000a) Characterization of Mbb1, a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast psbB/psbT/psbH gene cluster in Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A*, **97**, 14813-14818.

Vaistij, F.E., Goldschmidt-Clermont, M., Wostrikoff, K. and Rochaix, J.D. (2000b) Stability determinants in the chloroplast psbB/T/H mRNAs of Chlamydomonas reinhardtii. *Plant J*, **21**, 469-482.

Viitanen, P.V., Doran, E.R. and Dunsmuir, P. (1988) What is the role of the transit peptide in thylakoid integration of the light-harvesting chlorophyll a/b protein? *J Biol Chem*, **263**, 15000-15007.

Villarejo, A., Buren, S., Larsson, S., Dejardin, A., Monne, M., Rudhe, C., Karlsson, J., Jansson, S., Lerouge, P., Rolland, N., von Heijne, G., Grebe, M., Bako, L. and Samuelsson, G. (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat Cell Biol*, **7**, 1224-1231.

Wank, H., SanFilippo, J., Singh, R.N., Matsuura, M. and Lambowitz, A.M. (1999) A reverse transcriptase/maturase promotes splicing by binding at its own coding segment in a group II intron RNA. *Mol Cell*, **4**, 239-250.

Waters, M.T. and Langdale, J.A. (2009) The making of a chloroplast. *EMBO J*, **28**, 2861-2873.

Wei, L., Guo, J., Ouyang, M., Sun, X., Ma, J., Chi, W., Lu, C. and Zhang, L. (2010) LPA19, a Psb27 homolog in Arabidopsis thaliana, facilitates D1 protein precursor processing during PSII biogenesis. *J Biol Chem*, **285**, 21391-21398.

Wolf, P.G., Rowe, C.A. and Hasebe, M. (2004) High levels of RNA editing in a vascular plant chloroplast genome: analysis of transcripts from the fern Adiantum capillus-veneris. *Gene*, **339**, 89-97.

Wollman, F.A., Minai, L. and Nechushtai, R. (1999) The biogenesis and assembly of photosynthetic proteins in thylakoid membranes1. *Biochim Biophys Acta*, **1411**, 21-85.

Wolosiuk, R.A., Ballicora, M.A. and Hagelin, K. (1993) The reductive pentose phosphate cycle for photosynthetic CO2 assimilation: enzyme modulation. *FASEB J*, **7**, 622-637.

Woodson, J.D. and Chory, J. (2008) Coordination of gene expression between organellar and nuclear genomes. *Nat Rev Genet*, **9**, 383-395.

Wostrikoff, K., Girard-Bascou, J., Wollman, F.A. and Choquet, Y. (2004) Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of Chlamydomonas. *EMBO J*, **23**, 2696-2705.

Xie, Z., Culler, D., Dreyfuss, B.W., Kuras, R., Wollman, F.A., Girard-Bascou, J. and Merchant, S. (1998) Genetic analysis of chloroplast c-type cytochrome assembly in Chlamydomonas reinhardtii: One chloroplast locus and at least four nuclear loci are required for heme attachment. *Genetics*, **148**, 681-692.

Xie, Z. and Merchant, S. (1996) The plastid-encoded ccsA gene is required for heme attachment to chloroplast c-type cytochromes. *J Biol Chem*, **271**, 4632-4639.

Xie, Z. and Merchant, S. (1998) A novel pathway for cytochromes c biogenesis in chloroplasts. *Biochim Biophys Acta*, **1365**, 309-318.

Yamaguchi, K., Prieto, S., Beligni, M.V., Haynes, P.A., McDonald, W.H., Yates, J.R., 3rd and Mayfield, S.P. (2002) Proteomic characterization of the small subunit of Chlamydomonas reinhardtii chloroplast ribosome: identification of a novel S1 domain-containing protein and unusually large orthologs of bacterial S2, S3, and S5. *Plant Cell*, **14**, 2957-2974.

Yamaguchi, K., Beligni, M.V., Prieto, S., Haynes, P.A., McDonald, W.H., Yates, J.R., 3rd and Mayfield, S.P. (2003) Proteomic characterization of the Chlamydomonas reinhardtii chloroplast ribosome. Identification of proteins unique to th e70 S ribosome. *J Biol Chem*, **278**, 33774-33785.

Yehudai-Resheff, S., Hirsh, M. and Schuster, G. (2001) Polynucleotide phosphorylase functions as both an exonuclease and a poly(A) polymerase in spinach chloroplasts. *Mol Cell Biol*, **21**, 5408-5416.

Yohn, C.B., Cohen, A., Danon, A. and Mayfield, S.P. (1998) A poly(A) binding protein functions in the chloroplast as a message-specific translation factor. *Proc Natl Acad Sci U S A*, **95**, 2238-2243.

Yusupova, G.Z., Yusupov, M.M., Cate, J.H. and Noller, H.F. (2001) The path of messenger RNA through the ribosome. *Cell*, **106**, 233-241.

Zerges, W., Girard-Bascou, J. and Rochaix, J.D. (1997) Translation of the chloroplast psbC mRNA is controlled by interactions between its 5' leader and the nuclear loci TBC1 and TBC3 in Chlamydomonas reinhardtii. *Mol Cell Biol*, **17**, 3440-3448.

Zerges, W. (2000) Translation in chloroplasts. *Biochimie*, **82**, 583-601.

Zerges, W., Auchincloss, A.H. and Rochaix, J.D. (2003) Multiple translational control sequences in the 5' leader of the chloroplast psbC mRNA interact with nuclear gene products in Chlamydomonas reinhardtii. *Genetics*, **163**, 895-904.

Zhang, L., Paakkarinen, V., Suorsa, M. and Aro, E.M. (2001) A SecY homologue is involved in chloroplast-encoded D1 protein biogenesis. *J Biol Chem*, **276**, 37809-37814.

Zhang, L., Paakkarinen, V., van Wijk, K.J. and Aro, E.M. (2000) Biogenesis of the chloroplast-encoded D1 protein: regulation of translation elongation, insertion, and assembly into photosystem II. *Plant Cell*, **12**, 1769-1782.

Zhou, W., Cheng, Y., Yap, A., Chateigner-Boutin, A.L., Delannoy, E., Hammani, K., Small, I. and Huang, J. (2008) The Arabidopsis gene YS1 encoding a DYW protein is required for editing of rpoB transcripts and the rapid development of chloroplasts during early growth. *Plant J*.

Zito, F., Kuras, R., Choquet, Y., Kossel, H. and Wollman, F.A. (1997) Mutations of cytochrome b6 in Chlamydomonas reinhardtii disclose the functional significance for a proline to leucine conversion by petB editing in maize and tobacco. *Plant Mol Biol*, **33**, 79-86.

Zoschke, R., Liere, K. and Borner, T. (2007) From seedling to mature plant: arabidopsis plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant J*, **50**, 710-722.

Zoschke, R., Nakamura, M., Liere, K., Sugiura, M., Borner, T. and Schmitz-Linneweber, C. (2010) An organellar maturase associates with multiple group II introns. *Proc Natl Acad Sci U S A*, **107**, 3245-3250.

Zou, Z., Eibl, C. and Koop, H.U. (2003) The stem-loop region of the tobacco psbA 5'UTR is an important determinant of mRNA stability and translation efficiency. *Mol Genet Genomics*, **269**, 340-349.

VI. Manuscripts

Part A:

1) Dagmar Lyska, Kerstin Schult, Karin Meierhoff and Peter Westhoff (2011). **pAUL: A Gateway-based vector system for adaptive expression and flexible tagging of proteins in *Arabidopsis*.** Submitted to Journal of Experimental Botany for publication.

Part B:

2) Dagmar Lyska, Susanne Paradies, Karin Meierhoff and Peter Westhoff (2007). **HCF208, a homolog of *Chlamydomonas* CCB2, is required for accumulation of native cytochrome b_6 in *Arabidopsis thaliana*.** *Plant Cell Physiol,* **48**, 1737-1746.

3) **Molecular characterization of HCF208 localization and interactions.**

Part C:

4) **Analyses on protein function and complex formation of the PsbH synthesis factor HCF107.**

- Part A -

Manuscript 1

pAUL: A Gateway-based vector system for adaptive expression and flexible tagging of proteins in *Arabidopsis*

Running title

Vectors for differential expression and tagging

pAUL: A Gateway-based vector system for adaptive expression and flexible tagging of proteins in Arabidopsis

Authors

Dagmar Lyska*, Kerstin Schult, Karin Meierhoff, Peter Westhoff

Institut fuer Entwicklungs- und Molekularbiologie der Pflanzen, Heinrich-Heine-Universitaet, 40225 Duesseldorf, Germany

*Corresponding author:

Dagmar Lyska

Heinrich-Heine-Universitaet Duesseldorf

Institut fuer Entwicklungs- und Molekularbiologie der Pflanzen

Universitaetsstr. 1

D- 40225 Duesseldorf

Phone: +49 (0)211 8113428

Fax: +49 (0)211 8114871

E-mail: dagmar.lyska@uni-duesseldorf.de

Keywords: Gateway, adaptive expression, fusion protein, epitope tag, affinity purification, HCF

Summary

Determination of protein function requires tools that allow its detection and/ or purification. As generation of specific antibodies often is laborious and insufficient, protein tagging using epitopes that are recognized by commercially available antibodies and matrices appears more promising. Also, proper spatial and temporal expression of tagged proteins is required to prevent falsification of results. We developed a new series of binary Gateway cloning vectors named pAUL1-20 for C- and N-terminal in-frame fusion of proteins to four different tags: a single (i) HA epitope and (ii) Strep-tag*III*, (iii) both epitopes combined to a double tag, and (iv) a triple tag consisting of the double tag extended by a Protein A tag possessing a 3C protease cleavage site. Expression can be driven by either the 35S *CaMV* promoter or, for C-terminal fusions, promoters from genes encoding the chloroplast biogenesis factors HCF107, HCF136, or HCF173. Fusions of the four promoters to the GUS gene showed that endogenous promoter sequences are functional and drive expression more moderately and consistent throughout different transgenic lines when compared to the 35S *CaMV* promoter. By testing complementation of mutations affected in chloroplast biogenesis factors HCF107 and HCF208, we found that the effect of different promoters and tags on protein function strongly depends on the protein itself. Single-step and tandem affinity purification of HCF208 via different tags confirmed the integrity of the cloned tags.

Introduction

The majority of cellular processes is accomplished and regulated by proteins. To shed light on the precise function of a protein, tools for detection and/or determination of subcellular localization are required. Also, identification and characterization of interaction partners is of great importance as most proteins act in collaboration with other proteins either transiently or in stable complexes. To address all these questions diverse protein tagging strategies have been invented throughout the past years. In-frame translational fusions of the protein of interest and either a reporter protein (e.g. GFP; Pang et al., 1996) or an epitope tag (e.g. hemagglutinin; Field et al., 1988) are created and introduced into the investigated organism. The Gateway technology (Invitrogen) based on the site-specific recombination mechanism of phage lambda (Landy, 1989) allows rapid cloning of DNA sequences to vectors carrying designated tag sequences. Most of the published Gateway-compatible binary vectors (reviewed by Karimi et al., 2007) are designed for constitutive expression of transgenes therefore harboring the 35S promoter of cauliflower mosaic virus (*CaMV*, Odell et al., 1985) or the nopaline synthase (*Nos*) promoter of *Agrobacterium tumefaciens* (Depicker et al., 1982). This strategy is disadvantageous for purposes like purification of protein complexes via epitope tagged bait proteins because overexpressed proteins often are not associated with their binding partners. Nevertheless, most purification strategies so far rely on the overexpression of the bait protein in wild-type background with the transgenic protein competing for binding partners with the endogenous protein (e.g. Rubio et al., 2005). However, to assure proper function of tagged proteins they should be introduced into respective mutant backgrounds and reconstitute the wild-type phenotype. Furthermore, ubiquitous expression may affect complementation (Laufs et al., 2003), thus demanding for proper regulation of spatial and temporal expression. In this regard, the use of endogenous promoters or promoters of genes with similar expression profiles is more promising.

We designed the binary, Gateway-compatible “pAUL“ vector series for epitope tagging of proteins that are expressed under the control of endogenous *Arabidopsis thaliana* promoters or the 35S *CaMV* promoter (Odell et al., 1985). The primary application is supposed to be detection and purification of nuclear encoded proteins involved in chloroplast-related processes. Thus, vectors with C-terminal tags were generated in the first instance, as N-terminal fusions would be cleaved off toward chloroplast import. The C-terminal tags are combined with *A. thaliana* promoter sequences of genes known to participate in those processes, namely *HCF107* (Felder et al., 2001, Sane et al., 2005), *HCF136* (Meurer et al., 1998), and *HCF173* (Schult et al., 2007). To make the vectors applicable for proteins involved in other biological processes vectors harboring the 35S *CaMV* promoter combined with C-and N-terminal tags were also constructed.

Three epitope tags were utilized for four different C- or N-terminal fusions making possible single-, double- or triple-tagging of proteins of interest. The hemagglutinin (HA) epitope exhibits a small size (27 amino acids for 3x HA) and the availability of effective antibodies make it an ideal tool for detection. Purification can also be carried out in small scales via antibodies or anti-HA matrices and proteins can be eluted competitively by HA peptide or by low pH. The 28-amino acid Strep-tag*III* (Junttila et al., 2005) is an improvement of Strep-tag*II* and has not been described for purification of plant proteins so far. This tag has a strong binding affinity to Strep-Tactin, an engineered streptavidin derivate. Purifications can be performed under flexible binding conditions as Strep-tag*III* / Strep-Tactin interactions are resistant to detergents and varying salt concentrations and do not require the availability of cofactors (Juntilla et al., 2005). Also, the possibility of competitive elution via desthiobiotin makes it a suitable tool for protein and protein complex purification (Juntilla et al., 2005). In the pAUL vector system, the HA epitope and Strep-tag*III* can be used for single-tag-fusions of proteins of interest for detection (HA) or purification (HA and Strep-tag*III*). Double tagging includes both the HA epitope and Strep-tag*III* cloned in series and is supposed to serve for one-step purification via StrepTactin and subsequent detection via the HA epitope. Alternatively, two-step purification via StrepTactin and anti-HA affinity matrix may be carried out if required. Finally, we designed an alternative TAP (tandem affinity purification)-tag. The TAP tag originally developed in yeast consists of two immunglobulin-binding domains of protein A from *Staphylococcus aureus* (ProtA), a tobacco etch virus (TEV) cleavage site and a calmodulin binding site (CBP) (Rigaut et al., 1999), but has been modified in the past years (reviewed by Xu et al, 2010). Rohila et al. (2004) adapted this tag to plant applications and Rubio et al. (2005) further modified it. We exchanged the CBP by HA for efficient detection of the tagged protein and Strep-tag*III* for purification. The ProtA tag was retained since it displays a strong binding affinity to IgG Sepharose making it well suitable for protein purification. However, the large size of the tag (116 amino acids; ~13 kDa) may affect the function of the protein fused to it. The TEV cleavage site from the original TAP tag was replaced by the human rhinovirus (HRV) 3C protease cleavage site, which can be processed even at low temperatures according to Rubio et al. (2005).

Here, we describe cloning of the pAUL vector series. We test the 35S *CaMV* promoter and promoters from *HCF107*, *HCF136*, and *HCF173* for their activities by quantitative and histochemical GUS assays. Complementation with different promoter/ tag combinations is tested using *hcf107.2* (Felder et al., 2001) and *hcf208* (Lyska et al., 2007) mutants. Corresponding proteins are encoded in the nucleus and transported to chloroplast membranes where they affect thylakoid membrane biogenesis. Whereas HCF107 forms a low abundant high molecular weight complex (Sane et al., 2005) and is required for expression of the chloroplast-encoded photosystem II subunit PsbH, HCF208 is part of the

system IV c-type cytochrome maturation machinery for the b_6 subunit of the cytochrome b_6f complex (Kuras et al., 2007; Lezhneva et al., 2008) and fulfills its function as a stable heterodimer transiently interacting with other proteins (Saint-Marcoux et al., 2009). Finally, integrity of the different tags is tested by small-scale affinity purification of HCF208 from thylakoid membranes.

Results and Discussion

Design and Cloning of the pAUL Vector Series

We constructed a series of 20 binary Gateway-compatible vectors containing different combinations of promoters and tags, named pAUL1-20 (Figure 1). Four different single, double or triple tags were cloned into various backbone Gateway vectors from the pMDC series (Curtis and Grossniklaus, 2003) allowing either C- or N-terminal protein fusions to facilitate protein detection and purification. N-terminally tagged proteins can be expressed under the control of two copies of the 35S *CaMV* promoter, whereas vectors for C-terminal fusions contain either the two copies of the 35S *CaMV* promoter for ubiquitous and constitutive expression or one of three endogenous promoters from *A. thaliana*. The endogenous promoters were selected according to the function of the respective genes in chloroplast biogenesis and are described in the following section.

Single tags are the triple HA epitope and Strep-tag*III*. Both tags were combined in order to create the double tag. Addition of a 3C protease cleavage site and the ProtA tag made the triple tag. N- and C-terminal double and triple tags exhibit reverse orientations.

Sequences for C-terminal tags and a *Nos* terminator as well as the four promoters were cloned to the Gateway vector pMDC123 (Curtis and Grossiniklaus, 2003) resulting in vectors pAUL1 to pAUL12 (Figure 1A). Plasmid pMDC123 was used as the recipient for the tagging constructs, because this vector does not contain any preexisting promoter nor tag sequences around the Gateway att cassette but unique restriction sites (*Asc*I upstream of the att cassette, *Sac*I downstream of the att cassette) making it suitable for inserting promoter/ tag sequences. Moreover, it harbors a bar sequence encoding for phosphinothricin (Basta) resistance driven by a 35S *CaMV* pormoter.

The C-terminal Strep-tag*III* and promoter sequences were introduced into pMDC99 (Curtis and Grossniklaus 2003) which corresponds to pMDC123 except it carries a hygromycin resistance instead the bar gene. Those vectors were named pAUL13 to pAUL16 (Figure1A).

If required, promoters from pAUL1 to pAUL16 can be exchanged easily, as they were cloned after the tag sequences by the rare cutting restriction enzyme *Asc*I.

For vectors pAUL17 to pAUL20 N-terminal tags were inserted into pMDC32 (Curtis and Grossniklaus 2003), which contains a 35S *CaMV* promoter, a *Nos* terminator and a hygromycin resistance (Figure 1B). These vectors were not tested in this study as we investigated chloroplast-localized proteins whose N-termini are cleaved off upon import into the chloroplast.

Both, N- and C-terminal tag sequences were inserted into the expression cassette in a way that allows easy cloning of sequences according to the Gateway manual (Invitrogen).

Characterization of 35S CaMV, HCF107, HCF136, and HCF173 Promoters

The three different promoter regions from genes *HCF107*, *HCF136*, and *HCF173* were selected according to the respective mRNA profiles from the GENEVESTIGATOR database (Zimmermann et al., 2004). Experiments from Smith et al. (2004) indicate that *HCF136* and *HCF173* mRNAs accumulate to similar levels but are about 4-fold higher than *HCF107* mRNAs. Furthermore, *HCF107* and *HCF173* mRNAs are regulated diurnally, whereas the *HCF136* mRNA levels are stable throughout the day. The putative promoters were defined as sequences upstream of the transcription initiation site of the respective genes ending in regions of ~1500 bp or until a UTR of the previous gene is reached. 1525 bp of the sequence upstream of *HCF107* 5´UTR, 1401 bp upstream of *HCF136* 5´UTR, and 721 bp upstream of *HCF173* 5´UTR were cloned to pAUL vectors and are referred to as "*HCF107* promoter" (p*HCF107*), "*HCF136* promoter" (p*HCF136*), and "*HCF173* promoter" (p*HCF173*) in the following.

To test the ability of the selected sequences to serve as promoters and to compare them to the 35S *CaMV* promoter (p*35S CaMV*) quantitative and histochemical GUS assays were performed. The ß-glucoronidase gene was fused to the four promoter sequences and introduced into wild-type *A. thaliana* plants. As presented in Figure 2 all putative promoter sequences and p*35S CaMV* do function as promoters.

In intact *A. thaliana* seedlings p*35S CaMV* expression is detected in all plant organs, including cotyledons, hypocotyls, roots, and seed coat (Figure 2A). In contrast, p*HCF107*, p*HCF136*, and p*HCF173* are only active in cotyledons and hypocotyls representing the "green" tissue of the seedling but not in roots and seed. This is consistent with the function of HCF107, HCF136, and HCF173 in chloroplast biogenesis (Meurer et al., 1998; Sane et al., 2005; Schult et al., 2007). To address tissue specificity of the promoters inside leaves cross sections were prepared (Figure 2B). The 35S *CaMV* promoter is active in all cell types, which agrees with previous studies (Battraw and Hall, 1990). However, the staining pattern appears spotted, suggesting that expression is not uniform throughout cell layers and types. In contrast, all endogenous promoters display even staining patterns. Expression of GUS driven by p*HCF107*, p*HCF136*, and p*HCF173* is restricted to palisade and spongy mesophyll cells, which are the chloroplast possessing tissues. Together, these results show that all chosen endogenous promoter sequences are applicable for proper spatial expression of chloroplast-related proteins.

Since another aim of using endogenous promoters was to drive expression more moderately than the 35S *CaMV* promoter and to ensure proper temporal expression, promoters were also analyzed quantitatively. Data was generated for two different developmental stages (15 and 30 days after germination) from plant material always harvested at the same time of day.

In 15 day-old plants, p*HCF107* is the weakest of the endogenous promoters, since p*HCF136* and p*HCF173* exhibit ~ 2.3-fold and ~ 1.6-fold higher GUS activity than p*HCF107*, respectively (Figure 2C). p*35S CaMV* is the strongest promoter, presenting ~3.3-fold higher activity than p*HCF107* if median values are compared.

However, values for p*35S CaMV* are strongly dispersed with their maximum and minimum at 280 and 3 mmol MU/(mg protein*min), respectively, revealing a high variation of expression in individual lines. This characteristic of p*35S CaMV* has been reported previously (van Leeuwen et al., 2001). Furthermore, the presence of multiple 35S *CaMV* promoter copies is supposed to lead to silencing effects (Daxinger et al., 2008), which may occur in vectors that drive expression of the selectable marker by the same promoter and if plants are homozygous for the T-DNA. Expression by endogenous promoters appears more constant throughout different lines suggesting that they do not interfere with any other features of the pAUL vectors.

30 days after germination promoter activities appear to be decreased compared to values from 15 day-old plants (Figure 3C). p*HCF107* activity is only slightly reduced to ~ 80% and therefore is relatively stable. In contrast, p*HCF173* and p*HCF136* activities are drastically reduced to ~ 53% and ~ 25% respectively, obtaining values similar to p*HCF107*. However, p*35S CaMV* expression is relatively stable when individual values are taken into account rather than the median value, which again is not representative due to the high spread. Additionally, p*35S CaMV* driven expression is 3-fold higher than the endogenous promoters at that stage.

Altogether, it can be stated that p*HCF107* is suitable for constant and moderate expression of chloroplast-specific proteins, whereas p*HCF136* and p*HCF173* are optimal for proteins involved in early developmental stages (at least up to 15 days after germination) and which are not essential or would even be distracting in later stages. p*HCF136* is adequate for stronger and p*HCF173* for a more moderate early expression. Consequently, before starting cloning of sequences, one should check databases like GENEVESTIGATOR (Zimmermann et al., 2004), eFP Browser (Winter et al., 2007) to select the adequate vector.

Complementation Analyses of hcf107.2 and hcf208

Tagging of proteins, i.e. attachment of other protein sequences of up to several kDa (e.g. GFP: 27 kDa), and expression driven by foreign promoters may affect the function of fusion proteins. In order to test the influence of the presented tags on protein function and the differential expression by p*35S CaMV* or endogenous promoters, cDNAs of *A. thaliana* chloroplast biogenesis factors HCF107 and HCF208 were introduced into a subset of pAUL vectors (pAUL1, 2, 3, 6, 9), carrying either a single-, double-, or triple tag and *p35S CaMV* or

triple tags and p*HCF107* or p*HCF136* (Figure 3A). The constructs were transformed into heterozygous *hcf107.2*/*HCF107* or *hcf208*/*HCF208* background, as homozygous mutants cannot grow photoautotrophically. Transformants were screened for BASTA resistance and homozygous mutant backgrounds on sucrose-supplemented 0.5x MS medium and then were transferred to soil to test their capability to grow photoautotrophically.

The grade of complementation was tested by chlorophyll fluorescence measurements on 3 week-old plants (Figure 3B). Complementation of the photosystem II biogenesis factor *hcf107.2* was indicated by the Fv/Fm ratio, the crucial parameter for photosystem II activity (Genty et al., 1989) indicating an estimate of the maximum portion of absorbed quanta used in photosystem II reaction centers. Since photosystem II is intact in *hcf208* but the downstream cytochrome b_6f complex is strongly reduced, qP (photochemical quenching) values were determined for indication of complementation. qP displays reduction of variable fluorescence by photosynthetic electron transport processes. From each tested line proteins were isolated and analyzed by Western blot in order to determine levels of fusion proteins.

For HCF208, all promoter/ tag combinations are able to fully complement the mutant phenotype (Figure 3B). qP values of both, wild-type and complemented lines are ~ 0.9 compared to the drastically lower value 0.16 of *hcf208*. These results indicate that none of the three tags fused to the C-terminus of HCF208 affect its functionality and that all tested promoters are able to drive transgene expression in a way that is sufficient for HCF208 complementation. However, protein levels of transgenic HCF208 vary strongly depending on the construct, as revealed by Western blot analysis using the HA antibody (Figure 3C). Among the transgenes driven by p*35S CaMV*, all three independent HCF208pAUL1 lines accumulate very low protein levels unlike HCF208pAUL2- and HCF208pAUL3-constructs. One of the HCF208pAUL2 lines (line 5) also accumulates low amounts of HCF208 compared to the other two lines. There are two possible explanations for low accumulation of HCF208 in pAUL1 lines: either (i) unlike the double- and triple tag, the HA epitope destabilizes HCF208 when attached to its C-terminus or (ii) incidentally all three randomly selected HCF208pAUL1 lines and HCF208pAUL2-5 are silenced. Nevertheless, the residual amounts of HCF208 in these lines are sufficient to complement the mutant phenotype. Also, HCF208 levels in pAUL9 lines correspond to the low levels in HCF208pAUL1 and these lines, too, are fully complemented. As previously tested, the activity of p*HCF136* in pAUL9 decreases during plant development, which accounts very likely for the low protein levels in three week-old HCF208pAUL9 lines. In HCF208pAUL6 lines harboring the transgene driven by p*HCF107*, proteins accumulate to levels similar to HCF208pAUL3. Altogether, it can be stated that protein levels in HCF208pAUL2 (except line 5), -3, and also -6 represent an overexpression of HCF208 exceeding endogenous levels. Unfortunately, no HCF208-specific antibody was available preventing comparison of protein levels in complemented lines to the

wild-type situation. In conclusion, these results indicate that (i) the C-terminus of the integral membrane protein HCF208, which forms a large domain extending to the stroma (Lezhneva et al., 2008) is not prone to attachment of large tags, although it might be influenced by the HA epitope and (ii) expression of HCF208 can be driven by all promoters to complement the mutant phenotype, but the 35S *CaMV* promoter may be silenced in some lines.

The situation for *hcf107.2* is different in some ways (Figure 3B). Wild-type Fv/Fm values of ~ 0.83 compared to ~0.1 in the mutant *hcf107.2* are only reached by plants carrying HCF107pAUL1 and HCF107pAUL2 vectors, both driving expression by p*35S CaMV* and possessing small tags. Extension of the tag sequence by ProtA (pAUL3) decreases Fv/Fm values to ~ 0.74. Expression of the triple-tagged protein by p*HCF107* or p*HCF136* (HCF107pAUL6 and HCF107pAUL9) further decreases Fv/Fm values to ~ 0.65 and ~ 0.53 respectively (Figure 3B). According to these values, HCF107pAUL6 and HCF107pAUL9 plants are paler and smaller than wild type and HCF107pAUL1 to -3 plants. Even three weeks after germination, HCF107pAUL9 plants are very small and hardly produce seed, whereas the defect in HCF107pAUL6 is less severe (Figure 3B).

Western blot analysis was carried out using antibodies against the HA-epitope and HCF107, which was generated in our laboratory. As indicated in Figure 3C, HCF107pAUL1 and HCF107pAUL2 lines over-accumulate the fusion protein compared to wild-type levels, whereas levels of all triple-tagged proteins are significantly lower. Transgenic lines expressing triple-tagged HCF107 under p*35S CaMV* control (pAUL3) exhibit about wild-type amounts of HCF107 but, as indicated before, the Fv/Fm ratio displaying photosystem II activity is decreased. This points to an inhibitory effect of the triple tag on protein stability and function. Expression of the triple-tagged protein by pHCF107 and p*HCF136* results in protein levels below wild-type amounts. In HCF107pAUL9 plants HCF107 is hardly detectable.

Former experiments revealed that HCF107 forms a high molecular weight complex (Sane et al., 2005). Thus, it is possible that the restriction of protein function by the triple tag may be due to inefficient complex assembly and subsequent degradation of unassembled protein. On the other hand, the protein itself may be unstable independent of its assembly state. In order to achieve nearly wild-type situation, triple tagged HCF107 needs to be overexpressed.

This detailed complementation analysis leads to the conclusion that it strongly depends on the investigated protein which promoter and tag should be chosen for experiments. For HCF208, large tags and expression by endogenous promoters were suitable for complementation, as HCF208 seems to be not required in large amounts for its function and C-terminal tags do not impair protein function, irrespective of their size. If co-purification of assembly partners is performed, p*35S CaMV* driven expression would be disadvantageous, thus an endogenous promoter should be chosen. In case of HCF107 large tags impair protein function and/ or result in destabilization of the protein. Thus, only smaller tags are

suitable or overexpression of the incorporated gene is necessary to ensure complementation of the mutant phenotype. This also shows that protein analyses and purification should be carried out in mutant background if possible to show that the protein is not affected by its tag and ectopic expression.

Purification of HCF208 from Thylakoid Membranes

Integrity of the HA epitope, Strep-tag*III*, and the ProtA tag and purification via these tags were tested on HCF208pAUL1, -pAUL2, and -pAUL3 transgenic lines. The difficulty in purification of HCF208 lies in its feature to be an integral membrane protein. First, crude membranes were treated with 1% n-dodecyl-ß-D-maltoside to solubilize proteins. Subsequently, all purification steps had to be performed in the presence of 0.05 % n-dodecyl-ß-D-maltoside to keep proteins soluble.

The western blot analyses presented in the previous chapter indicate that the HA epitope is intact in all tag variants and that it is well suitable for detection using HA antibody. In contrast, detection of the Strep-tag*III* in plant protein extracts at least under our conditions produced an extremely high background. Purification efficiency via anti-HA matrix was tested using solubilized protein extract from HCF208pAUL1. Elution was carried out competitively by the HA-peptide (Figure 4A). The feasibility of the double-, and triple tag for tandem purifications was tested using proteins from HCF208pAUL2 and -pAUL3 transgenic lines. Double-tagged HCF208 protein was loaded on StrepTactin matrix first, eluted competitively by desthiobiotin, and then purified via anti-HA affinity matrix (Figure 4B). Triple-tagged proteins were purified via IgG Sepharose first and eluted by 3C protease cleavage. In the second step, the eluate was incubated with StrepTactin matrix and eluted as described above (Figure 4C). From all purifications significant amounts of HCF208 could be recovered and eluates exhibited no abundant signals (ATP Synthase and Ponceau staining). Successive purification of triple tagged HCF208 from HCF208pAUL3 results in a protein of lower molecular size in the eluate compared to the input, which is consistent with the lack of the ProteinA tag cleaved off by 3C protease treatment (Figure 4C).

These experiments show that both one step and tandem affinity purifications can be carried out with our tagging system always using same buffer conditions and in the presence of detergents. Further, we introduced Strep-tag*III* as a novel epitope that can be used for tagging of plant proteins and purification of proteins and protein complexes due to its property of binding at 4°C and competitive elution via desthiobiotin. Combination with the HA epitope allows easy detection of fusion proteins and, if desired, an additional low-scale purification step. Tandem affinity purification can be carried out using the triple tag via ProtA,

which can be cleaved off by 3C protease, and subsequently via Strep-tag*III* with all purification steps at 4°C.

Experimental Procedures

pAUL Vector Construction

Construction of pAUL 1 to pAUL12

Cassette C1 from the pMDC123 Gateway vector (Curtis and Grossniklaus, 2003) was modified prior to cloning of tags and promoters into the plasmid. The att cassette was amplified using primers pMDC123-H and pMDC123-R and then removed from the vector by digestion with *Asc*I and *Sac*I. The PCR product lacking 44bp stop codon-containing sequence between attR2 site and *Sac*I recognition site was digested using *Asc*I and *Sac*I and cloned into pMDC123 to generate pMDC123-(-)stop. The integrity of the att cassette was tested by sequencing.

The tags were first assembled in pBluescript II (pBSII) KS+ and each fused to the *Nos* terminator, which was obtained form the pC-TAPa plasmid (Rubio et al., 2005) by PCR reaction using primers T NOS*Pst*I-H and T NOS*Sac*I/*Hin*dIII-R. The *Pst*I/*Hin*dIII digested PCR product was ligated into the pBSII phagemid already containing the assembled tags, which were generated as follows.

For vectors containing only the 3x HA tag the DNA sequence from plasmid spa1g3xHA-pBS (provided by Ute Hocker; Sato and Wada, 1997) was amplified with primers 3xHA*Sac*I-H and 3xHAstop*Pst*I-R and cloned using *Sac*I and *Pst*I generating pBS-3xHA*Pst*I.

The DNA sequence of 3x HA tag cloned into the multiple tags was amplified using primers 3xHA*Sac*I-H and 3xHA*Bam*HI-R2 and cloned into pBSII after digestion with *Sac*I and *Bam*HI generating pBS+3xHA*Bam*HI.

Strep-tag*III* (Junttila et al., 2005) was obtained by annealing of the oligonucleotides HA/STREP-1 to -6 at 90°C for 15 minutes in annealing buffer (0.1M Tris/ HCl, pH 7.5; 1M NaCl; 10mM EDTA). The DNA sequence was adapted to plant codon usage. For the generation of 3x HA/Strep-tag*III* the DNA fragment was amplified using primers HA/STREP-7 and HA/STREP-9. The *Bam*HI/*Pst*I digested PCR product was cloned into pBS+3xHA*Bam*HI generating pBS+3xHA/Strep*III**Pst*I.

The 3xHA/Strep*III*/PA tag was created by amplification of the Strep-tag*III* DNA fragment with primers HA/Strep*III*-7 and 3xHAStrep*III**Xba*I-R and amplification of the 2xProteinA tag including 3C protease cleavage site from pC-TAPa using primers 3cIgG-BD*Xba*I-H and 3cIgG-BD*Pst*I-R. Both PCR products digested *Bam*HI/*Xba*I and *Xba*I/*Pst*I respectively were cloned into pBS+3xHA*Bam*HI generating pBS+3xHA/Strep*III*/PA*Pst*I.

The three tag constructs including nos terminator were removed from pBSII by digestion with *Sac*I and ligated into the pMDC123-(-)stop plasmid creating plasmids pMDC123-(-)stop-

3xHA, pMDC123-(-)stop-3xHA/Strep*III* and pMDC123-(-)stop-3xHA/Strep*III*/PA. The correct orientation of the tags was verified by restriction analysis.

The 2x35S *CaMV* promoter and the three endogenous promoters from *HCF107*, *HCF136* and *HCF173* were each cloned by PCR reaction and digestion with *Asc*I into the three pMDC123-(-)stop plasmids containing the tag constructs.
From plasmid pYL436 the 2x35S *CaMV* promoter was amplified using primers 35S*Asc*I-H and 35S*Asc*I-R. Predicted promoter sequences of genes *HCF107*, *HCF136* and *HCF173* were amplified from genomic DNA of *A. thaliana* Columbia-0 ecotype. The *HCF107* promoter was amplified by adapter PCR using primers *HCF107*-Prom-H1 and *HCF107*-Prom-R1 in the first step and primers *HCF107*-Prom-H5-2 and *HCF107*-Prom-R5 in the second step. Promoters of *HCF136* and *HCF173* were amplified with primers *HCF136*-Prom-H1/ *HCF136*-Prom-R1 and *HCF173*-Prom-H1/ *HCF173*-Prom-R1 respectively. *Asc*I digested PCR products were ligated to pMDC132-(-)stop-3xHA, pMDC132-(-)stop-3xHA/Strep*III* and pMDC123-(-)stop-3xHA/Strep*III*/PA generating vectors pAUL1 to pAUL12. Correct orientation was checked by restriction analyses. After completion of the vectors promoters and tags were checked by sequencing.

Construction of pAUL13 to pAUL16

For the creation of pAUL13 to pAUL16 the modified att cassette from pMDC123-stop was transferred to the Gateway vector pMDC99 (Curtis and Grossniklaus, 2003) generating pMDC99-(-)stop. The Strep-tag*III* sequence was amplified with primers HA/ Strep-9 and Strep*III*/ *NotSac*-H from vector pBS+3xHA/Strep*III**Pst*I. After digestion with *Not*I and *Pst*I the fragment was ligated to a pBluescript II vector already containing the *Nos* terminator sequence (cloned as described above). Correct fragments were removed from pBSII via a *Sac*I restriction site and ligated to pMDC99-(-)stop creating pMDC99-Strep. Promoters 2x35S *CaMV*, *HCF107*, *HCF136*, and *HCF173* were extracted from the above-described pAUL vectors via *Asc*I and fused to pMDC99-Strep generating pAUL13 to pAUL16. Correct orientation was checked by restriction analyses. After completion of the vectors promoters and tags were checked by sequencing.

Construction of pAUL 17 to pAUL20

N-terminal versions of the pAUL vector series were produced using pMDC32 (Curtis and Grossniklaus, 2003) containing a 2x35S *CaMV* promoter and nos terminator as a backbone.

3xHA and Strep-tag*III* sequences were amplified from the pAUL2 vector, pN-TAPa (Rubio et al., 2005) served as template for IgG-BD+3C protease cleavage site.

In order to create vectors containing the 3x HA only, the sequence was amplified with primers 3xHA-*Asc*I/*Pst*I-H and 3xHA-*Asc*I/*Hin*dIII-R. For double and triple tag primers 3xHA-*Pst*I-H and 3xHA-*Asc*I/*Hin*dIII-R used. Both PCR products were digested by *Pst*I and *Hin*dIII and ligated to pBSII to create pBS-3xHA-N and pBS-*Pst*I-3xHA-N respectively.

For the single tag Strep-tag*III* was amplified via primers Strep-*Asc*I-H2 and Strep-*Asc*I-R, restricted with *Asc*I and directly ligated to *Asc*I-linearized pMDC32. The Strep-tag*III* sequence needed for the 3xHA/Strep*III* double tag was amplified with primers Strep-*Sac*I/*Asc*I-H-2 and Strep-*Pst*I-R. The *Sac*I/*Pst*I digested fragment was fused to pBS-*Pst*I-3xHA-N resulting in pBS-Strep*III*/3xHA-N. Finally, primers Strep-*Xba*I-H1 and Strep-*Pst*I-R were employed for amplification of Strep-tag*III* needed for the triple tag. The fragment was restricted with *Xba*I and *Pst*I and introduced into pBS-*Pst*I-3xHA-N generating pBS-*Xba*I-Strep*III*/3xHA-N.

The sequence of the 2x Protein A tag and 3C protease cleavage site was amplified with primers IgG-*Asc*I/*Bst*XI-H and IgG-3C-*Bcu*I-R2 and digested with enzymes *Bst*XI and *Bcu*I. The fragment was ligated to pBS-*Xba*I-Strep*III*/3xHA-N digested with *Xba*I and *Hin*dIII (*Bcu*I and *Xba*I form compatible sticky ends) to create pBS-PA/Strep*III*/3xHA-N.

All tag sequences and pMDC32 were restricted with *Asc*I and ligated. After confirmation of the correct orientation the tags were checked by sequencing.

Cloning of Target Genes to pAUL Vectors

For characterization of the promoters the ß-glucuronidase (*GUS*) gene and the Nos terminator were amplified from the pBI121 vector (Jefferson et al., 1987) using primers *GUS*+Term-H and *GUS*+Term-R containing attB sites.

For complementation analyses *HCF107* was amplified with primers start107attB1-H and 107attB2-R using pPEX107PA (Sane et al., 2005) as a template. *HCF208* was amplified from cDNA obtained by reverse transcription of total RNA isolated from *A. thaliana* wild type Columbia-0 using primers start208attB1-H and 208attB2-R.

BP clonase reaction (Invitrogen) between the PCR products and pDONR221 were accomplished according to the Gateway manual creating pENTRY221+*GUS*, pENTRY+*HCF107*, and pENTRY+*HCF208*. Aliquots (5μl) were transformed into *Escherichia coli* strain *DH5α* using heat shock. The recombinants were selected on LB agar plates containing 50μg/ ml kanamycin. After sequence analysis of the recombined DNA sequence LR clonase reactions (Invitrogen) were performed (according to the Gateway manual) to introduce the genes into the respective pAUL vectors. The ß-glucuronidase gene was

introduced into pAUL1, pAUL4, pAUL7, and pAUL10 creating GUSpAUL1, GUSpAUL4, GUSpAUL7, and GUSpAUL10. *HCF107* and *HCF208* were recombined into pAUL1, pAUL2, pAUL3, pAUL6, and pAUL9. Resulting vectors were named HCF107pAUL1, HCF107pAUL2, HCF107pAUL3, HCF107pAUL6, HCF107pAUL9, HCF208pAUL1, HCF208pAUL2, HCF208pAUL3, HCF208pAUL6, and HCF208pAUL9.

Because pAUL vectors as well as the pENTRY221 vectors can only be selected on kanamycin containing media 5μl of each reaction mixture were digested *Hpa*I or *Eam*1105I respectively in order to linearize the pDONR221+*GUS* vector to avoid its transformation. The vectors were transformed into *DH5α* using heat shock and recombinants were selected on LB agar media supplemented with 50 μg/ml kanamycin. After restriction analyses the correct reading frame of each gene and the tags was checked by sequencing.

Plant material, growth conditions, and plant transformation

All constructs were transformed into *Agrobacterium tumefaciens* strain GV3101 and introduced to *A. thaliana* using the floral dip method (Clough and Bent, 1998). GUS constructs were transferred into wild-type Columbia-0 ecotype. pAUL vectors containing *HCF107* or *HCF208* cDNA were introduced into heterozygous *hcf107.2* (Wassilewskija ecotype, Felder et al., 2001) and *hcf208* (Columbia-0 ecotype, Lyska et al., 2007) plants respectively.

For seed production, protein extraction, spectroscopic measurements, and measurement of GUS activity, plants were grown on soil in a growth chamber operating at a 16h light/ 8h darkness period at a photon flux density (PFD) of ~50-70 μmol s^{-1} m^{-2} and a constant temperature of 21°C. Protein extracts used for affinity purification were isolated from plants grown under short-day conditions (8h light/ 16h darkness).

Homozygous *hcf208* and *hcf107.2* plants were grown on 0.5 x Murashige and Skoog (MS) medium (Murashige and Skoog, 1962) containing 2% (w/v) sucrose and 0.3% (w/v) gelrite (Roth, Karlsruhe, Germany). Seedlings were exposed to a 16h light/ 8h darkness period at a PFD of ~50-70 μmol s^{-1} m^{-2}. Selection of mutant plants exhibiting high chlorophyll fluorescence phenotype was performed in the dark under UV light (Meurer et al., 1996).

Transformants were selected on 0.5 x MS medium as described above containing 10 μg ml^{-1} phosphinothricin.

Measurement of GUS Activity and Histochemical Analyses

GUS analyses were carried out with T1 plants of *A.thaliana* harboring GUSpAUL1, GUSpAUL4, GUSpAUL7, and GUSpAUL10, respectively. Quantitative determination of GUS

activity was performed with 15- and 30-day-old plants harvested at midday after 6 hours of illumination according to Jefferson (1987) and Kosugi et al. (1990). The average values of the data are expressed by medians.

For histochemical analyses either intact 5-day-old seedlings or sections of 3-week-old plants cut manually with a razorblade were transferred into incubation buffer (100 mM Na_2HPO_4, pH 7.5; 10 mM EDTA; 50 mM K_4 [$Fe(CN)_6$]; 50 mM K_3[$Fe(CN)_6$]; 0.1% (v/v) Triton X-100; 2 mM 5-bromo-4-chloro-3-indolyl-β-D-glucuronid acid) and vacuum-infiltrated. Samples were incubated at 37°C until they stained blue and fixed in 75% ethanol and 25% acetic acid for 10 minutes. Subsequently, chlorophyll was removed by treatment with 70% ethanol.

Fluorescence Measurements

For fluorescence measurements complemented *hcf107.2* or *hcf208* plants carrying the respective wild-type gene in pAUL1, -2, -3, -6, or -9 vectors were employed. Chlorophyll fluorescence was imaged with a closed FluorCam FC 800-C controlled by FluorCam 6 software (Photon Systems Instruments) on 3-week-old plants. Experiments were carried out using pre-designed quenching protocol provided by the software.

Protein Extraction and Western Blot Analysis

Proteins were extracted from the same plants that were also used for fluorescence measurements described above according to Shen et al. (2007). 3 plants from each line were pooled, pestled in liquid nitrogen and immediately transferred to extraction buffer (10mM Tris/ HCl, pH 7.8; 4 M urea; 5% (w/v) SDS; 15% (w/v) glycerol; 10mM β-mercaptoethanol). The samples were boiled for 4 minutes and cleared by centrifugation at 15,000g for 5 minutes. Protein concentration was determined using RC DC Protein Assay (Bio-Rad). 50 μg total protein was separated on a 10% SDS-PAGE gel according to Schagger and von Jagow (1987), transferred to nitrocellulose membranes, and immunoblots were decorated with anti-HCF107 and Anti-HA-Peroxidase (Roche Applied Science).

Affinity purification of fusion proteins

Leaves of 4- to 6-week-old complemented *hcf208* carrying either HCF208pAUL1, HCF208pAUL2 or HCF208pAUL3 constructs were homogenized in lysis buffer (10 mM Hepes/ KOH, pH 7.8; 10 mM $MgCl_2$; and 25 mM KCl). Cell debris was separated by Miracloth filtration. The suspension was centrifuged at 4°C until a speed of 5,900g was reached and then stopped. Pelleted membranes were resuspended in Tris-buffered lysis buffer and

solubilized with 1% (w/v) n-dodecyl-ß-D-maltoside (Calbiochem) for 30 minutes at 4°C and a chlorophyll concentration of 1 mg ml^{-1}. Unsolubilized material was removed by centrifugation (20 minutes at 4°C and 15,000g). For each purification aliquots of 200 µl supernatant were added to matrices pre-equilibrated with washing buffer (20 mM Tris/ HCl, pH 7.8; 150 mM NaCl; 1mM EDTA; 0.05% (w/v) n-dodecyl-ß-D-maltoside).

Samples from HCF208pAUL1 (containing the 3x HA tag only) plants were incubated with a bed volume of 50 µl Anti-HA affinity matrix (Roche Applied Sciences) for 1 h at 4°C on a rotator. The matrix was washed with 20 volumes of washing buffer. Proteins were eluted by incubating the affinity matrix three times with 1 volume of elution buffer (1 mg/ ml HA peptide (Roche Applied Sciences)) in washing buffer for 15 minutes at 37°C.

HCF208pAUL2 (Strep-tag*III* and 3x HA tag) samples were rotated with 100 µl Strep-Tactin Macroprep (IBA) at 4°C for 1 h, washed with 5 times with 1 volume of washing buffer and eluted with 3 volumes of elution buffer (2.5 mM desthiobiotin (IBA) in washing buffer). Subsequently, the eluate was purified via Anti-HA affinity matrix as described above.

For purification of HCF208pAUL3 (2x ProtA, Strep-tag*III*, and 3x HA tag) samples, respective supernatants were incubated with 50 µl IgG Sepharose (GE Healthcare) for 1 h at 4°C on a rotator. After washing with 20 volumes of washing buffer the matrix was equilibrated with 5 volumes of cleavage buffer (20 mM Tris/ HCl, pH 7.8; 150 mM NaCl; 1mM EDTA; 0.05% (w/v) n-dodecyl-ß-D-maltoside; 1mM DTT). Cleavage was performed by incubation with 20 units of PreScission Protease (GE Healthcare) in cleavage buffer for 16 h at 4°C on a rotator. The following purification with Strep-Tactin Macroprep was performed according to the procedure described above.

All final eluates were precipitated with 15% (w/v) trichloroacetic acid and separated on 12.5% SDS-PAGE gels according to Schagger and von Jagow (1987). Proteins were transferred to nitrocellulose membranes and immunodecorated with Anti-HA-Peroxidase (Roche Applied Sciences) and anti-ATP Synthase.

Acknowledgements

We thank Stefanie Schulze for preparation and staining of leaf sections and Dr. Peter Jahns (Heinrich-Heine-Universitaet Duesseldorf) for his support with fluorescence measurements with the FlourCam system. Work described here was supported by the Deutsche Forschungsgemeinschaft (DFG) through Sonderforschungsbereich Transregio TR1.

Literature:

Battraw, M.J. and Hall, T.C. (1990) Histochemical analysis of CaMV 35S promoter-beta-glucuronidase gene expression in transgenic rice plants. *Plant Mol Biol*, **15**, 527-538.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J*, **16**, 735-743.

Curtis, M.D. and Grossniklaus, U. (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol*, **133**, 462-469.

Daxinger, L., Hunter, B., Sheikh, M., Jauvion, V., Gasciolli, V., Vaucheret, H., Matzke, M. and Furner, I. (2008) Unexpected silencing effects from T-DNA tags in Arabidopsis. *Trends Plant Sci*, **13**, 4-6.

Depicker, A., Stachel, S., Dhaese, P., Zambryski, P. and Goodman, H.M. (1982) Nopaline synthase: transcript mapping and DNA sequence. *J Mol Appl Genet*, **1**, 561-573.

Felder, S., Meierhoff, K., Sane, A.P., Meurer, J., Driemel, C., Plucken, H., Klaff, P., Stein, B., Bechtold, N. and Westhoff, P. (2001) The nucleus-encoded HCF107 gene of Arabidopsis provides a link between intercistronic RNA processing and the accumulation of translation-competent psbH transcripts in chloroplasts. *Plant Cell*, **13**, 2127-2141.

Field, J., Nikawa, J., Broek, D., MacDonald, B., Rodgers, L., Wilson, I.A., Lerner, R.A. and Wigler, M. (1988) Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method. *Mol Cell Biol*, **8**, 2159-2165.

Genty, B., Briantais, J.M. and Baker, N.R. (1989) The relationship between quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochim Biophys Acta*, **990**, 87-92.

Jefferson, R.A. (1987) Assaying chimeric genes in plants: The GUS gene fusion system. *Plant Mol Biol Rep*, **5**, 387–405

Jefferson, R.A., Kavanagh, T.A. and Bevan, M.W. (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J*, **6**, 3901-3907.

Junttila, M.R., Saarinen, S., Schmidt, T., Kast, J. and Westermarck, J. (2005) Single-step Strep-tag purification for the isolation and identification of protein complexes from mammalian cells. *Proteomics*, **5**, 1199-1203.

Karimi, M., Depicker, A. and Hilson, P. (2007) Recombinational cloning with plant gateway vectors. *Plant Physiol*, **145**, 1144-1154.

Kosugi, S., Ohashi, Y., Nakajima, K. and Arai, Y. (1990) An improved assay for beta-glucuronidase in transformed cells; Methanol almost completely suppresses a putative endogenous beta-glucuronidase activity. *Plant Sci*, **70**, 133–140.

Kuras, R., Saint-Marcoux, D., Wollman, F.A. and de Vitry, C. (2007) A specific c-type cytochrome maturation system is required for oxygenic photosynthesis. *Proc Natl Acad Sci U S A*, **104**, 9906-9910.

Landy, A. (1989) Dynamic, structural, and regulatory aspects of lambda site-specific recombination. *Annu Rev Biochem*, **58**, 913-949.

Laufs, P., Coen, E., Kronenberger, J., Traas, J. and Doonan, J. (2003) Separable roles of UFO during floral development revealed by conditional restoration of gene function. *Development*, **130**, 785-796.

Lezhneva, L., Kuras, R., Ephritikhine, G. and de Vitry, C. (2008) A novel pathway of cytochrome c biogenesis is involved in the assembly of the cytochrome b6f complex in arabidopsis chloroplasts. *J Biol Chem*, **283**, 24608-24616.

Lyska, D., Paradies, S., Meierhoff, K. and Westhoff, P. (2007) HCF208, a homolog of Chlamydomonas CCB2, is required for accumulation of native cytochrome b6 in Arabidopsis thaliana. *Plant Cell Physiol*, **48**, 1737-1746.

Meurer, J., Plucken, H., Kowallik, K.V. and Westhoff, P. (1998) A nuclear-encoded protein of prokaryotic origin is essential for the stability of photosystem II in Arabidopsis thaliana. *Embo J*, **17**, 5286-5297.

Odell, J.T., Nagy, F. and Chua, N.H. (1985) Identification of DNA sequences required for activity of the cauliflower mosaic virus 35S promoter. *Nature*, **313**, 810-812.

Pang, S.Z., DeBoer, D.L., Wan, Y., Ye, G., Layton, J.G., Neher, M.K., Armstrong, C.L., Fry, J.E., Hinchee, M.A. and Fromm, M.E. (1996) An improved green fluorescent protein gene as a vital marker in plants. *Plant Physiol*, **112**, 893-900.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M. and Seraphin, B. (1999) A generic protein purification method for protein complex characterization and proteome exploration. *Nat Biotechnol*, **17**, 1030-1032.

Rohila, J.S., Chen, M., Cerny, R. and Fromm, M.E. (2004) Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J*, **38**, 172-181.

Rubio, V., Shen, Y., Saijo, Y., Liu, Y., Gusmaroli, G., Dinesh-Kumar, S.P. and Deng, X.W. (2005) An alternative tandem affinity purification strategy applied to Arabidopsis protein complex isolation. *Plant J*, **41**, 767-778.

Saint-Marcoux, D., Wollman, F.A. and de Vitry, C. (2009) Biogenesis of cytochrome b6 in photosynthetic membranes. *J Cell Biol*, **185**, 1195-1207.

Sane, A.P., Stein, B. and Westhoff, P. (2005) The nuclear gene HCF107 encodes a membrane-associated R-TPR (RNA tetratricopeptide repeat)-containing protein involved in expression of the plastidial psbH gene in Arabidopsis. *Plant J*, **42**, 720-730.

Sato, M.H. and Wada, Y. (1997) Universal template plasmid for introduction of the triple-HA epitope sequence into cloned genes. *Biotechniques*, **23**, 254-256.

Schagger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem*, **166**, 368-379.

Schult, K., Meierhoff, K., Paradies, S., Toller, T., Wolff, P. and Westhoff, P. (2007) The nuclear-encoded factor HCF173 is involved in the initiation of translation of the psbA mRNA in Arabidopsis thaliana. *Plant Cell*, **19**, 1329-1346.

Shen, Y., Khanna, R., Carle, C.M. and Quail, P.H. (2007) Phytochrome induces rapid PIF5 phosphorylation and degradation in response to red-light activation. *Plant Physiol*, **145**, 1043-1051.

Smith, S.M., Fulton, D.C., Chia, T., Thorneycroft, D., Chapple, A., Dunstan, H., Hylton, C., Zeeman, S.C. and Smith, A.M. (2004) Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in Arabidopsis leaves. *Plant Physiol*, **136**, 2687-2699.

van Leeuwen, W., Ruttink, T., Borst-Vrenssen, A.W., van der Plas, L.H. and van der Krol, A.R. (2001) Characterization of position-induced spatial and temporal regulation of transgene promoter activity in plants. *J Exp Bot*, **52**, 949-959.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V. and Provart, N.J. (2007) An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One*, **2**, e718.

Xu, X., Song, Y., Li, Y., Chang, J., Zhang, H. and An, L. (2010) The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification. *Protein Expr Purif*, **72**, 149-156.

Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. and Gruissem, W. (2004) GENEVESTIGATOR. Arabidopsis microarray database and analysis toolbox. *Plant Physiol*, **136**, 2621-2632.

Table 1: Oligonucleotides used for pAUL vector construction

Oligonucleotide	Sequence (5´to 3´)
pMDC123-H	CCCTCGAGGCGCGCCAAGCTAT
pMDC123-R	GCGCGAGCTCTCGAACCACTTTGTACAAG
T NOS*Pst*I-H	GGCTGCAGCGGAAGATCGTTCAAACATTTG
TNOS*Sac*I/ *Hin*dIII-R	CTATGAAGCTTGAGCTCAATTCGATCTAGTAAC
3xHA*Sac*I-H	GACTGGAGCTCTACCCATATGACGTTCCAGAC
3xHAstop*Pst*I-R	GCACTGCAGTTATCAAGCGTAGTCAGGTACGTC
3xHA*Bam*HI-R2	CAGTGGATCCAGCGTAGTCAGGTACGTCG
HA/STREP-1	P-GATCCTGGTCTCATCCTCAATTCGAAAAGGGTGGA
HA/STREP-2	GAACCTCCACCCTTTTCGAATTGAGGATGAGACCAG
HA/STREP-3	P-GGTTCTGGAGGTGGATCAGGTGGTGGATCTTGG
HA/STREP-4	TGAGACCAAGATCCACCACCTGATCCACCTCCA
HA/STREP-5	P-TCTCATCCTCAATTCGAAAAGTGATAAGAGCTCG
HA/STREP-6	AATTCGAGCTCTTATCACTTTTCGAATTGAGGA
HA/STREP-7	CGGGATCCTGGTCTCATCCTCAATTC
HA/STREP-9	CGACTGCAGTTATCACTTTTCGAATTGAGGA
3xHAStrep*III*X*ba*I-R	GCTCTAGACTTTTCGAATTGAGGATGAGAC
3cIgG-BD*Xba*I-H	ACTCTAGACTGGAAGTTCTGTTCCAGGGGC
3cIgG-BD*Pst*I-R	CGGGCTGCAGTTATCATACCGAACTCGAATTC
35S*Asc*I-H	GTGAGCTCGGCGCGCCAAGCTTGCATGCCTGCAGGTC
35S*Asc*I-R	GCTCTAGAGGCGCGCCCCTCTCCAAATGAAATGAAC
HCF107-Prom-H1	CCGGATTTGGTAGCCACATTCAATGC
HCF107-Prom-R1	CCGGCTCGGGGAAGAAGAATGATGG
HCF107-Prom-H5-2	GGCGCGCCGGATTTGGTAGCCACATTCAATGC
HCF107-Prom-R5	GGCGCGCCGGCTCGGGGAAGAAGAATGATGG
HCF136-Prom-H1	CCCTGTTCATTGGAGTCATATCAAGTC
HCF136-Prom-R1	CCTCTCTTCTCTTTCTCTCTCCCGC
HCF173-Prom-H1	GGCGCGCCTCTCTTACATTTTTGGGCGAACTTG
HCF173-Prom-R1	GGCGCGCCAAATGCATACAATTGTTTGTTAAATGAATC
3xHA-*Asc*I/*Pst*I-H	GCACTGCAGGGCGCGCCATGTACCCATATGACGTTCCAGAC
3xHA-*Pst*I-H	GCACTGCAGTACCCATATGACGTTCCAGAC
3xHA-*Asc*I/*Hin*dIII-R	CTATGAAGCTTGGCGCGCCAGCGTAGTCAGGTACGTCG
Strep-*Xba*I-H1	ACTCTAGATGGTCTCATCCTCAATTCGAAAAGGGTGGAGGTTCTG
Strep-*Asc*I-H2	GACGTACCTGACTACGCTGGCGCGCCATGTGGTCT
Strep-*Sac*I/*Asc*I-H-2	GACTGGAGCTCGGCGCGCCATGTGGTCTCATCCTC
Strep-*Pst*I-R	CGGGCTGCAGCTTTTCGAATTGAGGATGAGAC
Strep-*Asc*I-R	CTGGAACAGAACTTCCAGGGCGCGCCCTTTTCG
IgG-*Asc*I/*Bst*XI-H	GACTGCCACCGCGGTGGCGCGCCATGGCCACCATGGCGCAAC
IgG-3C-*Bcu*I-R2	GACACTAGTGGGCCCCTGGAACAGAACTTCCAG

Table 2: Oligonuleotides used for cloning of target genes

Oligonucleotide	Sequence (5´to 3´)
GUS+Term-H	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGTTACGTCCTGTAGAAACC
GUS+Term-R	GGGGACCACTTTGTACAAGAAAGCTGGGTCGATCTAGTAACATAGATGACA
start107 attB1-H	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGCACTTCTTCTTCGTGCCG
107 attB2-R	GGGGACCACTTTGTACAAGAAAGCTGGGTCAGCACCATTTATTCTTCCTC
start208 attB1-H	GGGGACAAGTTTGTACAAAAAAGCAGGCTATGAGTATTCAAATTTGTAATTTC
208 attB2-R	GGGGACCACTTTGTACAAGAAAGCTGGGTCACCTCTGAATTTCTCAGCCATAG

Figure Legends

Figure 1: Schematic illustration of the Gateway compatible pAUL destination vector series, showing expression cassettes.

(A) C-terminal fusion vectors pAUL1-16. Expression is driven by either 2x p*35S CaMV* or endogenous promoter sequences from *A. thaliana* (p*HCF107*, p*HCF136*, p*HCF173*): pAUL1-3 and pAUL13 carry *p35S CaMV*; pAUL4-6 and pAUL14 carry p*HCF107*; pAUL7-9 and pAUL15 carry p*HCF136*; pAUL10-12 and pAUL16 carry p*HCF173*. Protein tags are: 3xHA single tag (pAUL1, 4, 7, 10); Strep-tag*III* single tag (pAUL13-16); 3xHA/Strep-tag*III* double tag (pAUL2, 5, 8, 11); and 3xHA/Strep-tag*III*/ProtA triple tag + 3C protease cleavage site (pAUL3, 6, 9, 12). (B) N-terminal fusion vectors pAUL17-20. Vectors carry coding sequences for 3xHA single tag (pAUL17); 3xHA/Strep-tag*III* double tag (pAUL18); 3xHA/Strep-tag*III*/ProtA triple tag + 3C protease cleavage site (pAUL19); and Strep-tag*III* single tag (pAUL20).

Figure 2: Characterization of promoters 2x p*35S CaMV*, p*HCF107*, p*HCF136*, and p*HCF173* fused to the *GUS* reporter gene.

(A) GUS staining of 5-day-old transgenic *A. thaliana* seedlings. (B) Histochemical localization of GUS activity in leaf sections of 3-week-old transgenic *A. thaliana* plants. UE, upper epidermis; PM, palisade mesophyll; SM, spongy mesophyll; LE, lower epidermis. (C) GUS activities in transgenic *A. thaliana* lines. In each case, 10 independent transgenic lines were tested 15 or 30 days after germination. Median values are shown as black bars and indicated at the top of each column. MU, 4-methylumbelliferone.

Figure 3: Complementation analysis of *hcf208*and *hcf107.2* with a representative pAUL vector set.

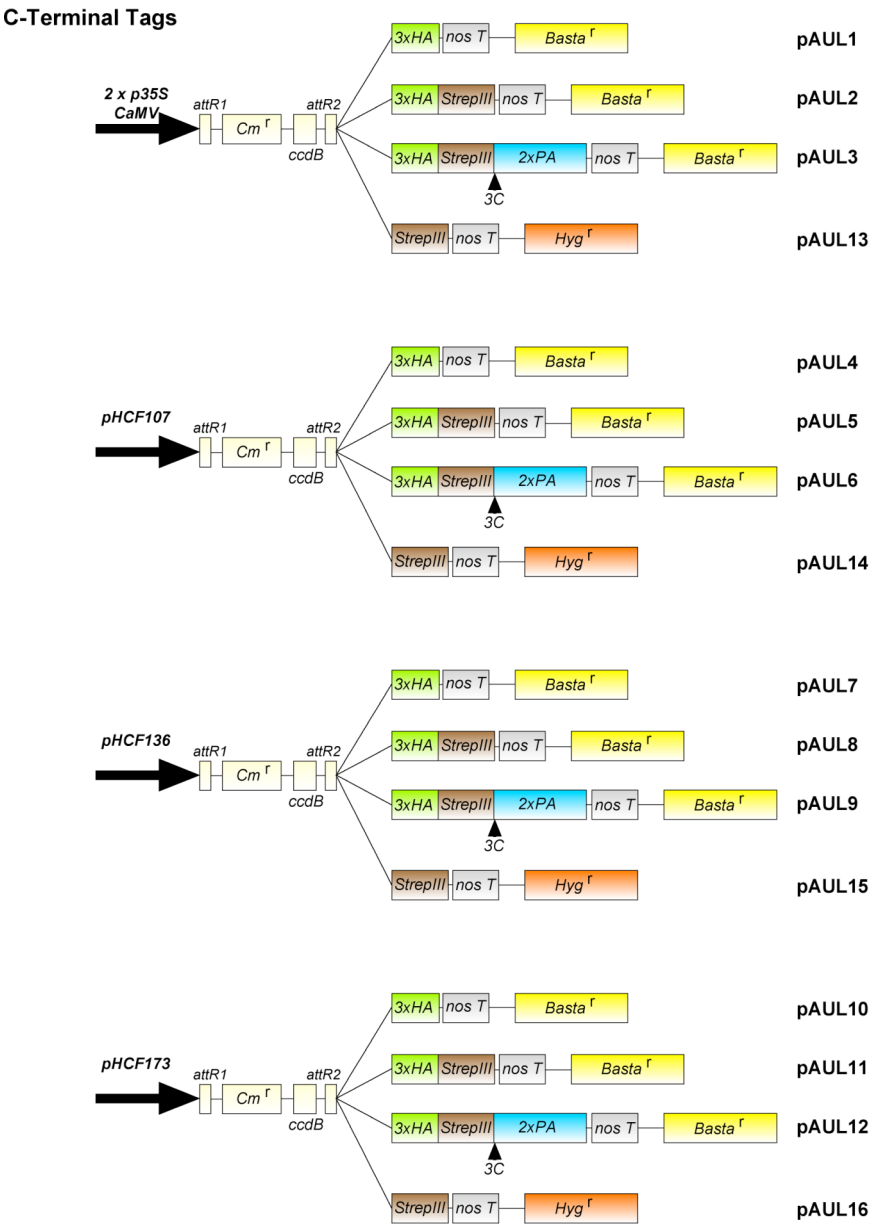
(A) Schematic illustration of promoter/ cDNA/ tag combinations generated for transformation of *hcf107.2* and *hcf208*. (B) Fluorometric analysis of HCF208- and HCF107 complemented plants, wild type and *hcf208* / *hcf107.2* mutant plants. Pseudo-color images of maximum quantum efficiency of photosystem II (Fv/Fm) are displayed for HCF107 and of photochemical quenching efficiency (qP) are displayed for HCF208. 3 independent transformants were tested for each construct. Values for each line investigated are illustrated in diagrams. (C) Western blot analysis of complemented lines, wild type and mutant plants. 50 μg of crude protein extract were loaded. Membranes were decorated with Anti-HA-Peroxidase antibody for HCF208 and HCF107; HCF107 was also visualized by an HCF107-specific antibody.

Figure 4: One step and tandem-purification of HCF208

100 (A) or 200 μg (B, C) chlorophyll aliquots of solubilized membrane proteins were applied for purification. Aliquots of 20 μg chlorophyll from extracts and total amounts of eluates were separated by SDS-PAGE, transferred to a nitrocellulose membrane and immunodecorated with antibodies against the HA tag (Anti-HA-Peroxidase) and ATP-Synthase as a control.

(A) One step purification of proteins from wild type and HCF208pAUL1 via the HA epitope and competitive elution. (B) Tandem purification of proteins from wild type and HCF208pAUL2 via Strep-tag*III* and 3xHA (C) Tandem purification of proteins from wild type and HCF208pAUL3 via ProtA tag + 3C protease cleavage and Strep-tag*III*.

(a)


C-Terminal Tags
2 x p35S CaMV
attR1
Cm r
attR2
ccdB
3xHA
nos T
Basta r
StrepIII
2xPA
3C
Hyg r
pAUL1
pAUL2
pAUL3
pAUL13
pHCF107
pAUL4
pAUL5
pAUL6
pAUL14
pHCF136
pAUL7
pAUL8
pAUL9
pAUL15
pHCF173
pAUL10
pAUL11
pAUL12
pAUL16

(b)

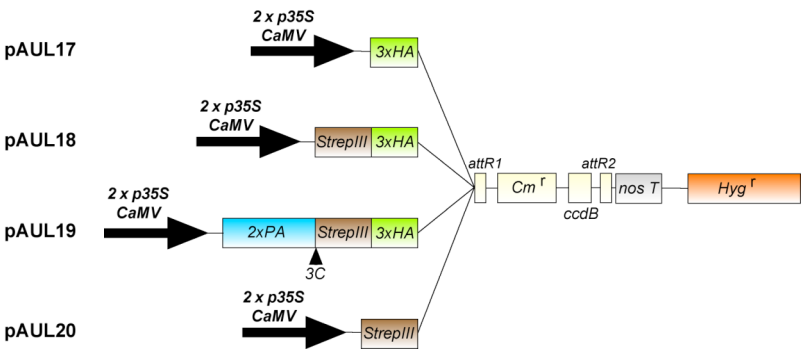

N-Terminal Tags
pAUL17
pAUL18
pAUL19
pAUL20
2 x p35S CaMV
3xHA
StrepIII
2xPA
3C
attR1
Cm r
attR2
ccdB
nos T
Hyg r

Figure 1

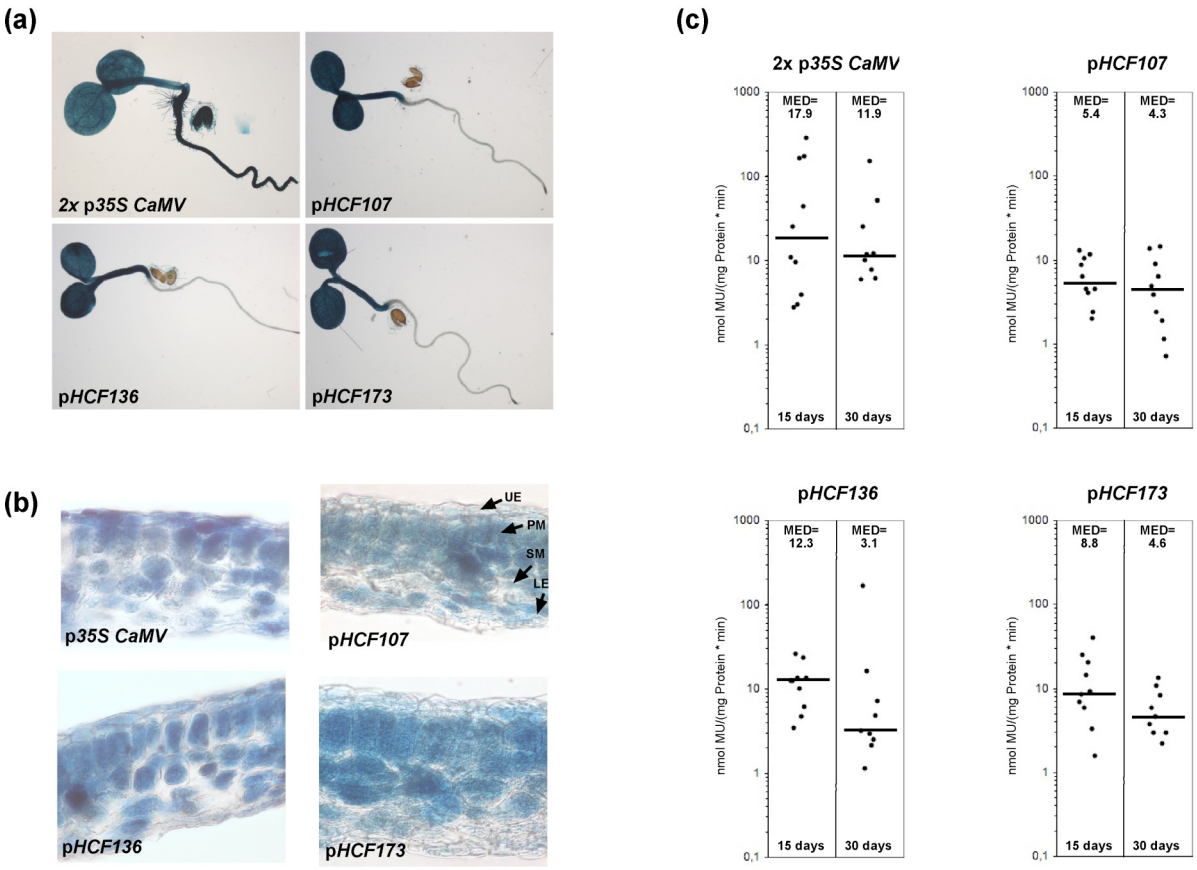

(a)
2x p35S CaMV
pHCF107
pHCF136
pHCF173
(b)
p35S CaMV
pHCF107
UE
PM
SM
LE
pHCF136
pHCF173
(c)
2x p35S CaMV
pHCF107
pHCF136
pHCF173
nmol MU/(mg Protein * min)
MED= 17.9
MED= 11.9
MED= 5.4
MED= 4.3
MED= 12.3
MED= 3.1
MED= 8.8
MED= 4.6
15 days
30 days
1000
100
10
1
0,1

Figure 2

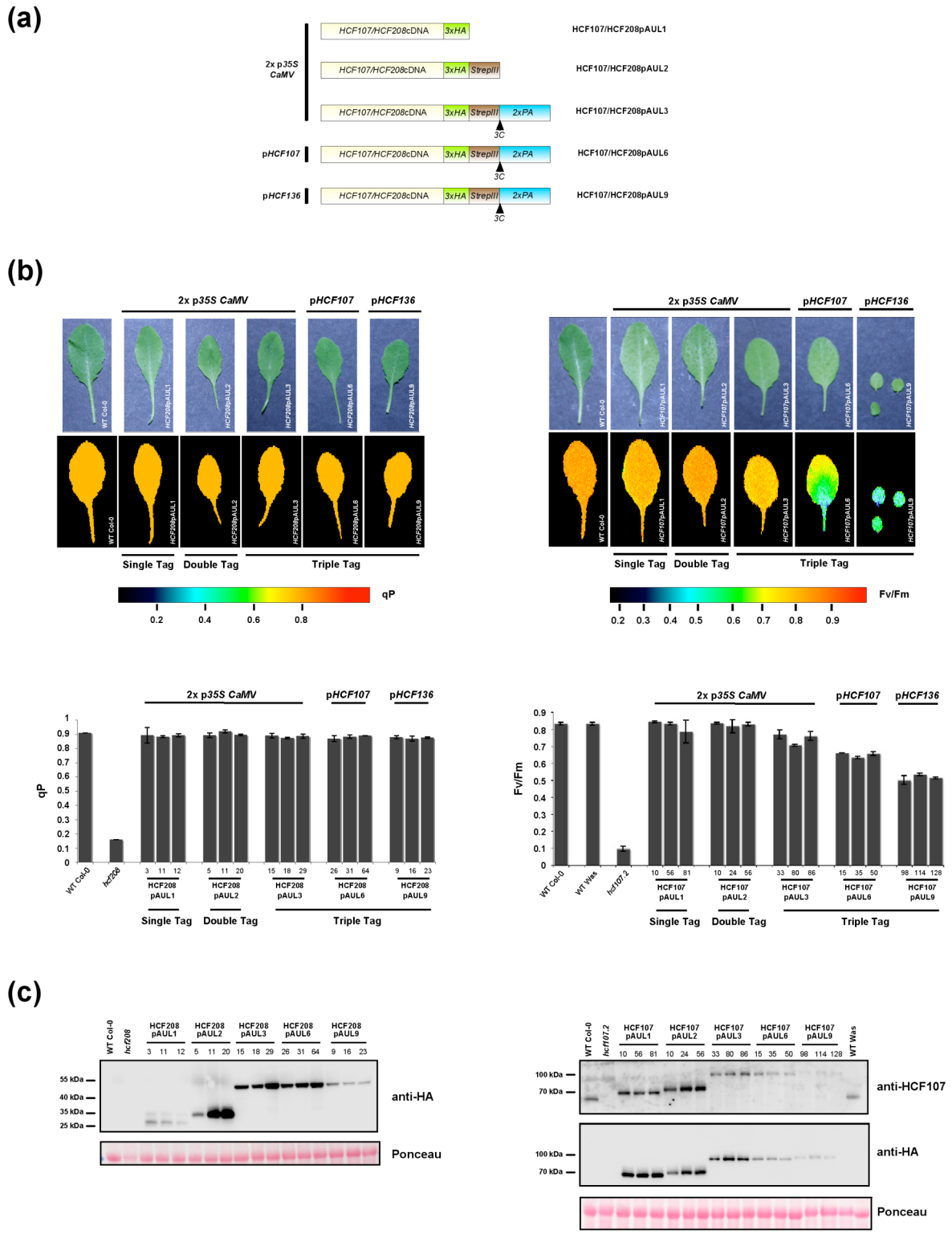

(a)
2x p35S CaMV
pHCF107
pHCF136
HCF107/HCF208cDNA
3xHA
StrepIII
2xPA
3C
HCF107/HCF208pAUL1
HCF107/HCF208pAUL2
HCF107/HCF208pAUL3
HCF107/HCF208pAUL6
HCF107/HCF208pAUL9
(b)
2x p35S CaMV
pHCF107
pHCF136
WT Col-0
HCF208pAUL1
HCF208pAUL2
HCF208pAUL3
HCF208pAUL6
HCF208pAUL9
HCF107pAUL1
HCF107pAUL2
HCF107pAUL3
HCF107pAUL6
HCF107pAUL9
Single Tag
Double Tag
Triple Tag
qP
0.2 0.4 0.6 0.8
Fv/Fm
0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9
qP
Fv/Fm
WT Col-0
hcf208
WT Was
hcf107.2
3 11 12
5 11 20
15 18 29
26 31 64
9 16 23
10 56 81
10 24 56
33 80 86
15 35 50
98 114 128
HCF208 pAUL1
HCF208 pAUL2
HCF208 pAUL3
HCF208 pAUL6
HCF208 pAUL9
HCF107 pAUL1
HCF107 pAUL2
HCF107 pAUL3
HCF107 pAUL6
HCF107 pAUL9
(c)
55 kDa
40 kDa
35 kDa
25 kDa
anti-HA
Ponceau
100 kDa
70 kDa
anti-HCF107
anti-HA
Ponceau

Figure 3

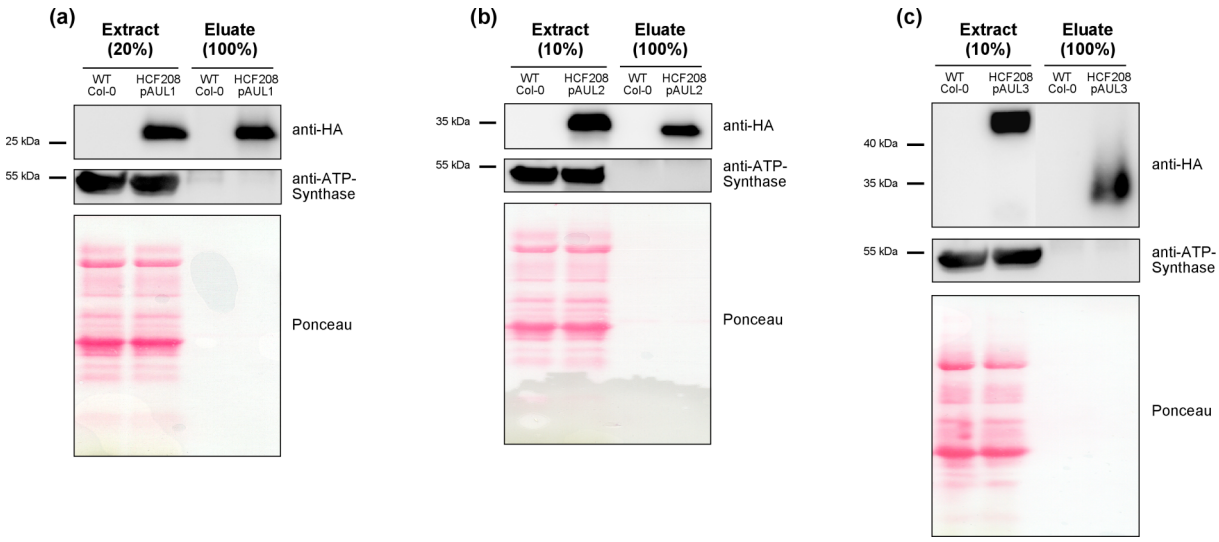

(a)
Extract (20%)
Eluate (100%)
WT Col-0
HCF208 pAUL1
WT Col-0
HCF208 pAUL1
25 kDa
55 kDa
anti-HA
anti-ATP-Synthase
Ponceau
(b)
Extract (10%)
Eluate (100%)
WT Col-0
HCF208 pAUL2
WT Col-0
HCF208 pAUL2
35 kDa
55 kDa
anti-HA
anti-ATP-Synthase
Ponceau
(c)
Extract (10%)
Eluate (100%)
WT Col-0
HCF208 pAUL3
WT Col-0
HCF208 pAUL3
40 kDa
35 kDa
55 kDa
anti-HA
anti-ATP-Synthase
Ponceau

Figure 4

Authors´ Contributions

DL performed: cloning of C-terminal tags and the 35S CaMV promoter, generation of HCF107- and HCF208 transgenic lines, quantitative GUS assays and histochemical analyses of seedlings, characterization of complemented plants, purification of HCF208 via tags. DL wrote the paper.
KS carried out cloning of N-terminal tags and endogenous promoters and generated GUS-lines. KM and PW participated in drafting of the manuscript.

- Part B -

Manuscript 2

HCF208, a homolog of *Chlamydomomas* CCB2, is required for accumulation of native cytochrome b_6 in *Arabidopsis thaliana*

Plant Cell Physiol. 48(12): 1737–1746 (2007)
doi:10.1093/pcp/pcm146, available online at www.pcp.oxfordjournals.org

HCF208, a Homolog of *Chlamydomonas* CCB2, is Required for Accumulation of Native Cytochrome b_6 in *Arabidopsis thaliana*

Dagmar Lyska, Susanne Paradies, Karin Meierhoff * and Peter Westhoff

Institut für Entwicklungs- und Molekularbiologie der Pflanzen, Heinrich-Heine-Universität, Universitätsstr. 1, D-40225 Düsseldorf, Germany

The cytochrome b_6 subunit of the cytochrome b_6f complex is a multiheme protein. Two b-type hemes are bound non-covalently to the protein, whereas the third heme (heme c_n) is covalently attached via an atypical thioether bond. To understand the maturation of cytochrome b_6 and to identify the assisting factors, we characterized the ethyl methanesulfonate-induced nuclear mutant *hcf208*. This *Arabidopsis* mutant shows a high chlorophyll fluorescence phenotype and does not accumulate the major cytochrome b_6f complex subunits. Transcript levels and patterns of the four major polypeptides of the complex are equal to those of the wild type. The mutant cytochrome b_6 polypeptide shows a faster migration behavior in SDS–PAGE compared with the wild type and it has no peroxidase activity. The *HCF208* locus was mapped and the gene was cloned. Sequence analysis revealed that HCF208 is a homolog of the *Chlamydomonas reinhardtii* CCB2 protein, which is a factor mediating attachment of heme c_n to the cytochrome b_6 polypeptide as part of a novel heme biogenesis pathway, called system IV. Blue Native PAGE revealed residual amounts of the cytochrome b_6f complex dimer in *hcf208*; however, this form is unable to participate in electron transport reactions.

Keywords: *Arabidopsis* — Chloroplast — Cytochrome b_6f complex — Heme attachment — Heme c_n.

Abbreviations: BAC, bacterial artificial chromosome; CCB: C = cofactor binding, C = cytochrome b_6f complex, B = subunit PetB; EMS, ethyl methanesulfonate; HCF, high chlorophyll fluorescence; PFD, photon flux density.

Introduction

The cytochrome b_6f complex plays a central role in electron transport in oxygenic photosynthesis. By its plastoquinol–plastocyanin oxidoreductase activity, the complex mediates electron transfer between PSI and PSII. In this linear process, protons are translocated through the thylakoid membrane, leading to the generation of an electrochemical gradient that is used for synthesis of ATP by the ATP synthase (Junge 1999, Allen 2002). In contrast to its mitochondrial and bacterial homolog, complex bc_1, the cytochrome b_6f complex is also involved in ferredoxin-dependent cyclic electron transport around PSI (Joliot and Joliot 2006).

The functional form of the cytochrome b_6f complex is a dimer (Huang et al. 1994, Breyton et al. 1997, Mosser et al. 1997). Each monomer consists of eight subunits: the four large polypeptides cytochrome *f* (PetA), cytochrome b_6 (PetB), Rieske FeS protein (PetC) and subunit IV (PetD), and the small polypeptides PetG, PetL, PetM and PetN. PetG, PetL and PetN were shown to play a role in assembly, stability and dimerization of the cytochrome b_6f complex in tobacco (Hager et al. 1999, Schöttler et al. 2007, Schwenkert et al. 2007). The role of PetM in regulatory processes mediated by the cytochrome b_6f complex has been studied in *Synechocystis* (Schneider et al. 2001), but not yet in higher plants. *PetC* and *PetM* are nuclear encoded, whereas the other subunits are plastid encoded. In spinach, the ferredoxin $NADP^+$ oxidoreductase could also be isolated as part of the cytochrome b_6f complex, indicating a role in ferredoxin-dependent cyclic electron transport (Zhang et al. 2001, Whitelegge et al. 2002).

Seven cofactors per monomer are associated with the cytochrome b_6f complex: one molecule of chlorophyll *a*, one molecule of β-carotene (Zhang et al. 1999, Pierre et al. 2003), one [2FE–2S] cluster bound to the Rieske protein, one heme covalently attached to cytochrome *f* and three heme groups in cytochrome b_6. Two of the heme groups, heme b_n at the electronegative (stromal) and heme b_p at the electropositive (lumenal) side of the complex, are non-covalently attached to cytochrome b_6 via pairs of histidines. Binding of these b hemes is supposed to be spontaneous (Robertson et al. 1994). The third, covalently bound heme (heme c_n or haem c_i) was recently identified in the crystal structures of the cytochrome b_6f complex of the thermophilic cyanobacterium *Mastigocladus laminosus* (Kurisu et al. 2003) and the green alga *Chlamydomonas reinhardtii* (Stroebel et al. 2003). The atypical heme was found to be located on the stromal side (Q_i site) of the complex where plastoquinone is reduced (Kurisu et al. 2003, Stroebel et al. 2003). It is supposed to be involved in cyclic electron transfer (Kurisu et al. 2003, Stroebel et al. 2003, Cramer and Zhang 2006, Yamashita et al. 2007) or Q_i site turnover (Baymann et al. 2007). In contrast to typical c-type hemes, which form two thioether bonds to a consensus sequence

*Corresponding author: E-mail, karin.meierhoff@uni-duesseldorf.de; Fax, +49-211-8114871.

(CxxCH), heme c_n is bound to a single Cys35 from cytochrome b_6 and has no axial amino acid ligands. Moreover, it is liganded by a water or hydroxide ion in a heme b_n propionate group and by plastoquinone (Kurisu et al. 2003, Stroebel et al. 2003, Yamashita et al. 2007). Hemes c_n and b_n strongly interact, as shown in electron paramagnetic resonance (EPR) spectroscopy experiments (Zatsman et al. 2006, Baymann et al. 2007).

Four nuclear-encoded factors, CCB1–CCB4, were found to be necessary for binding of heme c_n to cytochrome b_6 in *Chlamydomonas* (Kuras et al. 2007). These factors display a new c-type cytochrome maturation system called system IV. In spite of the identification of CCB1–CCB4, the pathway of heme c_n binding to cytochrome b_6 mediated by these factors remains unclear. However, as reported by Kuras et al. (2007), CCB1–CCB4 are conserved among all organisms performing oxygenic photosynthesis.

We report on the *Arabidopsis thaliana* homolog to CCB2 called HCF208. *Hcf208* mutant plants exhibit an *hcf* (high chlorophyll fluorescence) phenotype indicating an impaired photosynthetic electron flow. The four major subunits of the cytochrome b_6f complex are drastically reduced. The migration behavior of the cytochrome b_6 polypeptide is altered compared with the wild-type protein and it cannot be detected by its peroxidase activity. Nonetheless, the dimeric form of the cytochrome b_6f complex can be assembled in *hcf208* mutant seedlings.

Results

Electron transport between PSII and PSI is affected in hcf208

The mutant *hcf208* was selected from a collection of ethyl methanesulfonate (EMS)-induced mutants by its high chlorophyll fluorescence phenotype. *Hcf208* segregated as a single recessive nuclear mutation. Homozygous mutant seedlings of *hcf208* were unable to grow photoautotrophically on soil and were therefore grown on sucrose-supplemented media under light conditions of $\sim$50–70 $\mu mol\, s^{-1}\, m^{-2}$.

Measurements of chlorophyll fluorescence induction (Fig. 1A) and P700 redox kinetics (Fig. 1B) are used to determine the photosynthetic capacity of mutant plants. Chlorophyll fluorescence induction discloses the photosynthetic capacity of PSII. Its activity is reflected by the ratio of variable to maximum chlorophyll fluorescence (F_v/F_m; where $F_v = F_m - F_o$ is the variable fluorescence, F_m the maximum fluorescence, and F_o the minimum fluorescence). This value was reduced in the mutant *hcf208* ($F_v/F_m = 0.69 \pm 0.04$) compared with the wild type ($F_v/F_m = 0.78 \pm 0.01$), but still indicated an active PSII in the mutant *hcf208*. In contrast the fraction of PSII reaction centers in the open

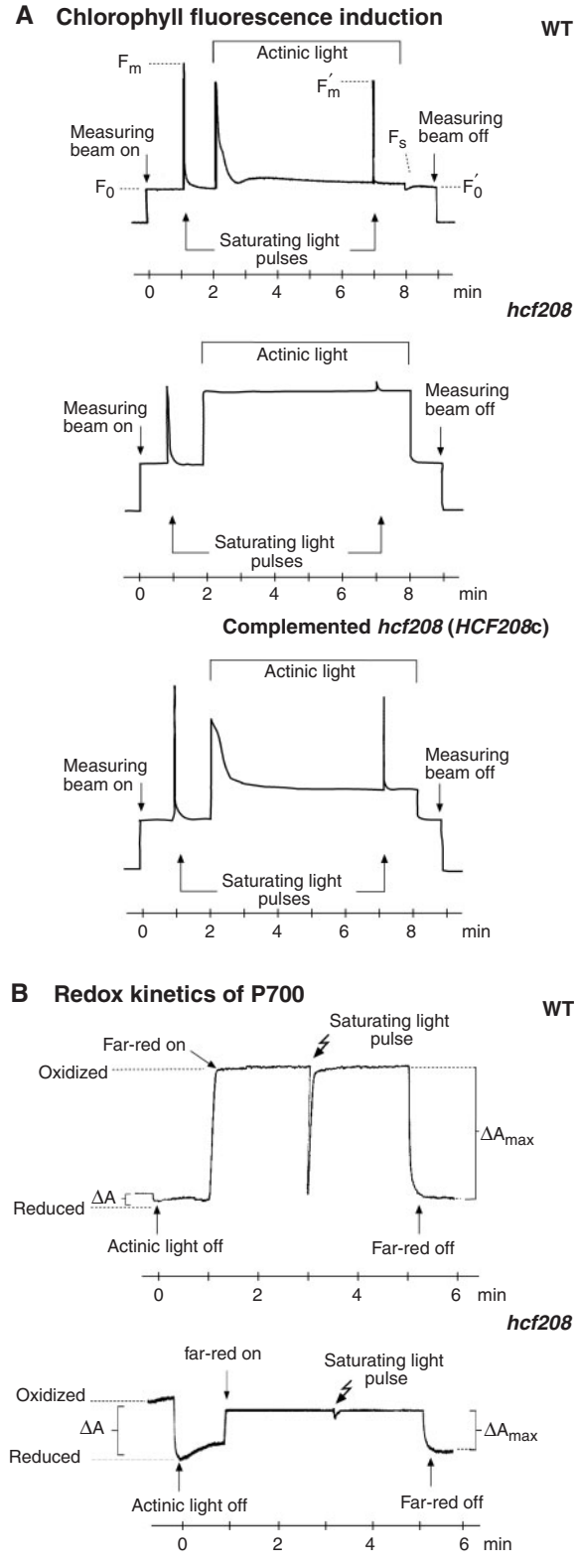


Fig. 1 Spectroscopic analyses of wild-type, *hcf208* mutant and complemented *hcf208* plants. Chlorophyll fluorescence induction (A) and P700 absorbance (B) measurement of 2- to 3-week-old plants.

state $\{qL = [(F'_m - F_s)/(F'_m - F'_o)] \times (F'_o/F_s)$; where F_s is steady-state fluorescence; Kramer et al. 2004$\}$ was drastically reduced in *hcf208* ($qL_{208} = 0.07 \pm 0.008$; $qL_{WT} = 0.95 \pm 0.06$), suggesting that electrons accumulated in the plastoquinone pool and thus were not transported downstream of the photosynthetic electron transport chain. To determine whether a defect in intersystem electron transport or in PSI is responsible for this blockade, we studied redox activity of the PSI reaction center by measuring the absorbance kinetics of P700 at 820–870 nm. In wild-type plants, P700 was reduced in the dark and became completely oxidized under illumination with far-red light. Under steady-state conditions in actinic light, the redox state of P700 was adjusted in between. A white light pulse in a far-red background led to a short re-reduction of P700 by electrons released from PSII. In *hcf208*, P700 was oxidized under illumination with actinic light to its maximum level, indicating the disruption of electron transport between PSII and PSI. A clear absorbance change induced by far-red light suggested the integrity of PSI. However, we could not observe a re-reduction of P700 by a saturating light pulse, which is indicative for a block in intersystem electron transport between the photosystems.

The spectroscopic measurements revealed integrity of PSII and PSI but an impaired electron transport between the two systems, indicating a defect of the intersystem electron transport chain, potentially at the level of the cytochrome b_6f complex, in *hcf208*.

Subunits of the cytochrome b_6f complex do not accumulate in hcf208

The block in the intersystem electron transport found in *hcf208* could be the result of reduced levels of the cytochrome b_6f complex. Thus, immunoblot analyses with antisera against the four major subunits of the complex were performed (Fig. 2A). We found that all tested subunits accumulated to about 12.5% of wild-type levels in *hcf208*. Moreover, the cytochrome b_6 subunit in *hcf208* migrated faster than in the wild type, indicating a change in the molecular weight of this protein. No cytochrome b_6 polypeptide in the range of the wild-type protein could be detected even after an extended exposure (Fig. 2B). Furthermore, we tested heme-associated peroxidase activities of the cytochrome b_6 and cytochrome *f* holoproteins. Fig. 2C shows that peroxidase activity of cytochrome b_6 cannot be detected in the *hcf208* mutant. Thus, the remaining 12.5% of the cytochrome b_6 polypeptide does not contain a heme group that is able to react with the enhanced chemiluminescence (ECL) components. In contrast, the cytochrome *f* polypeptide was detected by peroxidase activity at <6% compared with the wild type. As described above, the cytochrome *f* polypeptide accumulates to only about 12.5% of wild-type levels. Hence, <50% of the remaining cytochrome *f* polypeptide is present as mature holoprotein in the mutant. Nevertheless, a shifted cytochrome *f* protein band was not detectable.

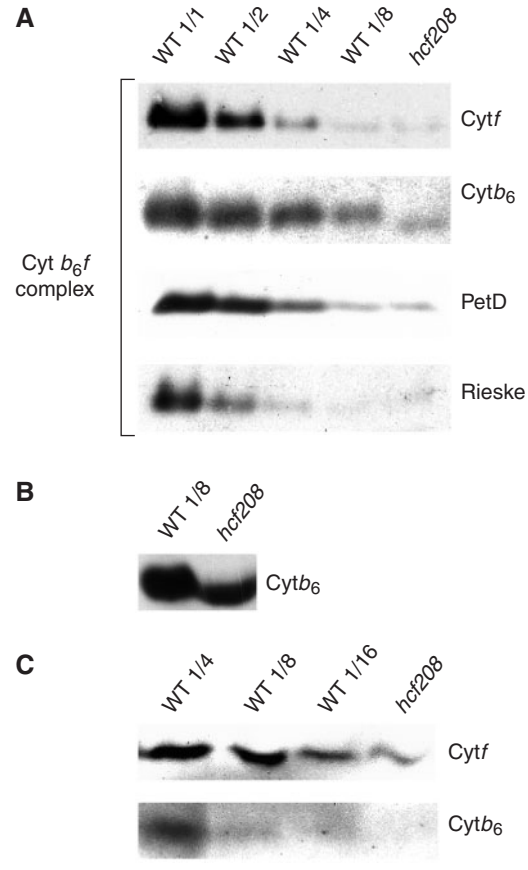


Fig. 2 Analysis of the cytochrome b_6f complex subunits from *hcf208* and wild type. Proteins were separated by SDS–PAGE and transferred to a nitrocellulose membrane. (A) Immunoblot analysis of the subunits cytochrome *f*, cytochrome b_6, PetD and the Rieske FeS protein. Lanes were loaded with 20 µg (WT 1/1 and *hcf208*), 10 µg (WT 1/2), 5 µg (WT 1/4) and 2.5 µg (WT 1/8) of protein. (B) Immunoblot analysis of the cytochrome b_6 protein. Lanes were loaded with 2.5 µg (WT 1/8) and 20 µg (*hcf208*) of protein and the nitrocellulose membrane was exposed five times longer than in (A) to visualize traces of mature cytochrome b_6. (C) Analysis of the peroxidase activity of the heme-binding proteins cytochrome *f* and cytochrome b_6. Aliquots of 25 µg (WT1/4), 12.5 µg (WT1/8), 6.25 µg (WT1/16) and 100 µg (*hcf208*) of protein were loaded. Peroxidase activity was detected by enhanced chemiluminescence.

We also tested representative subunits of PSII (D1, D2, CP47), PSI (PsaA/B) and ATP synthase (alpha-SU; Fig. 3). Polypeptides of PSII and the α-subunit of the ATP synthase accumulated to wild-type levels. The PsaA/B subunit was found to be reduced to 50% of wild-type amounts.

Taken together, spectroscopic and immunoblot analyses indicated that the mutation in *hcf208* leads to an impaired accumulation of the cytochrome b_6f complex. We assume that the mutation in *hcf208* primarily affects the

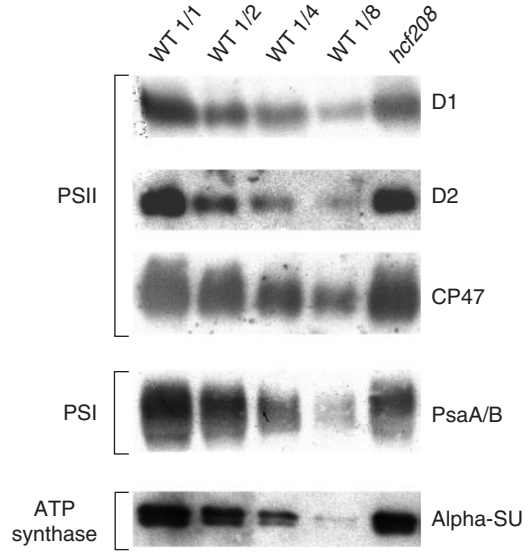


Fig. 3 Immunoblot analysis of subunits from PSII (D1, D2 and CP47), PSI (PsaA/B) and ATP-synthase (alpha-subunit). Aliquots of 20 µg (WT 1/1 and *hcf208*), 10 µg (WT 1/2), 5 µg (WT 1/4) and 2.5 µg (WT 1/8) of protein were separated by SDS–PAGE and transferred to a nitrocellulose membrane.

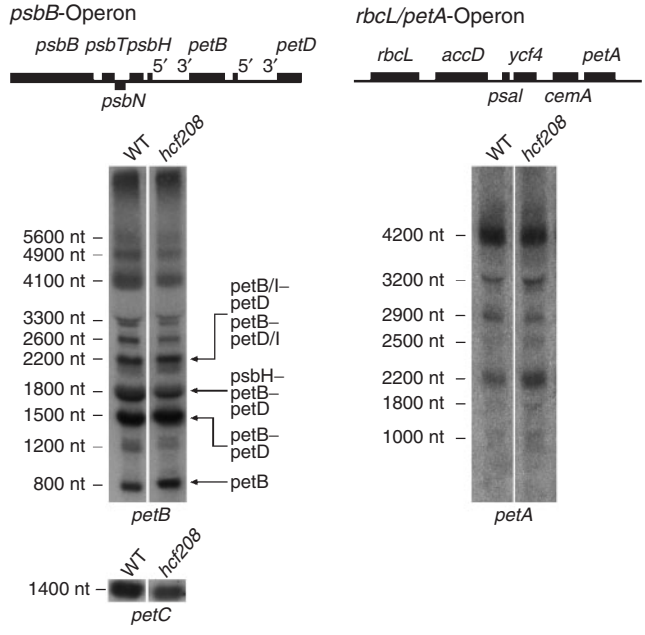


Fig. 4 RNA gel blot analyses of *petB*, *petA* and *petC* transcripts from *hcf208* and the wild type. A 5 µg aliquot of total leaf RNA was loaded per lane. Northern hybridization was performed as described in Materials and Methods. Sizes of the transcripts and transcripts detected with the *petB* probe containing the *petD* transcript are indicated. The introns of *petB* and *petD* are marked as I.

cytochrome b_6 subunit since its migration behavior is altered in comparison with the wild-type protein, and no peroxidase activity of this protein is detectable in the mutant.

Transcripts of the cytochrome b_6f complex subunits are not affected in hcf208

The reduction of the cytochrome b_6f complex may be caused either by reduced levels of the corresponding mRNAs or by decreased synthesis rates or increased degradation rates of the proteins. To differentiate between these possibilities, RNA gel blot hybridizations were performed using probes for *petA*, *petB* and *petC*.

The plastid-encoded genes *petA* (cytochrome *f*), *petB* (cytochrome b_6) and *petD* (PetD) are transcribed in polycistronic transcription units, resulting in complex RNA patterns (Barkan 1988, Westhoff and Herrmann 1988, Gray 1992, Yamazaki et al. 2004). *PetD* transcripts can be co-detected by the *petB* probe since both genes are part of the *psbB–psbT–psbH–petB–petD* transcription unit, and all *petD* transcripts also contain the *petB* transcript. The nuclear-encoded *PetC* (Rieske protein) is transcribed into a monocistronic transcript.

In *hcf208* transcript levels and patterns of the investigated genes were equal to those of the wild type (Fig. 4), suggesting that transcription, RNA processing and RNA stability are not affected. Thus, reduced amounts of the four large cytochrome b_6f complex subunits are caused by increased rates of protein degradation and/or reduced protein synthesis rates.

The cytochrome b_6f complex dimer accumulates in hcf208

The functional form of the cytochrome b_6f complex is a dimer (Huang et al. 1994, Breyton et al. 1997, Mosser et al. 1997). To determine whether the dimer accumulates in the absence of mature cytochrome b_6 protein, non-denaturing Blue Native PAGE (Schagger et al. 1994) was performed, and this first dimension was followed by separation of the complexes on a denaturing SDS gel. Subsequently proteins were blotted and immunodecorated using a cytochrome *f* antibody. As shown in Fig. 5, in both wild-type and *hcf208* mutant seedlings the cytochrome b_6f dimer was detectable in the range of about 250–300 kDa, whereas the monomeric form was found in the range of 140 kDa. The cytochrome b_6f monomer is a breakdown product occurring during preparation. Increasing detergent concentration causes decomposition of the active dimer into the inactive monomeric form (Breyton et al. 1997). Since equal detergent concentrations were used for complex preparation in wild-type and mutant plants, the ratio between detergent and complex was much higher in the

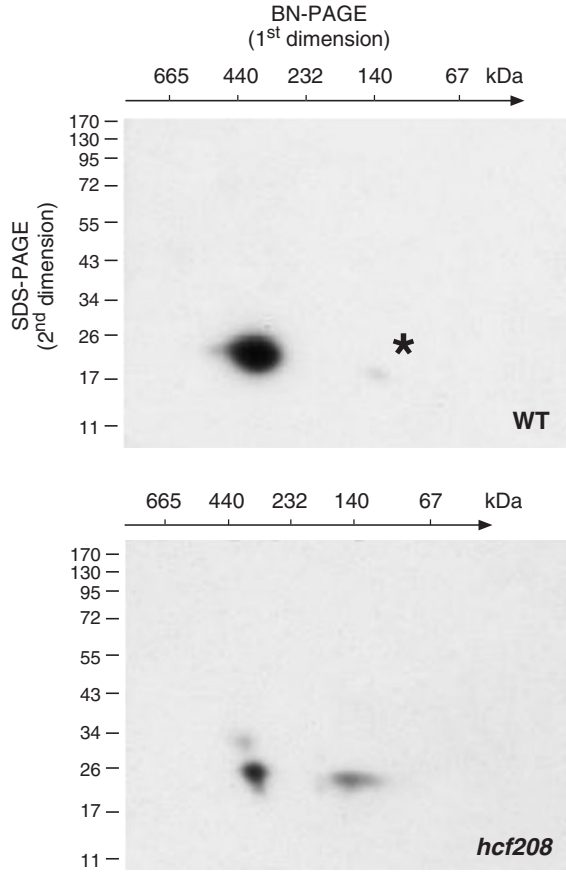


Fig. 5 Blue Native PAGE. Chloroplast membrane protein complexes from 3-week-old wild-type and *hcf208* plants were solubilized and separated by Blue Native PAGE (BN) and subsequent denaturing SDS–PAGE. Proteins were transferred to a nitrocellulose membrane and cytochrome *f* was immunodetected. The asterisk indicates the monomer in the wild type.

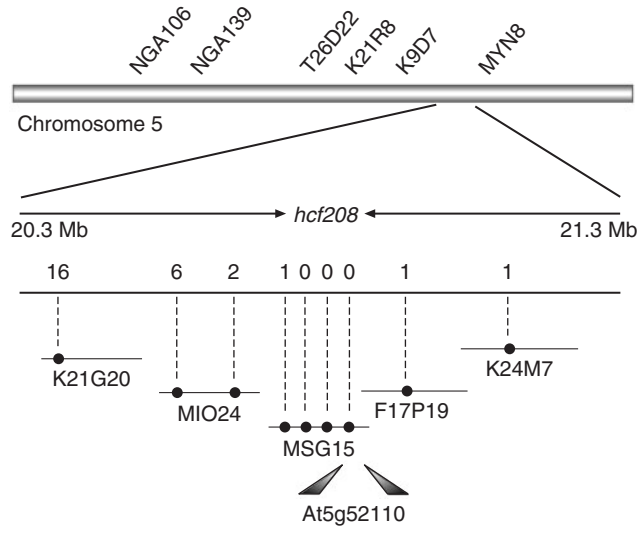


Fig. 6 Mapping of the *HCF208* locus. *HCF208* was mapped to chromosome 5 between the markers MSG15-3/4 on BAC clone MSG15 and F17P19 on BAC clone F11P19. Two genes localized in this region exhibited putative chloroplast transit peptides. Sequence analysis revealed At5g52110 as the *HCF208* locus.

mutant which very probably resulted in a disaggregation of the dimer.

Blue Native PAGE analysis revealed that the remaining subunits of the cytochrome b_6f complex could assemble into a dimeric complex in *hcf208* mutants.

Molecular cloning of HCF208

The mutant *hcf208* was generated by EMS mutagenesis of *Arabidopsis* seeds. Mapping of the *HCF208* gene to the lower arm of chromosome 5 was performed with the help of a set of microsatellite markers (Fig. 6). To determine the *HCF208* locus more precisely, new single sequence length polymorphism (SSLP) and cleaved amplified polymorphic sequence (CAPS) markers were created for genotyping of 320 plants of the mapping population. The region of the mutation locus was narrowed down between the markers MSG15-3/4 on bacterial artificial chromosome (BAC) clone MSG15 (one recombination) and F17P19 on BAC clone F17P19 (one recombination). This region (58 kb) was screened for genes that encode proteins with predicted chloroplast transit peptides. Two putative candidate genes were identified (*At5g52100* and *At5g52110*). Analysis using the Arabidopsis Information Resource database (www.arabidopsis.org) revealed that the *At5g52110* gene encoded a protein with unknown function. To test whether *At5g52110* was affected by the mutation, we amplified the whole gene by PCR using genomic DNA from *hcf208* mutants as a template. Sequence analysis revealed a point mutation in the form of a guanine to adenine exchange in the second exon at position 291 (Fig. 7A). This point mutation results in an exchange from the neutral amino acid glycine to the strongly basic amino acid arginine at position 68. RNA blot hybridization was performed in order to investigate whether the *At5g52110* transcript is present in the mutant. Transcript levels and patterns in the mutant *hcf208* did not differ from those of the wild type (data not shown), indicating that the observed point mutation leads to a non-functional At5g52110 protein.

To confirm that the mutant phenotype of *hcf208* is caused by the disruption of the identified gene *At5g52110*, the isolated cDNA under control of the 35S promoter was transferred into heterozygous *hcf208* plants by *Agrobacterium tumefaciens*-mediated transformation. We identified four independent transformants that were homozygous for the mutation *hcf208* and were able to grow photoautotrophically on soil. Chlorophyll

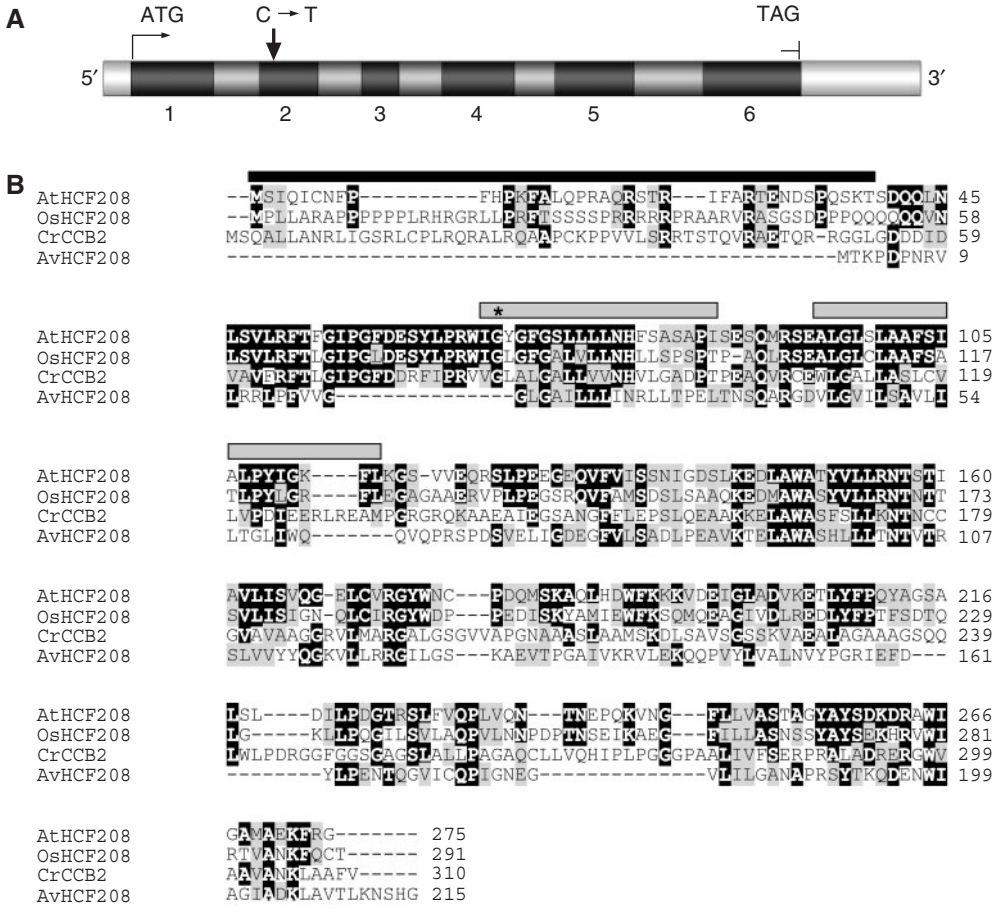


Fig. 7 Gene structure and sequence alignment of *HCF208*. (A) Scheme of *HCF208* and the mutant allele. The *HCF208* open reading frame of 1,585 bp includes six exons separated by five introns. The mutant allele carries a point mutation at position 291 in exon 2, which leads to an amino acid exchange at position 68 in the HCF208 protein. Exons are shown in black, introns are dark gray and untranslated regions are indicated in light gray. (B) Sequence alignment of *Arabidopsis* HCF208 (AtHCF208) and homologs from rice (OsHCF208), *C. reinhardtii* (CrCCB2) and *A. variabilis* (AvHCF208). White letters on a black background indicate amino acids identical to AtHCF208 and black letters on a gray background indicate amino acids similar to AtHCF208. The predicted chloroplast transit peptide of AtHCF208 is indicated by a black bar above the sequences, and the transmembrane domains are indicated by gray bars. The amino acid exchanged in the *hcf208* mutant is indicated by an asterisk.

fluorescence induction measurements indicated that the introduced gene was functional and able to replace the mutated gene (Fig. 1A). We therefore designated this gene as *HCF208*.

HCF208 encodes a protein similar to CCB2 from Chlamydomonas reinhardtii

The *HCF208* gene encodes a protein of 275 amino acids with a calculated molecular mass of 30.6 kDa. The protein exhibits a chloroplast transit peptide of 39 amino acids predicted by the WolfPSORT program (http://wolfpsort.seq.cbrc.jp). Using the topology prediction server HMMTOP (Version 2.0; http://www.enzim.hu/hmmtop; Tusnady and Simon 1998, Tusnady and Simon 2001) two putative transmembrane domains were identified extending from amino acids 67 to 86 and 95 to 114, respectively. As described above, amino acid 68 is exchanged from glycine to arginine in the mutant *hcf208*. Thus, the mutation affects the first predicted transmembrane domain (Fig. 7B).

Analyses using the BLAST algorithm (Altschul et al. 1994) revealed homologs of the HCF208 protein in rice (*Oryza sativa*), green algae (*Chlamydomonas reinhardtii*) and cyanobacteria (*Anabaena variabilis*, *Nostoc punctiforme*, *Synechococcus* sp., etc.; Fig. 7B). The mutation in HCF208 affects an amino acid conserved among *Arabidopsis*, rice and *Chlamydomonas* but absent in cyanobacteria. The HCF208 homolog from *Chlamydomonas* was recently published as CCB2 (C = cofactor binding, C = cytochrome b_6f complex, B = subunit PetB) by

Kuras et al. (2007). Pairwise alignment of CCB2 and HCF208 revealed 26% identical amino acid residues and a similarity of 48%. CCB2 is one of four proteins (CCB1–CCB4) that are required for covalent binding of the c-type heme to cytochrome b_6 in *Chlamydomonas* (Kuras et al. 2007).

Discussion

The cytochrome b_6f complex consists of two monomers each containing six plastid-encoded and two nuclear-encoded polypeptides, as well as seven prosthetic groups. To ensure linear and cyclic electron transport, synthesis and maturation of the subunits as well as the assembly of the complex need to be governed by assisting factors, which are mainly encoded by nuclear genes.

The nuclear-encoded factor HCF208 is the Arabidopsis homolog of the c_n-heme-binding factor CCB2 from Chlamydomonas

In *Chlamydomonas*, Kuras et al. (2007) found that binding of heme c_n to cytochrome b_6 and thus the maturation of the holoprotein is dependent on the four nuclear-encoded factors CCB1–CCB4. We have identified the *Arabidopsis* homolog of CCB2 named HCF208. CCB2 is supposed to be a membrane protein with a long C-terminal domain in the stroma. Analysis of HCF208 with the help of a topology prediction server revealed the same orientation of the *Arabidopsis* protein in the thylakoid membrane. The phenotype of the *ccb2* mutant is apparently caused by absence of the CCB2 factor as the *ccb2-1* mutant allele has a stop codon in the fourth of six exons (Kuras et al. 2007). In *hcf208* mutant plants, the *HCF208* gene is transcribed, implying that the protein is synthesized. The mutant allele carries a point mutation in the second exon, which leads to an amino acid exchange from glycine to arginine at position 68 of the HCF208 protein. The exchange affects the second amino acid of the first predicted transmembrane domain. Since neutral amino acids are a prerequisite for membrane-spanning regions, we suggest that the protein cannot be inserted into the thylakoid membrane as a result of the positive charge in the transmembrane domain caused by the arginine. The generation of an antibody against the HCF208 protein will solve this point in future experiments. The exchanged glycine does not occur in the homologous protein of cyanobacteria. Thus, it is unlikely that this residue is associated with a catalytical function.

Analysis of the *hcf208* phenotype revealed drastically reduced amounts of the cytochrome b_6 polypeptide (~12.5%). More importantly, the protein shows an altered migration behavior compared with the wild-type protein and there is no cytochrome b_6 peroxidase activity detectable in the mutant. These results are in accordance with the phenotype of cytochrome b_6 Cys35 mutant strains from *Chlamydomonas*, which are unable to bind heme c_n (de Vitry et al. 2004, Kuras et al. 2007). Due to the absence of heme c_n, the molecular mass of the cytochrome b_6 polypeptide is reduced by 636 Da (de Vitry et al. 2004). Reduced amounts of the cytochrome b_6 polypeptide are supposed to be due to an increased degradation of the apoprotein as a result of the missing heme (Kuras et al. 1997). The Cys35 mutant phenotype can be observed in *ccb* mutants, demonstrating the role of the factors CCB1–CCB4 in attachment of heme c_n to the cytochrome b_6 polypeptide (Kuras et al. 2007). Equal phenotypes of the *Chlamydomonas* Cys35 and *ccb* mutants and the *Arabidopsis hcf208* mutant argue for an identical function of CCB2 and HCF208 in *Chlamydomonas* and *Arabidopsis*, respectively. Thus, HCF208 is the structural and functional homolog of CCB2 in higher plants.

The major subunits of the cytochrome b_6f complex are reduced in hcf208 mutants

We also examined the effects of the mutational defect in *hcf208* on the other three major subunits of the cytochrome b_6f complex. PetD is reduced as strongly as cytochrome b_6, but the *petD* transcripts are present in wild-type amounts. This result is consistent with the assumption that stabilization of PetD is dependent on the interaction with cytochrome b_6 (Kuras and Wollman 1994). Cytochrome *f* was also reduced to about 12.5% of wild-type levels, which correlates with the results of *petB* mutants of *Chlamydomonas* and tobacco in which loss of *petB* leads to a decrease of the cytochrome *f* polypeptide to 10% of wild-type levels or even less due to a lower rate of synthesis (Kuras and Wollman 1994, Monde et al. 2000, Choquet et al. 2003). We also found that the cytochrome *f* holoprotein is present in *hcf208* mutants, suggesting that heme binding to cytochrome *f* occurs in the mutant as previously described by Kuras et al. (2007). However, the majority of the cytochrome *f* protein seems to accumulate as apoprotein. This result might indicate that HCF208 is also involved in heme attachment to cytochrome *f*. However, the finding of Kuras et al. (2007) argues against such a role for the HCF208 homolog CCB2 from *Chlamydomona*s, as levels of cytochrome *f* detected by immunoblot analysis and peroxidase activity correlate with each other, meaning that the holoprotein is generated in *ccb* mutants. The discrepancy between *Arabidopsis* and *Chlamydomonas* possibly implies a closer relationship between complex assembly and heme binding in higher plant chloroplasts.

Moreover, P700 absorbance kinetics and immunoblot analyses revealed a reduction of PSI to 50% of wild-type amounts. We cannot exclude that this reduction is caused by photooxidative damage of PSI. On the other hand, the

decreased PSI content could be due to altered photosynthetic redox control possibly coupled with a modified retrograde signaling. Signaling pathways are triggered by intermediates of the tetrapyrrole biosynthetic pathway, plastid gene expression and the redox state of photosynthetic electron transport components (e.g. plastoquinone, cytochrome b_6f complex and reactive oxygen species; reviewed by Nott et al. 2006). The latter is known to control the expression of plastid-encoded (*psaA/B*) and nuclear-encoded (*PsaD*, *PsaF*) PSI components (reviewed by Pfannschmidt 2003).

The cytochrome b_6f complex can be assembled to its dimeric form in hcf208 mutant plants

The assembly of the cytochrome b_6f complex dimer is known to be dependent on the presence of mature subunits (Kuras et al. 1995). Our data show that there is no mature cytochrome b_6 polypeptide detectable in the mutant *hcf208*. By Blue Native PAGE analysis we could show that small amounts of the cytochrome b_6f complex dimer are present in the mutant, which has the same molecular mass as the wild type, demonstrating that all subunits are present in the complex.

We cannot rule out that there are trace amounts of mature cytochrome b_6 below our detection limit, which contribute to complex assembly. Alternatively, it is possible that cytochrome b_6 lacking heme c_n is incorporated into nascent cytochrome b_6f complexes. These complexes accumulate but are unable to participate in photosynthetic electron transport, as confirmed by our spectroscopic measurements. This is somewhat in contrast to other mutants defective in biogenesis of the cytochrome b_6f complex, e.g. *hcf164* or *hcf152-1* (Lennartz et al. 2001, Meierhoff et al. 2003). In spite of an equal reduction of the major subunits, there was a residual activity in *hcf164* and *hcf152-1* mutant plants, indicating that the remaining proteins are assembled into an active complex that is capable of linear electron transport.

Our data reveal that the assembly of the cytochrome b_6f complex might occur to some extent without attachment of heme c_n to cytochrome b_6.

The mutational analysis of the *Arabidopsis* cytochrome b_6f complex mutant *hcf208* confirmed the existence of the additional heme-binding pathway in higher plants, which was recently described for *Chlamydomonas* and called system IV (Kuras et al. 2007). Analysis of the CCB factors in *Chlamydomonas* and *Arabidopsis* should permit the functional characterization of this pathway in the near future. Moreover, the presence of a cytochrome b_6f complex dimer in the *hcf208* mutant, which presumably does not contain heme c_n, might help in studying the role of this heme in oxygenic photosynthesis.

Materials and Methods

Growth conditions and mutant selection

The mutant *hcf208* was selected from a collection of EMS-induced mutants. Surface-sterilized seeds were sown on Petri dishes containing 0.3% (w/v) gelrite (Roth, Karlsruhe, Germany) supplemented with 0.5 MS nutrients and 2% (w/v) sucrose. Seedlings were grown under a 16 h light/8 h darkness period at a photon flux density (PFD) of ~50–70 $\mu mol\, s^{-1}\, m^{-2}$. Mutant plants that exhibited high chlorophyll fluorescence were selected in the dark under UV light as described by Meurer et al. (1996). All plant material for the different experiments was grown in Petri dishes.

For seed production, plants were grown in a growth chamber operating at a 16 h light/8 h darkness period at a PFD of ~50–70 $\mu mol\, s^{-1}\, m^{-2}$ and a constant temperature of 21°C. For seed collection, the Arasystem (Beta Tech, Gent, Belgium) was used.

Spectroscopic measurements

Chlorophyll fluorescence induction and P700 absorbance measurements were performed as described by Meurer et al. (1996). Saturating light pulses had a length of 1 s and actinic light had a PFD of 110 $\mu mol\, s^{-1}\, m^{-2}$. For the P700 absorbance kinetics, the pulse amplitude-modified fluorometer was equipped with the dual-wavelength emitter detector unit ED-P700DW-E (Waltz).

SDS–PAGE of proteins, heme staining and immunoblotting

Isolation of crude leaf proteins from wild-type and mutant plants, SDS–PAGE (Schägger and von Jagow 1987) and immunodecoration of electroblotted proteins followed the methods described by Meurer et al. (1996). For the detection of heme peroxidase activity, proteins were blotted onto nitrocellulose membranes, incubated with the reagent of the chemiluminescence assay (ECL Plus; GE Healthcare, Uppsala, Sweden) and exposed to X-ray film for 4 h.

RNA isolation and gel blot analysis

Isolation of total leaf RNA was performed as described previously by Westhoff et al. (1991) and Meurer et al. (1996). The RNA was analyzed by Northern hybridization according to Westhoff et al. (1991) and Meurer et al. (1996) using specific probes listed by Westhoff and Herrmann (1988) and Meurer et al. (1996). Poly(A)$^+$ RNA was isolated using the Oligotex mRNA Midi Kit (Qiagen, Hilden, Germany) according to the protocol supplied by the manufacturer.

Two-dimensional Blue Native/SDS–PAGE

Blue Native PAGE was performed according to Schägger et al. (1994). Leaves of 3-week-old seedlings were homogenized in lysis buffer (10 mM HEPES-KOH, 10 mM $MgCl_2$, 25 mM KCl, pH 7.8) and centrifuged until a speed of 5,900 × *g* was reached. The chloroplast membranes were resuspended in ACA buffer (750 mM ε-aminocapronic acid, 50 mM Bis-Tris, 0.5 mM EDTA, pH 7) and solubilized with 1% (w/v) digitonin in ACA buffer for 30 min at 4°C and a chlorophyll concentration of 1 $mg\, ml^{-1}$. Unsolubilized material was removed by centrifugation (20 min at 4°C and 15,000 × *g*). The supernatant was supplemented with 0.5% (w/v) *n*-dodecyl-β-D-maltoside and one-tenth volume of sample buffer [50 mM Bis-Tris, 750 mM ε-aminocapronic acid, 30% (w/v) sucrose, 5% (w/v) Coomassie Serva G] and analyzed on a 6–12% gradient gel by Blue Native PAGE. A 20 µg aliquot of

chlorophyll was loaded per lane. Separated protein complexes were analyzed by SDS–PAGE in the second dimension. Immunoblot analysis was performed as described above.

Mapping of hcf208

An F_2 mapping population was produced by crossing wild-type plants of the ecotype Landsberg erecta with Columbia plants heterozygous for the *hcf208* mutation and selfing of the resulting F_1 plants. DNA from homozygous mutant F_2 plants was screened with various microsatellite markers to localize the gene on chromosome 5. For fine-mapping, the mapping population was examined with markers listed in the Arabidopsis Information Resource database and markers designed with the information of the Cereon Arabidopsis Polymorphism Collection (www.arabidopsis.org.cereon; Jander et al. 2002). The region of the mutation locus was narrowed down between the markers MSG15-3/4 (MSG15-3/4-H, 5′-CTCCTCTTCGAGATGCACATG-3′; MSG15-3/4-R, 5′-CTCTCCCTTCCACGGCAG-3′; polymorphism, *Tai*I restriction site in *Col*) and F17P19 (F17P19-H, 5′-CCAACAGACCGACCGGAG-3′; F17P19-R, 5′-CACCAC ACACTTGATACTTATC-3′; polymorphism, *Rsa*I restriction site in *Col*).

Complementation of the mutant phenotype

For complementation of the mutant phenotype, a construct was produced using Gateway technology (Hartley et al. 2000). *HCF208* cDNA was obtained by reverse transcription of total RNA isolated from the wild type. The cDNA was amplified using primers start208attB1-F (5′-GGGGACAAGTTTGTACAAAAA AGCAGGCT*ATG*AGTATTCAAATTTGTAATTTC-3′; the attB1 site is underlined, the start codon is shown in italics) and 208stop-attB2-R (5′-GGGGACCACTTTGTACAAGAAAGCTG GGT*CTA*ACCTCTGAATTTCTCAGC-3′; the attB2 site is underlined, the stop codon is shown in italics). A BP clonase reaction (Invitrogen, Karlsruhe, Germany) between the PCR product and donor vector pDONR221 was accomplished according to the Gateway manual, creating pENTRY221 + *HCF208*. After sequence analysis of the recombined DNA sequence, an LR clonase reaction (Invitrogen, Karlsruhe, Germany) was performed to introduce the *HCF208* gene into the binary Ti destination vector pGWB2 (35S promoter, no tag; Nakagawa et al. 2007), generating pGWB2 + *HCF208*. Subsequently, the construct was transferred to heterozygous *hcf208* plants by *Agrobacterium tumefaciens*-mediated transformation according to the floral dip method (Clough and Bent 1998). Complemented transgenic plants were named *HCF208*c.

Accession numbers

The accession number of the protein that is encoded by *HCF208* is NP_200024. Accession numbers of proteins described in the text are as follows: ABB47906 (OsHCF208), ABP57442 (CrCCB2) and YP_322396 (AvHCF208).

Acknowledgment

Funding was provided by the German Science Foundation (SFB-TR1 to P.W.)

References

Allen, J. (2002) Photosynthesis of ATP—electrons, proton pumps, rotors, and poise. *Cell* 110: 273–276.

Altschul, S.F., Boguski, M.S., Gish, W. and Wootton, J.C. (1994) Issues in searching molecular sequence databases. *Nat. Genet.* 6: 119–129.

Barkan, A. (1988) Proteins encoded by a complex chloroplast transcription unit are each translated from both monocistronic and polycistronic mRNAs. *EMBO J.* 7: 2637–2644.

Baymann, F., Giusti, F., Picot, D. and Nitschke, W. (2007) The ci/bH moiety in the b6f complex studied by EPR: a pair of strongly interacting hemes. *Proc. Natl Acad. Sci. USA* 104: 519–524.

Breyton, C., Tribet, C., Olive, J., Dubacq, J.P. and Popot, J.L. (1997) Dimer to monomer conversion of the cytochrome b6 f complex. Causes and consequences. *J. Biol. Chem.* 272: 21892–21900.

Choquet, Y., Zito, F., Wostrikoff, K. and Wollman, F.A. (2003) Cytochrome f translation in Chlamydomonas chloroplast is autoregulated by its carboxyl-terminal domain. *Plant Cell* 15: 1443–1454.

Clough, S.J. and Bent, A.F. (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* 16: 735–743.

Cramer, W.A. and Zhang, H. (2006) Consequences of the structure of the cytochrome b6f complex for its charge transfer pathways. *Biochim. Biophys. Acta* 1757: 339–345.

de Vitry, C., Desbois, A., Redeker, V., Zito, F. and Wollman, F.A. (2004) Biochemical and spectroscopic characterization of the covalent binding of heme to cytochrome b6. *Biochemistry* 43: 3956–3968.

Gray, J.C. (1992) Cytochrome f: structure, function and biosynthesis. *Photosynth. Res.* 34: 359–374.

Hager, M., Biehler, K., Illerhaus, J., Ruf, S. and Bock, R. (1999) Targeted inactivation of the smallest plastid genome-encoded open reading frame reveals a novel and essential subunit of the cytochrome b6f complex. *EMBO J.* 18: 5834–5842.

Hartley, J.L., Temple, G.F. and Brasch, M.A. (2000) DNA cloning using in vitro site-specific recombination. *Genome Res.* 10: 1788–1795.

Huang, D., Everly, R.M., Cheng, R.H., Heymann, J.B., Schagger, H., Sled, V., Ohnishi, T., Baker, T.S. and Cramer, W.A. (1994) Characterization of the chloroplast cytochrome b6f complex as a structural and functional dimer. *Biochemistry* 33: 4401–4409.

Jander, G., Norris, S.R., Rounsley, S.D., Bush, D.F., Levin, I.M. and Last, R.L. (2002) Arabidopsis map-based cloning in the post-genome era. *Plant Physiol.* 129: 440–450.

Joliot, P. and Joliot, A. (2006) Cyclic electron flow in C3 plants. *Biochim. Biophys. Acta* 1757: 362–368.

Junge, W. (1999) ATP synthase and other motor proteins. *Proc. Natl Acad. Sci. USA* 96: 4735–4737.

Kramer, D.M., Johnson, G., Kiirats, O. and Edwards, G.E. (2004) New fluorescence parameters for the determination of QA redox state and excitation energy fluxes. *Photosynth. Res.* 79: 209–218.

Kuras, R., Buschlen, S. and Wollman, F.A. (1995) Maturation of pre-apocytochrome f in vivo. A site-directed mutagenesis study in. *Chlamydomonas reinhardtii*. *J. Biol. Chem.* 270: 27797–27803.

Kuras, R., Saint-Marcoux, D., Wollman, F.A. and de Vitry, C. (2007) A specific c-type cytochrome maturation system is required for oxygenic photosynthesis. *Proc. Natl Acad. Sci. USA* 104: 9906–9910.

Kuras, R. and Wollman, F.A. (1994) The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga. *Chlamydomonas reinhardtii*. *EMBO J.* 13: 1019–1027.

Kuras, R., de Vitry, C., Choquet, Y., Girard-Bascou, J., Culler, D., Buschlen, S., Merchant, S. and Wollman, F.A. (1997) Molecular genetic identification of a pathway for heme binding to cytochrome b6. *J. Biol. Chem.* 272: 32427–32435.

Kurisu, G., Zhang, H., Smith, J.L. and Cramer, W.A. (2003) Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science* 302: 1009–1014.

Lennartz, K., Plucken, H., Seidler, A., Westhoff, P., Bechtold, N. and Meierhoff, K. (2001) HCF164 encodes a thioredoxin-like protein involved in the biogenesis of the cytochrome b(6)f complex in Arabidopsis. *Plant Cell* 13: 2539–2551.

Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N. and Schuster, G. (2003) HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. *Plant Cell* 15: 1480–1495.

Meurer, J., Meierhoff, K. and Westhoff, P. (1996) Isolation of high-chlorophyll-fluorescence mutants of Arabidopsis thaliana and their characterisation by spectroscopy, immunoblotting and northern hybridisation. *Planta* 198: 385–396.

Monde, R.A., Zito, F., Olive, J., Wollman, F.A. and Stern, D.B. (2000) Post-transcriptional defects in tobacco chloroplast mutants lacking the cytochrome b6/f complex. *Plant J.* 21: 61–72.

Mosser, G., Breyton, C., Olofsson, A., Popot, J.L. and Rigaud, J.L. (1997) Projection map of cytochrome b6 f complex at 8 Å resolution. *J. Biol. Chem.* 272: 20263–20268.

Nakagawa, T., Kurose, T., Hino, T., Tanaka, K., Kawamukai, M., Niwa, Y., Toyooka, K., Matsuoka, K., Jinbo, T. and Kimura, T. (2007) Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation. *J. Biosci. Bioeng.* 104: 34–41.

Nott, A., Jung, H.S., Koussevitzky, S. and Chory, J. (2006) Plastid-to-nucleus retrograde signaling. *Annu. Rev. Plant Biol.* 56: 739–759.

Pfannschmidt, T. (2003) Chloroplast redox signals: how photosynthesis controls its own genes. *Trends Plant Sci.* 8: 33–41.

Pierre, Y., Chabaud, E., Herve, P., Zito, F. and Popot, J.L. (2003) Site-directed photochemical coupling of cytochrome b6f-associated chlorophyll. *Biochemistry* 42: 1031–1041.

Robertson, D.E., Farid, R.S., Moser, C.C., Urbauer, J.L., Mulholland, S.E., Pidikiti, R., Lear, J.D., Wand, A.J., DeGrado, W.F. and Dutton, P.L. (1994) Design and synthesis of multi-haem proteins. *Nature* 368: 425–432.

Schagger, H., Cramer, W.A. and von Jagow, G. (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* 217: 220–230.

Schagger, H. and von Jagow, G. (1987) Tricine–sodium dodecyl sulfate–polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166: 368–379.

Schneider, D., Berry, S., Rich, P., Seidler, A. and Rogner, M. (2001) A regulatory role of the PetM subunit in a cyanobacterial cytochrome b6f complex. *J. Biol. Chem.* 20: 16780–16788.

Schottler, M.A., Flugel, C., Thiele, W. and Bock, R. (2007) Knock-out of the plastid-encoded PetL subunit results in reduced stability and accelerated leaf age-dependent loss of the cytochrome b6f complex. *J. Biol. Chem.* 282: 976–985.

Schwenkert, S., Legen, J., Takami, T., Shikanai, T., Herrmann, R.G. and Meurer, J. (2007) Role of the low-molecular-weight subunits PetL, PetG, and PetN in assembly, stability, and dimerisation of the cytochrome b6f complex in tobacco. *Plant Physiol.* 144: 1924–1935.

Stroebel, D., Choquet, Y., Popot, J.L. and Picot, D. (2003) An atypical haem in the cytochrome b(6)f complex. *Nature* 426: 413–418.

Tusnady, G.E. and Simon, I. (1998) Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J. Mol. Biol.* 283: 489–506.

Tusnady, G.E. and Simon, I. (2001) The HMMTOP transmembrane topology prediction server. *Bioinformatics* 17: 849–850.

Westhoff, P. and Herrmann, R.G. (1988) Complex RNA maturation in chloroplasts. The psbB operon from spinach. *Eur. J. Biochem.* 171: 551–564.

Westhoff, P., Offermann-Steinhard, K., Höfer, M, Eskins, K., Oswald, A. and Streubel, M. (1991) Differential accumulation of plastid transcripts encoding photosystem II in the mesophyll and bundle-sheath cells of monocotyledonous NADP-malic enzyme-type C4 plants. *Planta* 184: 377–388.

Whitelegge, J.P., Zhang, H., Aguilera, R., Taylor, R.M. and Cramer, W.A. (2002) Full subunit coverage liquid chromatography electrospray ionization mass spectrometry (LCMS+) of an oligomeric membrane protein: cytochrome b(6)f complex from spinach and the cyanobacterium *Mastigocladus laminosus*. *Mol. Cell Proteomics* 1: 816–827.

Yamashita, E., Zhang, H. and Cramer, W.A. (2007) Structure of the cytochrome b6f complex: quinone analogue inhibitors as ligands of heme cn. *J. Mol. Biol.* 370: 39–52.

Yamazaki, H., Tasaka, M. and Shikanai, T. (2004) PPR motifs of the nucleus-encoded factor, PGR3, function in the selective and distinct steps of chloroplast gene expression in Arabidopsis. *Plant J.* 38: 152–163.

Zatsman, A.I., Zhang, H., Gunderson, W.A., Cramer, W.A. and Hendrich, M.P. (2006) Heme–heme interactions in the cytochrome b6f complex: EPR spectroscopy and correlation with structure. *J. Am. Chem. Soc.* 128: 14246–14247.

Zhang, H., Huang, D. and Cramer, W.A. (1999) Stoichiometrically bound beta-carotene in the cytochrome b6f complex of oxygenic photosynthesis protects against oxygen damage. *J. Biol. Chem.* 274: 1581–1587.

Zhang, H., Whitelegge, J.P. and Cramer, W.A. (2001) Ferredoxin : NADP+ oxidoreductase is a subunit of the chloroplast cytochrome b6f complex. *J. Biol. Chem.* 276: 38159–38165.

(Received August 10, 2007; Accepted October 23, 2007)

Authors´ Contributions

DL performed all experiments except fine mapping of the mutant locus (carried out by SP). DL wrote the paper. KM and PW participated in drafting of the manuscript.

Manuscript 3

Molecular characterization of HCF208 localization and interactions

Introduction

The cytochrome b_6f complex of organisms performing oxygenic photosynthesis mediates linear electron transfer between PSII and PSI by acting as a plastoquinol-plastocyanin oxidoreductase. The electron transport is coupled to translocation of protons through the thylakoid membrane and therefore generation of an electrochemical gradient that is used for ATP synthesis (Junge, 1999; Allen 2002). The cytochrome b_6f complex is also involved in ferredoxin-dependent cyclic electron transport around PSI and in photoprotection (Joliot and Joliot, 2006; Shikanai, 2007; DalCorso et al., 2008).

The active form of the cytochrome b_6f complex is a dimer (Huang et al., 1994; Breyton et al., 1997; Mosser et al., 1997). Each monomer consists of eight subunits: the four large polypeptides cytochrome *f* (PetA), cytochrome b_6 (PetB), the Rieske protein (PetC) and subunit IV (PetD), and the small polypeptides PetG, PetL, PetM and PetN (Kurisu et al., 2003; Stroebel et al., 2003). Additionally, the complex is associated with overall seven co-factors per monomer, of which five are redox-active and contribute to electron transport (Kurisu et al., 2003; Stroebel et al., 2003; Cramer et al., 2006). Those co-factors bind to three of the large polypeptides: the Rieske protein harbors a [2Fe-2S] cluster, cytochrome *f* binds one c-type heme and cytochrome b_6 is associated with two b-type and one c-type hemes (Kurisu et al., 2003; Stroebel et al., 2003). The two b-type hemes, heme b_n at the electronegative (stromal) and heme b_p at the electropositive (lumenal) side of the complex, are non-covalently attached to cytochrome b_6 via pairs of histidines. Binding of these hemes is supposed to occur spontaneously (Robertson et al., 1994). In contrast, the c-type hemes of cytochrome b_6 (heme c_n) and cytochrome *f* are covalently attached to their respective proteins.

C-type hemes typically form two thioether bonds to the cysteine residues of a highly conserved CXXCH heme binding-motif and are liganded by the motifs histidinyl residue (Giegé et al., 2008). Unlike cytochrome *f*, which possesses the typical c-type heme binding motif, cytochrome b_6 binds heme c_n in an atypical manner via a single cysteine residue (Cys35) and provides no axial amino acid ligands (Kurisu et al., 2003; Stroebel et al., 2003). Moreover, heme c_n is liganded by a water or hydroxide ion in a heme b_n propionate group and by plastoquinone (Kurisu et al., 2003; Stroebel et al., 2003; Yamashita et al., 2007).

Three distinct pathways referred to as systems I to III operating in bacteria, mitochondria and chloroplasts have been described for biogenesis of c-type cytochromes (Kranz et al., 2009). Those systems commonly act on the electropositive side of membranes and require components that mediate heme delivery, heme and cysteine ligand reduction and ligation of heme to the apocytochrome (Kranz et al., 2009). In chloroplasts system II facilitates heme attachment to cytochrome *f* at the electropositive (lumenal) side of the thylakoid membrane

(Kranz et al., 2009), whereas cytochrome b_6 binds its atypical heme c_n on the electronegative (stromal) side of the complex (stroma; Kurisu et al., 2003; Stroebel et al., 2003) independently from system II.

Previously, the nuclear *high* chlorophyll *f*luorescence mutant *hcf208* from *Arabidopsis thaliana* was described (Lyska et al., 2007). In this mutant, the cytochrome b_6 protein exhibited an altered migration behavior when compared to the wild-type situation and did not possess peroxidase activity, which pointed to the lack of the covalent heme c_n (Lyska et al., 2007). It was demonstrated that the HCF208 locus encodes for the *Arabidopsis* homolog of *Chlamydomonas reinhardtii CCB2.* CCB2 is one of four nuclear-encoded factors (CCB1-CCB4) that are supposed to act together in maturation of cytochrome b_6 since respective mutants exhibited the same heme c_n-deficient phenotype (Kuras et al., 2007). CCB1-CCB4 were therefore referred to as “system IV”, a novel c-type cytochrome maturation pathway specific for cytochrome b_6 (Kuras et al., 2007). The four CCB proteins are conserved among all organisms performing oxygenic photosynthesis (Kuras et al., 2007).

To shed light on HCF208 function (hereafter also referred to as AtCCB2) an therefore on the operating mode of system IV in *Arabidopsis*, the CCB homologs AtCCB1, AtCCB3 and AtCCB4 were included into the present studies. Different mutant alleles of AtCCB1, HCF208/ AtCCB2 and AtCCB4 (Lyska et al., 2007; Lezhneva et al., 2008) were analyzed in terms of mutant phenotype, photosynthetic capacity (chlorophyll fluorescence induction) and accumulation of cytochrome b_6. Interaction partners of HCF208/ AtCCB2 were identified by different approaches showing that HCF208/ AtCCB2 plays a central role in the novel cytochrome b_6 maturation pathway.

Results

Comparison of *Arabidopsis ccb* mutants

The mutant *hcf208* was selected from a collection of ethyl methansulfonate (EMS)-induced mutants and carries a point mutation in exon 2 of the HCF208/ AtCCB2 gene (Figure 1; Lyska et al., 2007). The *atccb1*_SALK, *atccb2*_SALK and *atccb4*_SALK mutants harbor T-DNA insertions, which lead to disruption of respective genes and loss of protein function (Figure 1; Lezhneva et al., 2008). Further, another mutant allele of AtCCB1 (*atccb1*_GK) was isolated from the GABI-Kat collection harboring a T-DNA insertion in the 5´ UTR of AtCCB1 (Figure 1) and also exhibiting an *hcf* mutant phenotype. Mutants affected in AtCCB3 were not available. Homozygous mutant plants were unable to grow photoautotrophically on soil and were therefore grown on sucrose-supplemented media. Plants from all mutant lines except *atccb1*_GK were smaller and slightly paler when compared to the wild type (Figure 2 A). A previously described pale green phenotype of the SALK mutants (Lezhneva et al., 2008) could not be confirmed in this study since plants were grown under long day conditions (16h light/ 8h darkness) rather than continuous light as in previous studies by Leznheva et al. (2008).

The photosynthetic capacity of the mutants and the wild type was determined by chlorophyll fluorescence induction measurements (Figure 2 B). PSII activity is reflected by the ratio of variable to maximum chlorophyll fluorescence (F_v/F_m; where $F_v = F_m – F_0$ is the variable fluorescence, F_m is the maximum fluorescence, and F_0 is the minimum fluorescence) indicating an estimate of the maximum portion of absorbed quanta used in PSII reaction centers. The F_v/F_m values of all mutants were slightly reduced when compared to wild type, but still pointed to the overall integrity of PSII (Figure 2 B). The fraction of still open PSII reaction centers is reflected by the qL value (photochemical quenching according to the “lake model”; $qL = [(F_m´-F_s)/(F_m´- F_0´)] \times (F_0´/F_s)$; where F_s is steady state fluorescence; Kramer et al., 2004) indicating the capacity of electron transport downstream of PSII. The qL values of all mutant alleles were drastically reduced when compared to the wild type (Figure 2 B). *Atccb1*_SALK, *hcf208*, *atccb2*_SALK and *atccb4*_SALK exhibited the most dramatic reductions of qL values (5%, 6%, 9%, and 12% of wild type, respectively). Fluorescence hardly decreased after switching-on actinic light pointing to an almost complete loss of electron transfer downstream of PSII (Figure 2 B). In contrast, the qL value (15% of wild type) and the slight reduction of fluorescence after switching-on of actinic light indicated a somewhat less severe defect in electron transport downstream of PSII in the *atccb1*_GK mutant.

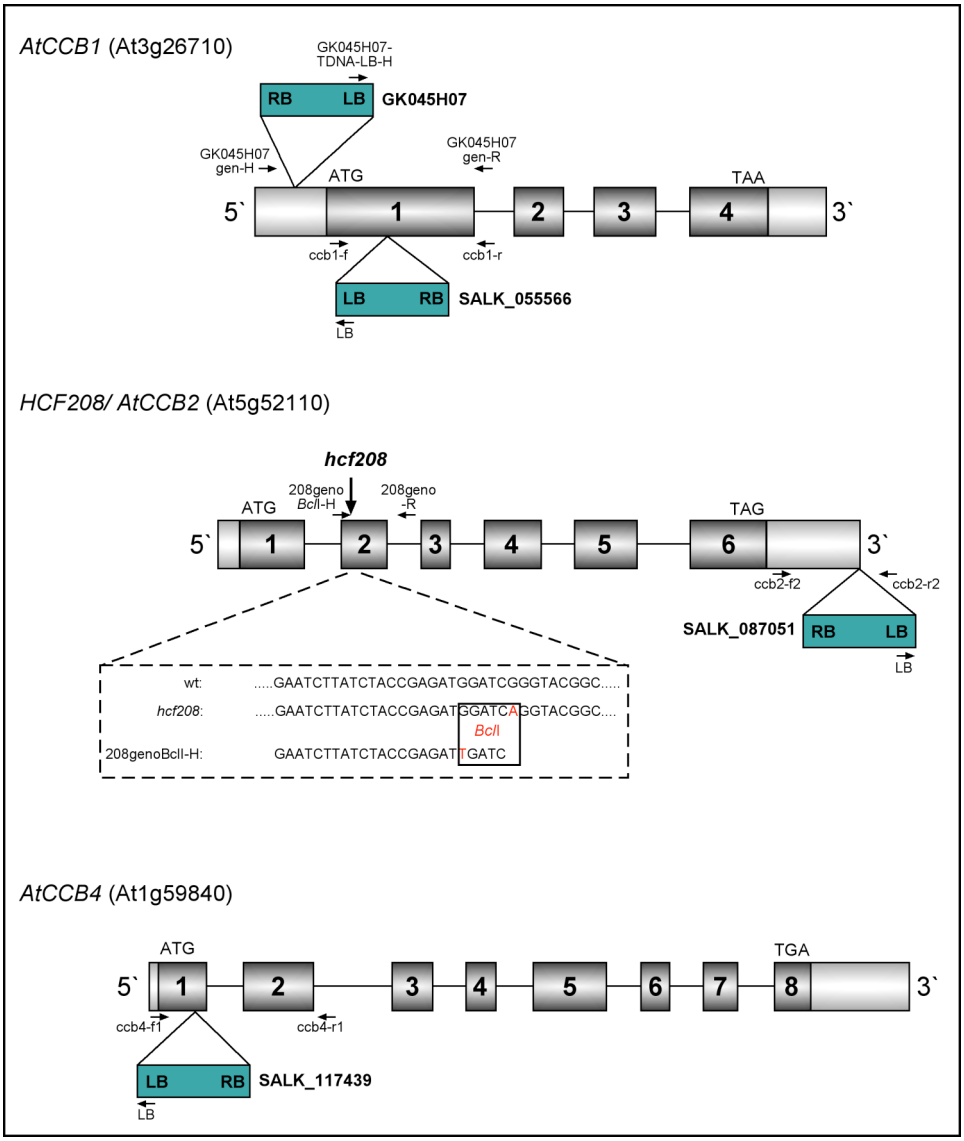
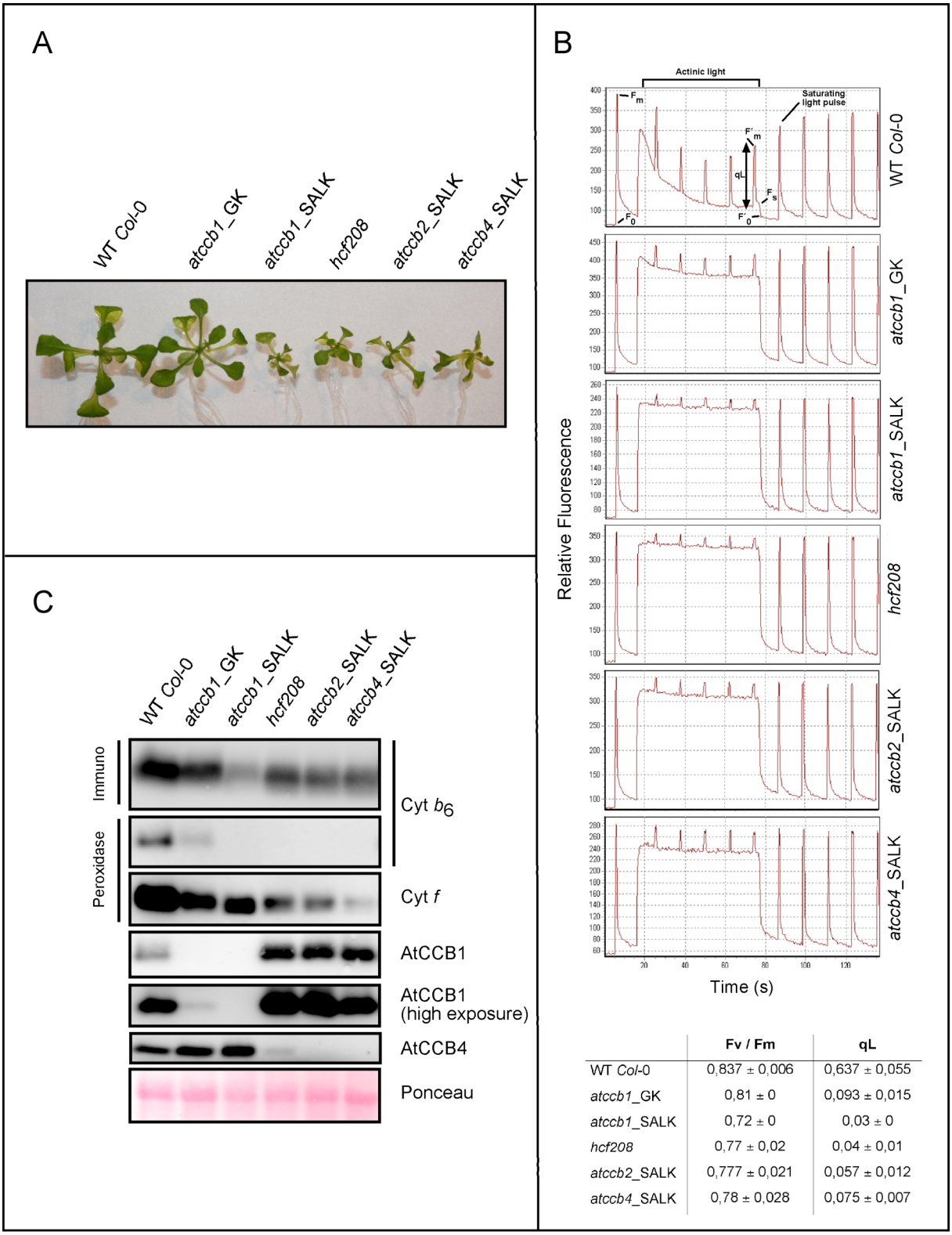


Figure1: Schematic representations of the *AtCCB1*, *HCF208/ AtCCB2* and *AtCCB4* genes.
The sites of T-DNA insertions and point mutation are shown for each gene. The sites of primers used for PCR analyses are indicated by small arrows.

In previous studies on mutants affected in the CCB genes, cytochrome b_6 was found to be drastically reduced, migrate faster in SDS-PAGE than wild type and to not possess peroxidase activity due to the lack of heme c_n (Kuras et al., 2007; Lyska et al., 2007; Lezhneva et al., 2008). Direct comparison of all abovementioned mutant alleles confirmed reduction of cytochrome b_6 levels and lack of peroxidase activity of the remaining protein (Figure 2 C). The only exception was the *atccb1*_GK mutant. It exhibited traces of cytochrome b_6 peroxidase activity and accumulated relatively high cytochrome b_6 amounts when compared to the other mutants (Figure 2 C) pointing to residual AtCCB1 activity in *atccb1*_GK. In contrast, the *atccb1*_SALK mutant exhibited the greatest decrease of cytochrome b_6 levels (Figure 2 C). In *hcf208*, *atccb2*_SALK and *atccb4*_SALK, cytochrome b_6 was reduced to about equal amounts (Figure 2 C). Cytochrome *f* peroxidase activity was



	Fv / Fm	qL
WT *Col*-0	0,837 ± 0,006	0,637 ± 0,055
*atccb1*_GK	0,81 ± 0	0,093 ± 0,015
*atccb1*_SALK	0,72 ± 0	0,03 ± 0
hcf208	0,77 ± 0,02	0,04 ± 0,01
*atccb2*_SALK	0,777 ± 0,021	0,057 ± 0,012
*atccb4*_SALK	0,78 ± 0,028	0,075 ± 0,007

Figure 2: Characterization of wild type (WT) Col-0, *atccb1*-, *hcf208*/*atccb2*- and *atccb4* mutant alleles.
(A) Phenotypes, (B) chlorophyll fluorescence induction charts and table of crucial values for photosynthetic capacity, and (C) Western blot analysis of 3-week-old wild-type and mutant plants grown under long day conditions (16h light/ 8h darkness). For Western blot 30μg of membrane proteins were analyzed using antibodies against AtCCB1, AtCCB4 and cytochrome b_6 or via detection of peroxidase activities of cytochrome b_6 and cytochrome *f* by enhanced chemoluminescence.

also reduced, but still present in all mutants, which is in accordance to previous results (Lyska et al., 2007; Lezhneva et al., 2008).

Taken together, characterization of mutants affected in the AtCCB1, HCF208/ AtCCB2 and AtCCB4 genes revealed overall similar defects in terms of growth phenotype, electron transport capacity and protein accumulation. This data confirms involvement of these proteins in heme c_n attachment to cytochrome b_6 and points to an equal contribution of at least three out of the four known CCB-factors in this process.

Interdependent accumulation of AtCCB proteins

Using polyclonal antibodies raised against AtCCB1 and AtCCB4 (Supplemental data 1) accumulation of the respective proteins was tested in the five mutants affected in AtCCB1, HCF208/ AtCCB2 and AtCCB4. Production of antibodies against HCF208/ AtCCB2 and AtCCB3 did not succeed. In the *atccb1*_SALK mutant the AtCCB1 protein was absent, whereas in *atccb1*_GK a slight band was visible upon high exposure (Figure 2 C). Therefore, the weak *atccb1*_GK mutant phenotype described above is due to accumulation of residual AtCCB1 amounts and function and a loss-of-function mutant phenotype is represented only by *atccb1*_SALK. In *hcf208*, *atccb2*_SALK and *atccb4*_SALK mutants the AtCCB1 levels were elevated when compared to wild type (Figure 2C).

Similarly, the AtCCB4 protein accumulated to slightly higher amounts in the two *atccb1* mutant alleles when compared to the wild type. AtCCB4 was not detected in *atccb4*_SALK mutants, wherefore the mutant phenotype displays complete loss of AtCCB4 function. Interestingly, in *atccb2*_SALK mutants AtCCB4 was also absent and accumulated to only small amounts in *hcf208* (Figure 2 C) indicating a dependency of AtCCB4 accumulation on the presence of HCF208/ AtCCB2. Unfortunately, the reciprocal situation could not be determined due to the lack of a HCF208/ AtCCB2-specific antibody.

HCF208/ AtCCB2 associated with chloroplast membranes

The HCF208/ AtCCB2 protein was predicted to harbor an N-terminal chloroplast transit peptide and two transmembrane domains (Lyska et al., 2007; Lezhneva et al., 2008). Targeting of the protein to the chloroplast was confirmed previously (Lezhneva et al., 2008), wherefore the association with chloroplast membranes was addressed in the present study. Since no sufficient antibody against the HCF208/ AtCCB2 protein could be generated (Supplemental data 1), transgenic lines expressing a C-terminally hemagglutinin (HA)/Strep-tag*III* tagged version of the protein in *hcf208* mutant background under the control of the 35S promoter were used to allow its detection (HCF208pAUL2; Lyska et al., 2011). Chloroplasts

from HCF208pAUL2 lines were isolated and fractionated into stroma and membranes. Membranes were treated with different salts in order to determine the strength of membrane-association. The HCF208/AtCCB2 fusion protein was only detected in chloroplast membranes but not in the stroma (Figure 3). The stromal protein RbcL served as a control for this fraction. Washing of membranes with sodium chloride and sodium carbonate, which remove peripheral membrane proteins (Fujiki et al., 1982), did not release HCF208/ AtCCB2 from membranes (Figure 3) indicating that HCF208/ AtCCB2 is a chloroplast-localized integral membrane protein as predicted.

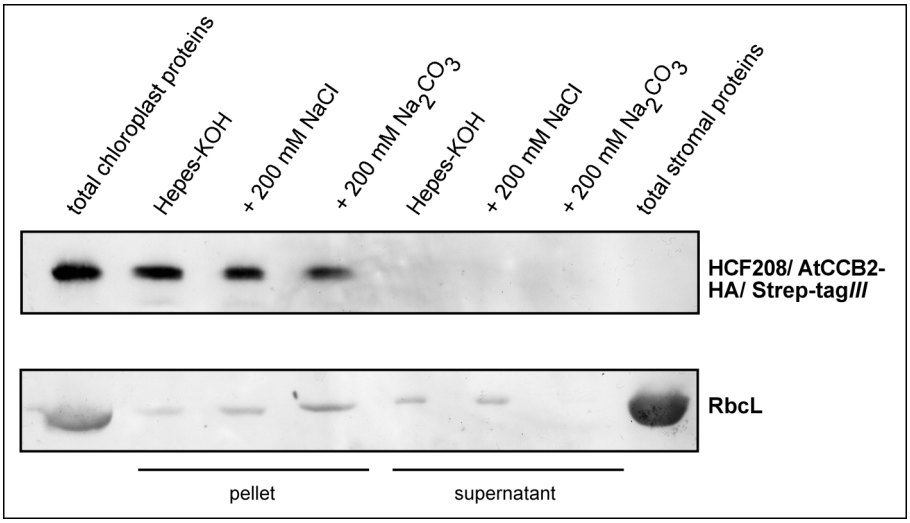


Figure 3: Immunolocalization of HCF208/ AtCCB-HA/Strep-tag*///*.
Chloroplasts were isolated from four week-old HCF208pAUL2 plants by Percoll gradient centrifugation, lysed, and separated stroma and membrane proteins. Membranes were washed with 200 mM NaCl and 250 mM Na_2CO_3 respectively. Equivalents of 20 μg chlorophyll were used for SDS-PAGE and Western blot analyses.

HCF208/ AtCCB2 forms distinct complexes

Since HCF208/ AtCCB2 acts functionally interacts with the three factors AtCCB1, AtCCB3, and AtCCB4 its association with other proteins was tested by two-dimensional Blue Native (BN)-PAGE (Schägger et al., 1994). Solubilized membrane proteins from HCF208pAUL2 were analyzed by nondenaturating BN-PAGE in the first dimension followed by separation on denaturating SDS gels in the second dimension. Figure 4 presents immunoblot analysis of two 2D-BN-PAGE experiments. In the first dimension the majority of the HCF208/ AtCCB2 fusion protein was detected at low molecular weight with two strong signals below and slightly above the 67 kDa band of the molecular weight standard. Since the HCF208/ AtCCB2-HA/ Strep-tag*///* fusion protein migrates as a ~35 kDa protein in SDS-PAGE (see Lyska et al., 2011), these two spots may display a HCF208/ AtCCB2 monomer and a dimeric form (homo- or heterodimer). The large amounts of the monomer presumably originate from the overexpression of the fusion protein by the 35S promoter. Other characteristic spots were

detected at ~170 kDa and in the range of ~300 kDa (Figure 4). Therefore, the integral membrane protein HCF208/ AtCCB2 is able to form homo- or heterodimers and larger complexes suggesting interaction with other proteins.

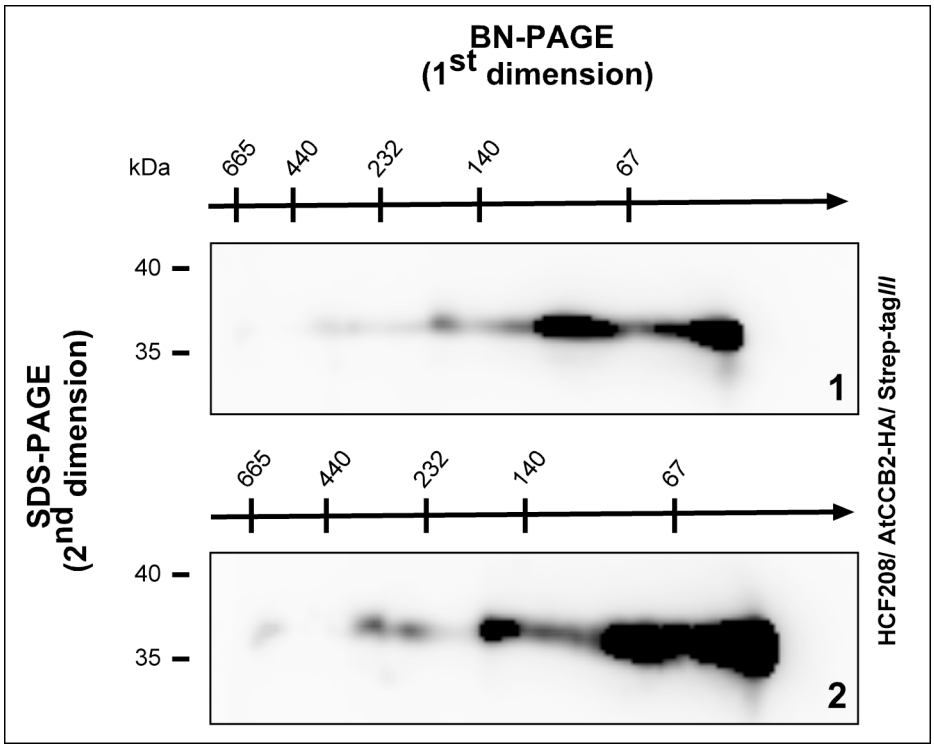


Figure 4: Identification of HCF208/ AtCCB2-HA/Strep-tagIII-containing complexes by Two-Dimensional Blue Native/SDS-PAGE.
Membranes from two week-old HCF208pAUL2 plants were solubilized and separated by Blue Native-PAGE according to their molecular weights (1st dimension) and subsequently by denaturating SDS-PAGE using (2nd dimension). Two independent experiments are shown.

AtCCB4 co-precipitates with HCF208/ AtCCB2

To find interaction partners of HCF208/ AtCCB2 *in planta* co-precipitation analyses were performed using HCF208pAUL6 lines expressing a C-terminally HA/Strep-tag*III*/ProteinA-tagged HCF208/ AtCCB2 protein (Lyska et al., 2011). In this line, expression is controlled by the *HCF107* promoter to prevent overaccumulation of the fusion protein (Lyska et al., 2011) and therefore to increase the probability to co-purify interaction partners. Large-scale affinity purifications (3 experiments) via IgG Sepharose were performed using solubilized membranes. Coomassie staining of eluates separated by SDS-PAGE revealed the presence of two specific bands in HCF208/ AtCCB2 eluates that were not present in the controls (wild type Col-0 and HCF107pAUL3; Lyska et al., 2011) at ~ 40 kDa and in the range of ~23 kDa (asterisks). To identify all proteins in the three eluates mass spectrometry (MS) was performed. Proteins identified in the controls were subtracted from those identified in HCF208pAUL6 eluates (Supplemental table 1) to restrict data to HCF208/ AtCCB2-specific

potential interaction partners. HCF208/ AtCCB2 was detected in all three eluates via 2 or 1 peptides, respectively. Also, peptides corresponding to AtCCB4 were identified in the HCF208pAUL6 eluates from all three independent experiments (6, 10 and 3 peptides,

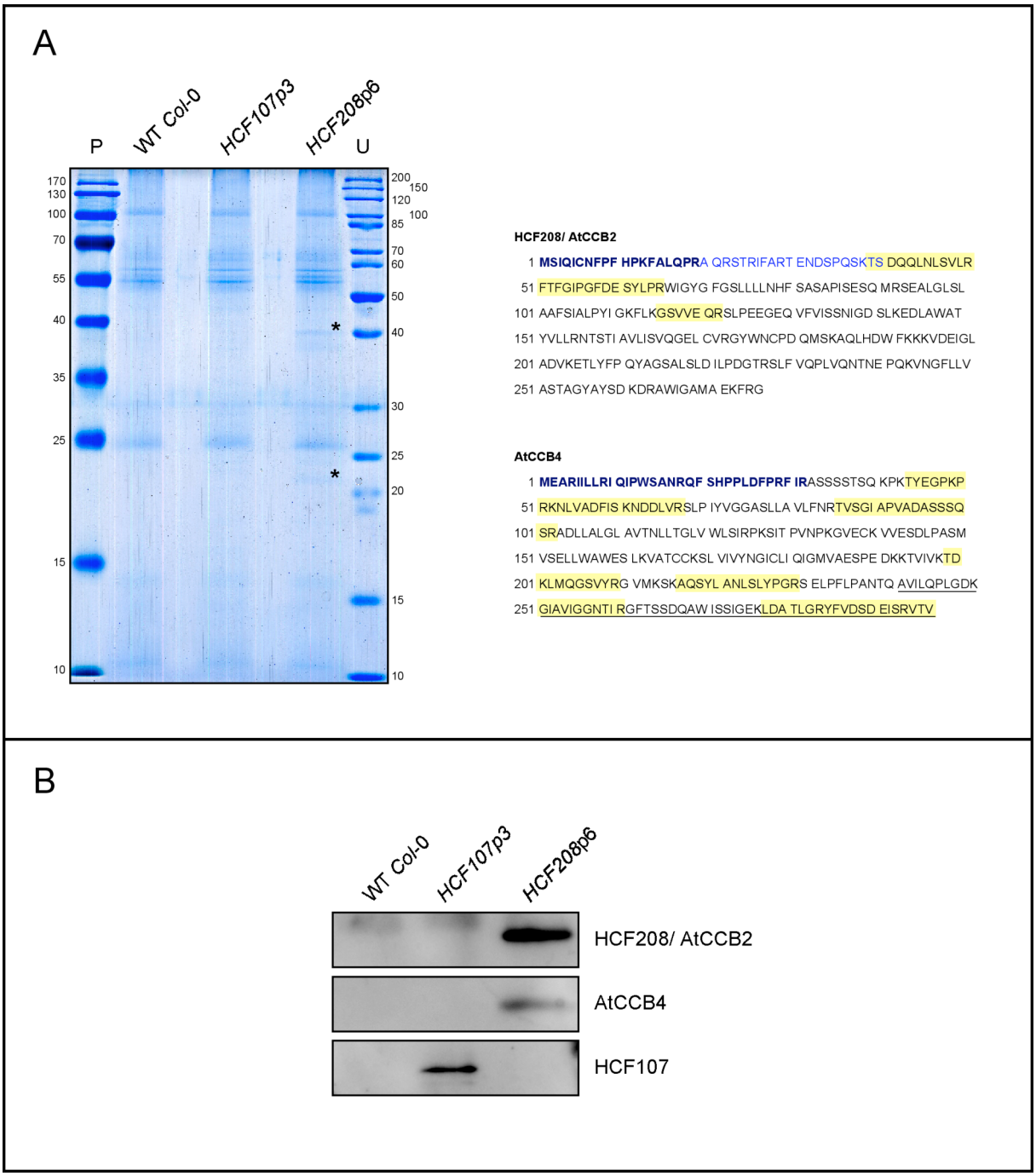


Figure 5: Affinity purification of HCF208/ AtCCB2-HA/Strep-tag*III*.
Solubilized membrane proteins from four-week-old wild type (WT) Col-0, HCF107pAUL3 and HCF208pAUL6 plants were purified via IgG Sepharose. (B) Coomassie staining of eluates according to 10 mg chlorophyll (one out of three experiments). Bands specific to the HCF208pAUL6 eluate are marked by asterisks. Lanes were diced into small pieces, digested and analyzed by LC-ESI-MS/MS. Peptides identified from HCF208/ AtCCB2 and AtCCB4 are highlighted in yellow. The last 57 amino acids that would lack according to the AtCCB4.2 gene model (Lezhneva et al., 2008) are underlined. Predicted cTPs are indicated in blue (dark blue: ChloroP; light blue: WoLF PSORT). (A) Immunoblot of eluates according to 1 mg chlorophyll. Proteins were detected by anti-HA-peroxidase and an anti-AtCCB4 antibody.

respectively; Supplemental table 1) indicating a HCF208/ AtCCB2-AtCCB4 interaction. The AtCCB4 peptides cover the entire protein as predicted from the AtCCB4.1 gene model (Figures 1 and 5 A). An alternative gene model (AtCCB4.2) encoding for a shorter protein lacking the last 57 amino acids (Figure 5 A, underlined) that was described earlier (Lezhneva et al., 2008) is therefore improbable. Besides AtCCB4 no other AtCCB protein co-purified with HCF208/ AtCCB2 under the chosen experimental conditions. Also, the overall weak recovery rate of peptides in all eluates did not allow conclusions on other, HCF208/ AtCCB2 eluate-specific proteins.

To confirm the HCF208/ AtCCB2-AtCCB4 interaction small-scale affinity purification of HCF208/ AtCCB2 from HCF208pAUL6 lines via IgG Sepharose using the above-mentioned controls was performed. Eluates were subjected to SDS-PAGE and immunoblot analyses. Using the AtCCB4-specific antibody, the protein was specifically detected in HCF208pAUL6 eluates but not in wild type and HCF107pAUL3 (Figure 5 B). This data substantiates the presence of a stable HCF208/ AtCCB2-AtCCB4 interaction as indicated by MS analyses.

Interactions of the AtCCB proteins in yeast split ubiquitin assays

Affinity purification of HCF208/ AtCCB2 indicated an interaction with AtCCB4. To further investigate whether, in addition to their functional interaction, the four AtCCB proteins physically interact with each other and with their target cytochrome b_6, a yeast two-hybrid assay was employed. Since all AtCCB proteins possess transmembrane domains (Figure 6 A; Lyska et al., 2007; Lezhneva et al., 2008) and HCF208/ AtCCB2 was experimentally shown to be tightly associated with membranes, the yeast split ubiquitin system was used, which is also applicable for chloroplast proteins (Johnson and Varshavsky, 1994; Pasch et al., 2005). In this assay, bait and prey proteins are fused to the C- or N-terminal half of ubiquitin (Cub and Nub), respectively, which is reconstituted upon protein interaction. Hence, a transcription factor fused to Cub is cut off and migrates to the nucleus to switch on reporter gene expression. Since the Cub and Nub moieties must be exposed to the cytoplasm (corresponding to the chloroplast stroma) conclusions on protein topologies can be drawn from yeast split ubiquitin experiments.

Interactions were displayed by the ability of yeast cells co-transformed with bait and prey constructs to grow autotrophically on stringent selection media. No interaction of AtCCB1, neither as Cub nor as Nub fusion, with any of the other AtCCB proteins or cytochrome b_6 was detected (Figure 6 B). However, the interaction of AtCCB1 fused to Cub at its C-terminus with the positive control $Alg5^{NubI}$ confirmed the predicted stromal exposition of the AtCCB1 C-terminus (Figures 6 A and B; Lezhneva et al., 2008). Cytochrome b_6, which was only used as a prey in the studies, did not interact with any of the AtCCB proteins either (Figure 6 B).

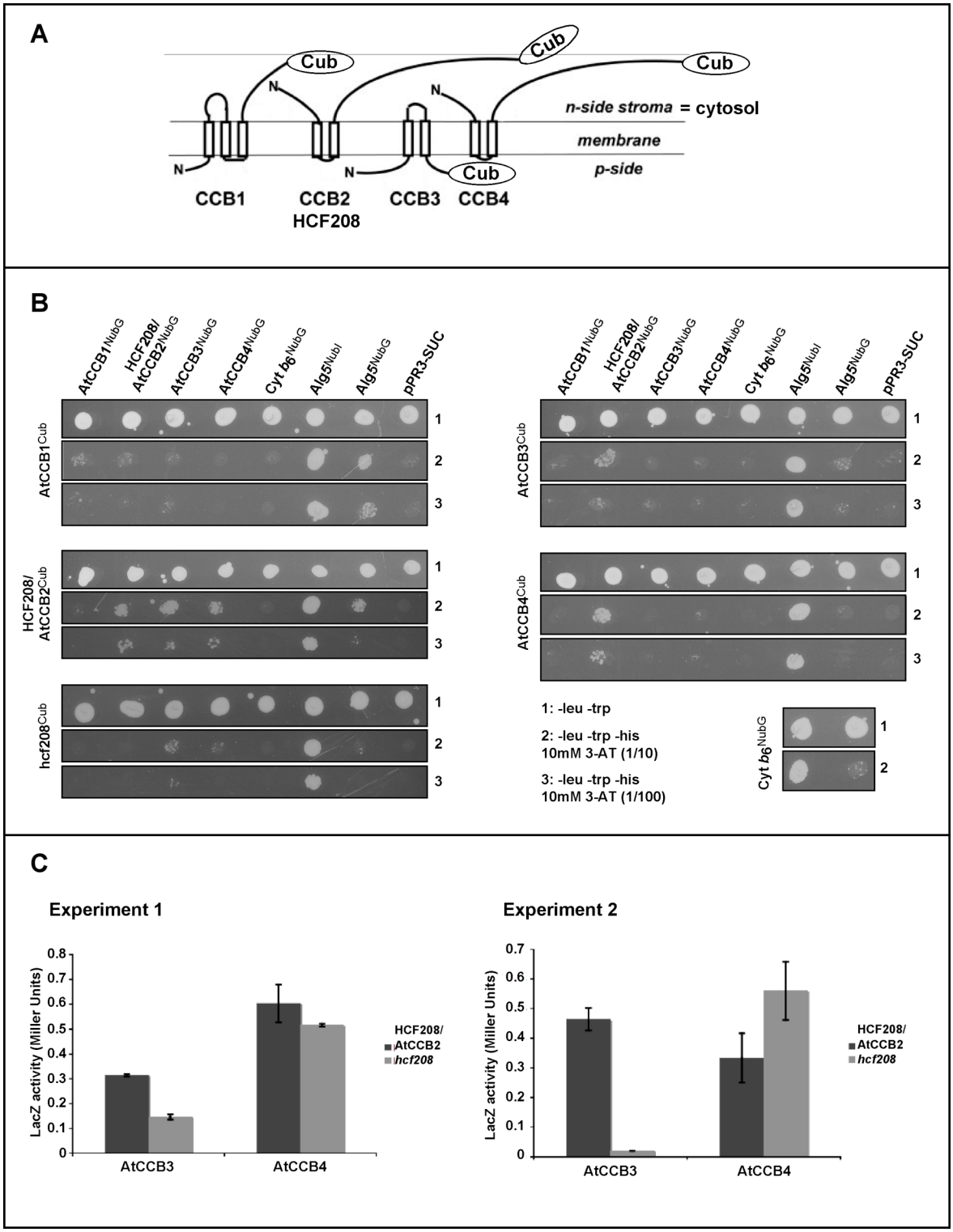


Figure 6: Yeast Split Ubiquitin analysis of AtCCB1, HCF208/ AtCCB2, AtCCB3, AtCCB4, hcf208 and cytochrome b_6.

(A) Predicted membrane orientations of AtCCB1, HCF208/ AtCCB2, AtCCB3 and AtCCB4 (Modified from Lezhneva et al., 2008) and location of the Cub-TF moiety. (B) Interaction analysis. Co-transformed yeast cells harboring bait- and prey plasmids were selected on media lacking leucin and tryptophane; interactions were tested on media lacking leucine, tryptophane and histidine. –leu –trp-his media contained 10 mM 3-AT to suppress *HIS3*-autoactivation. Two dilutions (1/10 and 1/100) were tested. (C) β-galactosidase (*lacZ*) assay of interactions between HCF208/ AtCCB2 and the hcf208 mutant protein, respectively, with AtCCB3 and AtCCB4. LacZ activity was calculated in Miller Units. Background activities of co-transformation with the negative control prey vector pDL2-Alg5(NubG) were subtracted.

Co-transformation of HCF208/ AtCCB2 with both AtCCB3 and AtCCB4 enabled the yeast to grow on selection media (Figure 6 B). The HCF208/ AtCCB2-AtCCB3 and -AtCCB4 interactions occurred independent of the proteins being employed as bait or prey, therefore substantiating this finding (Figure 6 B). An interaction of HCF208/ AtCCB2 with itself was indicated by autotrophy of co-transformed yeast (Figure 6 B). However, auto-activation tested by co-transformation with the negative control $Alg5^{NubG}$, was high when HCF208/ AtCCB2 served as a bait (Figure 6 B) making no clear conclusion possible. As for AtCCB1, the predictions of C-termini exposed to the stroma were confirmed for HCF208/ AtCCB2 and AtCCB4 since C-terminal Cub fusions were applied in this study (Figures 6 A and B). Surprisingly, the C-terminal Cub fusion of the AtCCB3 protein was also able to restore ubiquitin upon interaction with other proteins, which stands in disagreement to the predicted topology (C-terminus in the lumen; Figures 6 A and B) indicating either an incorrect prediction or an altered topology of AtCCB3 when heterologously expressed in yeast.

The *hcf208* mutant allele encodes for a protein carrying an amino acid exchange in the first transmembrane domain of HCF208/ AtCCB2 that is unable to facilitate cytochrome b_6 maturation (Lyska et al., 2007). To test whether this mutation influences the previously determined interactions, both the wild-type and mutant allele serving as baits were analyzed. On selection media, yeast co-transformed with hcf208 and both AtCCB3 and AtCCB4 were able to grow but interactions appeared weaker since cell growth was restricted when using higher dilutions (Figure 6 B). To test this observation, interactions were quantified using a β-galactosidase (*lacZ*) assay. Quantification revealed that AtCCB4 interactions were not altered using mutated hcf208 compared to the wild-type allele (Figure 6 C), but the hcf208-AtCCB3 interaction was drastically reduced by at least two fold when compared to wild type (Figure 6 C) pointing to a role of the first transmembrane domain in HCF208/ AtCCB2-AtCCB3 interaction.

In summary, yeast split ubiquitin data suggests that HCF208/ AtCCB2 is a central player in the pathway of heme attachment to cytochrome b_6 since it interacted with two other proteins of this pathway, AtCCB3 and AtCCB4. In contrast, there was no evidence for direct interactions between other AtCCB proteins or of cytochrome b_6 with any AtCCB protein.

Discussion

The cytochrome b_6f complex of organisms performing oxygenic photosynthesis contains the two c-type cytochromes cytochrome *f* and cytochrome b_6. Heme is attached to cytochrome *f* via two thioether bonds to a characteristic consensus sequence CXXCH and ligation occurs at the lumenal, electropositive side of the thylakoid membrane by system II. In contrast, cytochrome b_6 binds heme c_n at the stromal, electronegative side of the thylakoid membrane in an unusual manner via a single cysteine residue (Cys35). Four nuclear loci (CCB1-CCB4) from *Chlamydomonas* are required for covalent heme attachment cytochrome b_6 and are referred to as c-type cytochrome maturation system IV, which is genetically conserved in all organisms performing oxygenic photosynthesis (Kuras et al., 2007). Analyses on the *Arabidopsis* nuclear mutant *hcf208*, which is affected in the *Arabidopsis* homolog of CCB2 (HCF208/ AtCCB2), confirmed the functional conservation of system IV between *Chlamydomonas* and *Arabidopsis* (Lyska et al., 2007). This was substantiated by an independent study by Lezhneva et al. (2008) who described a heme c_n-deficient phenotype of T-DNA insertion mutants of HCF208/ AtCCB2 (*atccb2*_SALK), AtCCB1 (*atccb1*_SALK) and AtCCB4 (*atccb4*_SALK). In the present study, the *atccb1*_SALK and *atccb4*_SALK mutants were shown to completely lack respective proteins and therefore display knock-out mutant lines.

A comparison of mutant phenotypes of the *atccb* mutants and *hcf208* constrained an overall equal contribution of the AtCCB factors in cytochrome b_6 maturation since the lines exhibited similar changes in chlorophyll fluorescence and lacked cytochrome b_6 peroxidase activity. However, in *atccb1*_SALK knock-out-lines cytochrome b_6 accumulated to lower levels than in all other mutant lines. Since the decrease of cytochrome b_6 originates from protein instability rather than reduced expression (Lyska et al., 2007; Saint-Marcoux et al., 2009), AtCCB1 could have a stabilizing effect on apocytochrome b_6 to some extent. Mutants affected in HCF208/ AtCCB2 and AtCCB4 accumulated equal cytochrome b_6 levels and, interestingly, the AtCCB4 protein did not accumulate in *hcf208* and *atccb2*_SALK. Accordingly, the CCS1 protein of c-type cytochrome maturation system II does not accumulate in mutants affected in CcsA with which it forms a heme transport and attachment complex (Dreyfuss et al., 2003; Hamel et al., 2003; Kranz et al., 2009). Taking this into account the interdependency of HCF208/ AtCCB2-AtCCB4 accumulation points to a close functional, and eventually physical interaction of the two proteins.

Membrane association of HCF208/ AtCCB2 and topologies of the AtCCB proteins

Different prediction algorithms indicate that the four CCB proteins from both *Arabidopsis* and *Chlamydomonas* are integral membrane proteins possessing three (CCB1) or two (HCF208/ CCB2, CCB3, CCB4) transmembrane helices (Kuras et al., 2007; Lyska et al., 2007; Lezhneva et al., 2008). HCF208/ CCB2 and CCB4 are predicted to expose both termini to the stroma with the C-terminus forming a large soluble domain. The C-terminus of CCB1 is also predicted in the stroma, whereas its N-terminus and both termini of CCB3 are supposed in the lumen. Immunolocalization studies using HCF208-HA/Strep-tag*III* confirmed tight association HCF208/ AtCCB2 with chloroplast membranes in *Arabidopsis*. Accordingly, the CCB1-4 proteins were resistant to dissociating treatments of membranes in *Chlamydomonas* (Saint-Marcoux et al., 2009). A hint on the orientation of the HCF/ AtCCB C-termini is given by yeast split ubiquitin assays performed in this study. The HCF208/ AtCCB1-4 proteins C-terminally fused with the Cub moiety interacted with the positive control ($Alg5^{NubI}$) indicating their C-terminito be exposed to the cytosol, which corresponds to the stromal side of chloroplast membranes. This stands in accordance with the predicted topologies of AtCCB1, HCF208/ AtCCB2 and AtCCB4 and with the transmembrane topologies of *Chlamydomonas* CCB1, CCB2 and CCB4, which were analyzed by protease (trypsin) accessibility and yeast split ubiquitin assays (Saint-Marcoux et al., 2009). For AtCCB3, the data presented here is contradictory to the predicted orientation. The C-terminus of CCB3 in *Arabidopsis* and *Chlamydomonas* is supposed to be exposed to the lumen and Saint-Marcoux et al. (2009) provided experimental evidence supporting this prediction. However, the lumenal localization of the C-terminus of CCB3 was only indicated by the inability to detect a CCB3 protease digestion product lacking three C-terminal amino acids cleaved off by trypsin (Saint-Marcoux et al., 2009). which is hard to detect in Western blot analysis. Therefore, the C-terminus of AtCCB3 may also be exposed to the stroma as indicated from yeast split ubiquitin experiments presented here. Nevertheless, it cannot be excluded that AtCCB3 topology is somehow different when the protein is expressed in yeast.

Topologies of the CCB proteins forming system IV in *Arabidosis* and *Chlamydomonas* stand in accordance with a function in heme ligation to cytochrome b_6 at the stromal side of the thylakoid membrane. As described by Kuras et al. (2007) the majority of conserved amino acids in all CCB proteins are located at the stromal side of the membrane indicating that these sites may play an important role in holo-cytochrome b_6 formation. The C-terminal soluble extensions of CCB1, CCB2 and CCB4 protruding into the stroma may function in heme delivery and ligation to apo-cytochrome b_6 similar to the CcsA/ CCS1 proteins operating in system II in the lumen (Kranz et al., 2009). According to CCB1, CCB2 and CCB4, CcsA and CCS1 are integral membrane proteins with large domains exposed to the

soluble compartment. However, CcsA exhibits a heme binding site in one of these domains (WWD motif; Rurek, 2008; Kranz et al., 2009), which is absent from the CCB proteins. Also, the CCBs lack conserved histidine residues that could be involved in transient or covalent heme binding (Saint-Marcoux et al., 2009).

Stable and transient interactions of the HCF208/ AtCCB2 protein

Based on their origin from a unique cyanobacterial ancestor (CCB2/4; Kuras et al., 2007; Supplemental Figure S3) and interdependent accumulation, the paralogs HCF208/ CCB2 and CCB4 were supposed to form heterodimers. Co-immunoprecipitation and yeast split ubiquitin analyses presented here disclosed a stable, direct interaction between HCF208/ AtCCB2 and AtCCB4, which is confirmed by the identification of an HCF208/ AtCCB2-containing dimeric complex at ~70 kDa in 2D-BN-PAGE. This stable heterodimer was also identified for the homologous Chlamydomonas proteins (Saint-Marcoux et al., 2009).
In yeast, HCF208/ AtCCB2 also directly interacted with AtCCB3. Since AtCCB3 was absent from HCF208/ AtCCB2-eluates after affinity purification the HCF208/ AtCCB2-AtCCB3 interaction may be only weak and/ or transient *in planta*. However, using 2D-BN-PAGE an HCF208/ AtCCB2-containing complex of ~170 kDa was identified, which, as at least in *Chlamydomonas,* includes CCB3, CCB4 and cytochrome b_6 (Saint-Marcoux et al., 2009). Unlike affinity purification, 2D-BN-PAGE was performed using material from young plants (two-week-old) suggesting an interconnection between the HCF208/ AtCCB2-AtCCB3 interaction and the developmental stage of the chloroplast. The cytochrome b_6f complex is very stable with lifetimes in the range of several days (Gong et al., 2001), wherefore de novo synthesis and consequently cytochrome b_6 maturation may be diminished in mature plants. This is supported by the findings of Saint-Marcoux et al. (2009) who could not detect the ~170 kDa complex in wild type, but only in *Chlamydomonas* mutants affected in the heme c_n-binding site of cytochrome b_6 (petB-C35V) arresting in the last step before heme ligation. Therefore, the HCF208/ AtCCB2-AtCCB3 interaction is transient and only occurs during heme c_n attachment to cytochrome b_6, whereas HCF208/ AtCCB2 and AtCCB4 form a stable, permanent heterodimer.
Using the mutated *hcf208* mutant allele of HCF208/ AtCCB2, which encodes for a protein carrying an amino acid exchange in the first transmembrane helix (Gly68 → Arg68; Lyska et al., 2007), the AtCCB4-interaction was maintained, whereas the affinity to AtCCB3 was drastically reduced. Alignment of HCF208/ AtCCB2 homologs from *Arabidopsis*, rice, *Chlamydomonas* and the CCB2/4 protein from the cyanobacteria *Anabena variabilis* and *Synechococcus sp.* previously indicated that Gly68 is not present in cyanobacteria and therefore not conserved (Lyska et al., 2007). However, when including the sequences of the

paralog CCB4 into alignments Gly68 appears as a highly conserved residue in the first transmembrane helix of all CCB2, CCB4 and CCB2/4 proteins (Supplemental Figure S3). Glycine residues in transmembrane helices are supposed to be predominantly oriented towards helix-helix interfaces and to mediate their interactions (Javadpour et al., 1999). Therefore, interaction between HCF208/ AtCCB2 and AtCCB3 requires the first transmembrane helix of HCF208/ AtCCB2 involving Gly68, whereas the HCF208/ AtCCB2-AtCCB4 interaction is mediated by a different region, presumably the second transmembrane helix.

An interesting aspect rising from the yeast split ubiquitin experiments is the lack of a direct interaction between AtCCB3 and AtCCB4. HCF208/ AtCCB2 and AtCCB4 derive from a common cyanobacterial ancestor protein (CCB2/4) to which AtCCB4 is more closely related than HCF208/ AtCCB2 (Kuras et al., 2007; Supplemental Figure S3). Since heme c_n is also present in the cytochrome b_6f complex of cyanobacteria (Kurisu et al., 2003) the CCB2/4 protein may have a similar function like CCB2 and CCB4 in chloroplasts. Based on the presented data, this function would have to include interaction with CCB3. Thus, after gene duplication, the CCB4 protein supposedly lost its ability to directly interact with CCB3, whereas this function was maintained for CCB2. Whether CCB2 and CCB4 also have distinct functions in different steps of heme-attachment to cytochrome b_6 and if they have to strictly form a heterodimer for their functionremains to be elucidated.

The role of CCB proteins in holo-cytochrome b_6 formation in *Arabidopsis* and *Chlamydomonas*

Interaction studies revealed that CCB2 and CCB4 form a stable heterodimer, which interacts with CCB2 and with cytochrome b_6 presumably at early developmental stages during cytochrome b_6f complex formation. Analyses in *Chlamydomonas* disclosed further interactions and interaction dynamics, that could not be analyzed in *Arabidopsis* due to the lack of an AtCCB3 mutant and antibody. By co-immunoprecipitation and 2D-BN-PAGE, evidence was provided for a transient CCB3-cytochrome b_6 interaction and subsequent binding of the CCB2-CCB4 heterodimer (Saint-Marcoux et al., 2009) thereby forming the abovementioned ~170 kDa CCB2-CCB3-CCB4-cytochrome b_6 complex. Also CCB1 co-precipitated with cytochrome b_6 and this interaction was supposed to occur before the formation of the CCB2-CCB3-CCB4-cytochrome b_6 complex (Saint-Marcoux et al., 2009). However, in yeast split ubiquitin assays performed in this study and by Saint-Marcoux et al. (2009) no interactions between cytochrome b_6 and AtCCB3 or AtCCB1 were detected, which may indicate the requirement of additional (co-) factors that are only present *in planta*.

According to the reduced amounts of cytochrome b_6 in the *atccb1*_SALK mutant when compared to mutants affected in other AtCCB proteins, and its association with cytochrome b_6 as a first step in maturation of this protein, the function of CCB1 may be cytochrome b_6-stabilization/ chaperoning prior to attachment of heme. The complex formed by CCB2, CCB3 and CCB4 is supposed to display the "heme ligation complex" that delivers and incorporates heme c_n to cytochrome b_6 (Saint-Marcoux et al., 2009).

Comparison of c-type cytochrome maturation system IV displayed by the CCB proteins with known components of systems I and II revealed no common features of these pathways (Saint-Marcoux et al., 2009). These systems consist of several components mediating heme delivery, heme and cysteine ligand reduction and ligation of heme to the apocytochrome (Kranz et al., 2009). Since heme cn attachment occurs on the electronegative, stromal side of the thylakoid membrane, where heme is synthesized no transport or thiooxidation/ thioreduction mechanisms are required. Therefore, a system for stabilization/ chaperoming of cytochrome and heme ligation may be sufficient for heme cn attachment. However, it cannot be excluded that other so far unidentified proteins also act in this pathway. The identification of a ~ 300 kDa complex in 2D-BN-analyses performed in this study may reflect association of at least HCF208/ AtCCB2 with other factors.

Taken together, the data presented for the *Arabidopsis* CCB proteins coincide with the tentative model for holo-cytochrome b_6 formation via CCB1-CCB4 based on analyses from *Chlamydomonas* (Saint-Marcoux et al., 2009) and confirm a functional conservation of this pathway. Further, formation of the putative heme ligation complex and therefore cytochrome b_6f complex formation was found to be likely dependent on the developmental stage of plants. Future analyses on the CCB complexes will help revealing the temporal changes of cytochrome b_6f complex synthesis and eventually identify new components of the system IV pathway for covalent heme attachment to cytochrome b_6.

Materials and Methods

Growth conditions and mutant selection

The mutant *hcf208* was selected from a collection of EMS-induced mutants (Lyska et al., 2007). The *ccb1*_GK mutant was isolated from the line GK045H07 of the GABI-Kat collection (Rosso et al., 2003). The *atccb1*_SALK, *atccb2*_SALK and *atccb4*_SALK mutants were isolated from the SALK collection and provided by Catherine de Vitry (CNRS, France, Paris; Lezhneva et al., 2008). The transgenic lines HCF208pAUL2, HCF208pAUL6 and HCF107pAUL3 were described in Lyska et al. (2011). All mutant and transgenic lines are all in *Arabidopsis* Col-0 background.

Surface sterilized seeds were plated on 0.5x MS medium (Murashige and Skoog, 1962) supplemented with 1% (w/v) sucrose for mutant selection and 2D-BN-PAGE. For all other purposes seeds were sown on Floraton I-soil (Florogard, Oldenburg, Germany).

Seeds were stratified in 4°C for two to five days an then grown in a growth chamber at a constant temperature of 21°C and a PFD of 50-70 μmol s^{-1} m^{-2}. For mutant selection, 2D-BN-PAGE, protein extraction and seed production plants were grown under long day conditions with a 16 h light/ 8 h darkness period. Plants used for chloroplast isolation and affinity purification were grown under short day conditions (8 h light/ 16 h darkness).

Selection of mutant plants exhibiting high chlorophyll fluorescence phenotype was performed in the dark under UV light (Meurer et al., 1996).

Genotyping of mutant lines

Genotyping of plants was necessary for identification of heterozygous *hcf/ ccb*-mutant lines for seed production and for verification of mutant phenotypes.

Genotyping of the T-DNA-insertion lines *atccb1*_SALK, *atccb2*_SALK and *atccb4*_SALK lines was performed as previously described by Lezhneva et al. (2008). Positions of T-DNA insertions and primers are indicated in Figure 1 and primer sequences are listed in Supplemental table 2. The *atccb1*_GK mutation in line GK045H07 is caused by a T-DNA insertion in the 5´UTR of the gene. Genotyping was performed using primer pairs GK045H07-TDNA-LB-H and GK045H07-gen-R for the mutant/ T-DNA allele and GK045H07-gen-H and GK045H07-gen-R for the wild-type allele in two separate PCRs.

The *hcf208* mutation was induced by EMS-mutagenesis, which resulted in the exchange of guanine to adenine in the second exon at 291 bp downstream of the start-codon. For genotyping, a forward primer was designed (208geno*Bcl*I-H) carrying an incomplete *Bcl*I restriction site ending one nucleotide upstream the point mutation (Figure 1). The *Bcl*I

restriction site is completed in combination with the adenine from the mutant allele leading to cleavage of the product resulting from PCR with primers 208geno*Bcl*I-H and 208geno-R (Supplemental table 2; Figure 1). The PCR product resulting from amplification of the wild-type allele is not recognized.

Fluorescence measurements

Chlorophyll fluorescence and was imaged with a closed FluorCam FC 800-C controlled by FluorCam 6 software (Photon Systems Instruments) on 3-week-old plants. Experiments were performed using pre-designed quenching protocol provided in the software. Nomenclature of fluorescence parameters was adapted according to van Kooten and Snel (1990).

SDS-PAGE of proteins, Coomassie staining of gels, immuno blotting and heme staining

Crude leaf proteins (Meurer et al., 1996) from wild-type and mutant plants were separated in discontinuous gel systems after Schägger (Schägger and von Jagow, 1987) or Laemmli (1970). Immunodecoration of electroblotted proteins followed the methods described by Meurer et al. (1996). The fusion proteins HCF208-HA/Strep-tag*III*, HCF208-HA/Strep-tag*III*/ProteinA and HCF107-HA/Strep-tag*III*/ProteinA were detected by anti-HA-peroxidase (high affinity (3F10); Roche, Mannheim, Germany). SDS-polyacrylamide gels were stained using PageBlue^{TM} Protein Staining Solution (Fermentas, St. Leon-Rot, Germany) based on the Coomassie Brilliant Blue G-250 dye.
For the detection of heme peroxidase activity proteins were blotted onto nitrocellulose membranes, incubated with the reagent of the chemoluminescence assay (SuperSignal® West Femto Maximum Sensitivity kit (Thermo Scientific, Rockford, USA)) and exposed in a LAS-4000 mini (GE Healthcare, Piscataway, USA) imaging system.

Immunolocalization studies

Plant material for isolation of intact chloroplasts was grown under short day conditions for four to six weeks and kept in darkness for 20 hours prior to harvesting. Plant material was homogenized in homogenization buffer (330 mM sorbitol; 50 mM Hepes/ KOH, pH 7.8; 2 mM EDTA; 1 mM $MgCl_2$; 0.25% (w/v) BSA; 0.1% (w/v) sodium ascorbate) in a Waring blender. Cell debris was removed by Miracoth filtration. The suspension was centrifuged at 4°C until a speed of 5900 x *g* was reached. The pellet was resuspended in homogenization buffer and applied to a 85/45% percoll gradient (85 % (v/v) Percoll; 2.55 % (w/v) PEG 6000; 0.85 %

(w/v) BSA; 0.85 % (w/v) Ficoll; 330 mM sorbitol; 50 mM Hepes/KOH, pH 7.8; 2 mM EDTA; 1 mM $MgCl_2$; 2 mM DTE and 45 % (v/v) Percoll; 1.35 % (w/v) PEG 6000; 0.45 % (w/v) BSA; 0,45 % (w/v) Ficoll; 330 mM sorbitol; 50 mM Hepes/KOH, pH 7.8; 2 mM EDTA; 1 mM $MgCl_2$; 2 mM DTE). Intact chloroplasts were collected, washed with resuspension buffer (330 mM sorbitol; 50 mM Hepes/ KOH (pH 7.8); 10 mM $MgCl_2$), resuspended in lysis buffer (10 mM Hepes/ KOH (pH 7.8); 10 mM $MgCl_2$; 25 mM KCl) and incubated on ice for 10 minutes. Membranes and stroma were separated by 15 minutes of centrifugation at 15000 x *g* and 4°C. Supernatants were precipitated with 15% (w/v) trichloroacetic acid.

Strength of the membrane association of HCF208-HA/Strep-tag*III* was analyzed by resuspending pelleted membranes in lysis buffer or lysis buffer supplemented with 200 mM NaCl or 250 mM Na_2CO_3, respectively. Membranes were centrifuged again for 15 minutes at 15000 x *g* and 4°C. Supernatants were precipitated with 15% (w/v) trichloroacetic acid.

Samples were separated by SDS-PAGE and subsequently subjected to immunoblot analysis.

Two-dimensional Blue Native/ SDS-PAGE

Blue Native PAGE was performed according to Schägger et al. (1994). 2-week-old seedlings were homogenized in lysis buffer (10 mM Hepes/ KOH, pH 7.8; 10 mM $MgCl_2$; 25 mM KCl) in a Waring blender. After filtration through two layers of Miracoth, membranes were centrifuged 4°C until a speed of 5900 x *g* was reached. The membranes were resuspended in ACA buffer (750 mM ε-aminocaproic acid; 50 mM Bis-Tris, pH 7; 0,5 mM EDTA) and solubilized with 1% (w/v) digitonin in ACA buffer for 30 minutes at 4°C and a chlorophyll concentration of 1 mg ml^{-1}. Unsolubilized material was removed by centrifugation (20 min at 4°C and 15,000 x *g*). The supernatant was supplemented with 0.5% (w/v) n-dodecyl-β-D-maltoside and one-tenth volume of sample buffer (50 mM Bis-Tris, pH 7; 750 mM ε-aminocapronic acid; 30% (w/v) sucrose; 5% (w/v) Coomassie Serva G). Aliquots corresponding to 20 µg of chlorophyll were analyzed on a 6 to 12% gradient gel by Blue Native PAGE. Separated protein complexes were analyzed by SDS-PAGE according to Laemmli (1970) in the second dimension. Immunoblot analysis was performed as described above.

Affinity purification of HCF208-HA/Strep-tag*III*/ProteinA

Leaves of 4- to 6-week-old wild-type, HCF197pAUL3 and HCF208pAUL6 plants were homogenized in lysis buffer (10 mM Hepes/ KOH, pH 7.8; 10 mM $MgCl_2$; and 25 mM KCl). Cell debris was separated by Miracloth filtration. The suspension was centrifuged at 4°C until

a speed of 5,900 x *g* was reached. Pelleted membranes were resuspended in Tris-buffered lysis buffer and solubilized with 1% (w/v) n-dodecyl-ß-D-maltoside (Calbiochem) for 30 minutes at 4°C and a chlorophyll concentration of 1 mg ml^{-1}. Unsolubilized material was removed by centrifugation (20 minutes at 4°C and 15,000 x *g*). Supernatants were incubated with one tenth volume of IgG-Sepharose 6 Fast Flow (GE Healthcare, Piscataway, USA) pre-equilibrated with washing buffer (10 mM Hepes/ KOH, pH 7.8; 150 mM NaCl; 0.05% (w/v) n-dodecyl-β-D-maltoside) for 1 hour at 4°C on an overhead mixer. Subsequently, IgG-Sepharose was washed with 20 volumes of washing buffer and proteins were eluted with 1 matrix-volume of 1% (w/v) SDS. Eluates were precipitated with 15% (w/v) trichloroacetic acid and analyzed by SDS-PAGE according to Laemmli (1970). Gels were subjected to Coomassie staining or immunoblot analyses.

Sample preparation for mass spectrometry

Protein samples obtained from affinity purification were subjected to SDS-PAGE according to Laemmli (1970) and each lane was diced into small pieces. In gel digestion was performed according to a modified protocol from Shevchenko et al. (2006). After digestion, dried peptides were resuspended in 3% acetonitrile 0.2% trifluoretic acid and cleaned up using Sepak Cartridges (Waters, Milford, Massachusetts, USA). Clean samples were dried and resuspended in 12μl 3% acetonitrile 0.2% formic acid for Mass spectrometry.

Analysis by LC-ESI-MS/MS (Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry) and data mining

Dried peptides were resuspended in 3% acetonitrile, 0.2% formic acid and analyzed on a LTQ Orbitrap mass spectrometer (Thermo Fischer Scientific, Bremen, Germany) coupled to an Eksigent-Nano-HPLC system (Eksigent Technologies, Dublin (CA), USA). Peptide mixtures were loaded onto laboratory made capillary columns (75 µm inner diameter (BGB Analytik, Boeckten, Switzerland), 8 cm length, packed with Magic C18 AQ beads, 3 µm, 100 Å (Michrom BioResources, Auburn, CA, USA). Peptides were eluted from the column by an increased acetonitrile concentration in the mobile phase from 5% acetonitrile, 0.2% formic acid to 40% acetonitrile, 0.2% formic acid over 74 minutes, followed by a 10 minutes wash step at 5% acetonitrile, 0.2% formic acid. Full-scan MS spectra (300−2000 m/z) were acquired with a resolution of 60000 at 400 m/z after accumulation to a target value of 500000. Collision induced dissociation (CID) MS/MS spectra were recorded in data dependent manner in the ion trap from the six most intense signals above a threshold of 500, using a normalized collision energy of 35% and an activation time of 30 ms. Charge state

screening was enabled and singly charge states were rejected. Precursor masses already selected for MS/MS were excluded for further selection for 120 s and the exclusion window was set to 20 ppm. The size of the exclusion list was set to a maximum of 500 entries.
MS/MS spectra were searched with Mascot (Matrix Science, London, UK) version 2.2.04 against the *Arabidopsis thaliana* TAIR9 protein database (download on June 29th, 2009) with a concatenated decoy database supplemented with contaminants (67'079 entries). The search parameters were: requirement for tryptic ends, one missed cleavage allowed, mass tolerance = ± 5 ppm. Beside carbamidomethylation of cysteines as fixed modification, oxidation of methionine was included as variable modification. Individual ions scores higher than 22 indicated identity or extensive homology. Peptide identification was accepted with a minimal Mascot ion score of 23 and a Mascot expectation value below 0.05. The spectrum false discovery rate was calculated by dividing the number of decoy database spectrum assignments by the number of spectrum assignments in the final dataset and was below 1 % for all measured experiments.

Yeast split ubiquitin assays

The yeast split ubiquitin assay is a yeast two-hybrid assay that allows monitoring of interactions between transmembrane proteins. It is based on the ability of reconstitution of the C-terminal half of ubiquitin (Cub) with the N-terminal half of ubiquitin (NubI), even when fused to different, non-interacting proteins. A mutated version of NubI (Ile3 → Gly3; NubI → NubG) possesses lower affinity to Cub. Using the NubG moiety, ubiquitin can only be reconstituted when the fusion proteins interact with each other. The Cub moiety binds a transcription factor (TF), which towards interaction of the fusion proteins is released and migrates to the nucleus to activate transcription of reporter genes. The Cub-TF and Nub moieties fused to analyzed proteins must be located on the cytosolic side of the membrane to allow migration of the released transcription factor to the nucleus.
Interaction studies were performed using the Dual-Membrane-kit 3 (Dualsystems Biotech AG, Schlieren, Switzerland). Sequences encoding mature AtCCB1, HCF208/ AtCCB2, AtCCB3 and AtCCB4 proteins (without cTPs as predicted by WoLF PSORT; Supplemental Figure 1) and cytochrome b_6 were amplified from cDNA obtained from wild-type RNA using primers listed in Supplemental table 2. The coding sequence of mature hcf208 was amplified from cDNA obtained from *hcf208* mutant plants using the same primers as for HCF208/ AtCCB2. Via an *Sfi*I restriction site all fragments were introduced to the pBT3-SUC and the pPR3-SUC vectors harboring the coding sequences for Cub-TF and NubG, respectively for C-terminal fusions. Bait and prey constructs were co-transformed into the yeast NMY51 strain and transformants were selected on synthetic "drop-out" media lacking leucin and

tryptophane (SD-leu-trp) for three days at 37°C. Interactions were monitored by the ability of co-transformed yeast cells to grow on SD-leu-trp media lacking histidine in the presence of 10 mM 3-aminotriazole. For quantification of interactions, the expression of the lacZ reporter gene was monitored by the rate of ONPG (o-nitrophenyl-β-D-galactopyranoside) metabolized by β-galactosidase reflected by an increase of the absorption at 420 nm. LacZ activity was calculated in Miller Units (Miller Units LacZ = 1000 x OD_{420} / (5 x t x OD_{600}), t being the incubation time in minutes).

Co-transformation of bait constructs with a construct encoding for an ER-localized protein from yeast (Alg5) fused to the NubI (pAI-Alg5) moiety served as a positive prey control displaying the correct orientation of the Cub-TF fusions. Constructs expressing Alg5 fused to NubG (pDL2-Alg5) and the empty prey vectors (pPR3-SUC) were used as negative controls.

Accession numbers

Accession numbers of proteins described in the text are as follows: NP_200024 (HCF208/ AtCCB2), NP_566797 (AtCCB1), NP_198461 (AtCCB3), NP_974051 (AtCCB4), ABP57441 (CrCCB1), ABP57442 (CrCCB2), ABP57443 (CrCCB3), ABP57444 (CrCCB4), YP_322396 (AvCCB2/4), and YP_001735959 (SynCCB2/4).

Literature

Allen, J. (2002) Photosynthesis of ATP-electrons, proton pumps, rotors, and poise. *Cell*, **110**, 273-276.

Breyton, C., Tribet, C., Olive, J., Dubacq, J.P. and Popot, J.L. (1997) Dimer to monomer conversion of the cytochrome b6 f complex. Causes and consequences. *J Biol Chem*, **272**, 21892-21900.

Cramer, W.A. and Zhang, H. (2006) Consequences of the structure of the cytochrome b6f complex for its charge transfer pathways. *Biochim Biophys Acta*, **1757**, 339-345.

Dreyfuss, B.W., Hamel, P.P., Nakamoto, S.S. and Merchant, S. (2003) Functional analysis of a divergent system II protein, Ccs1, involved in c-type cytochrome biogenesis. *J Biol Chem*, **278**, 2604-2613.

Fujiki, Y., Hubbard, A.L., Fowler, S. and Lazarow, P.B. (1982) Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasmic reticulum. *J Cell Biol*, **93**, 97-102.

Giege, P., Grienenberger, J.M. and Bonnard, G. (2008) Cytochrome c biogenesis in mitochondria. *Mitochondrion*, **8**, 61-73.

Gong, X.S., Chung, S. and Fernandez-Velasco, J.G. (2001) Electron transfer and stability of the cytochrome b6f complex in a small domain deletion mutant of cytochrome f. *J Biol Chem*, **276**, 24365-24371.

Hamel, P.P., Dreyfuss, B.W., Xie, Z., Gabilly, S.T. and Merchant, S. (2003) Essential histidine and tryptophan residues in CcsA, a system II polytopic cytochrome c biogenesis protein. *J Biol Chem*, **278**, 2593-2603.

Huang, D., Everly, R.M., Cheng, R.H., Heymann, J.B., Schagger, H., Sled, V., Ohnishi, T., Baker, T.S. and Cramer, W.A. (1994) Characterization of the chloroplast cytochrome b6f complex as a structural and functional dimer. *Biochemistry*, **33**, 4401-4409.

Javadpour, M.M., Eilers, M., Groesbeek, M. and Smith, S.O. (1999) Helix packing in polytopic membrane proteins: role of glycine in transmembrane helix association. *Biophys J*, **77**, 1609-1618.

Johnsson, N. and Varshavsky, A. (1994) Split ubiquitin as a sensor of protein interactions in vivo. *Proc Natl Acad Sci U S A*, **91**, 10340-10344.

Joliot, P. and Joliot, A. (2006) Cyclic electron flow in C3 plants. *Biochim Biophys Acta*, **1757**, 362-368.

Junge, W. (1999) ATP synthase and other motor proteins. *Proc Natl Acad Sci U S A*, **96**, 4735-4737.

Kramer, D.M., Johnson, G., Kiirats, O. and Edwards, G.E. (2004) New fluorescence parameters for the determination of q(a) redox state and excitation energy fluxes. *Photosynth Res*, **79**, 209-218.

Kranz, R.G., Richard-Fogal, C., Taylor, J.S. and Frawley, E.R. (2009) Cytochrome c biogenesis: mechanisms for covalent modifications and trafficking of heme and for heme-iron redox control. *Microbiol Mol Biol Rev*, **73**, 510-528, Table of Contents.

Kuras, R., Saint-Marcoux, D., Wollman, F.A. and de Vitry, C. (2007) A specific c-type cytochrome maturation system is required for oxygenic photosynthesis. *Proc Natl Acad Sci U S A*, **104**, 9906-9910.

Kurisu, G., Zhang, H., Smith, J.L. and Cramer, W.A. (2003) Structure of the cytochrome b6f complex of oxygenic photosynthesis: tuning the cavity. *Science*, **302**, 1009-1014.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lezhneva, L., Kuras, R., Ephritikhine, G. and de Vitry, C. (2008) A novel pathway of cytochrome c biogenesis is involved in the assembly of the cytochrome b6f complex in arabidopsis chloroplasts. *J Biol Chem*, **283**, 24608-24616.

Lyska, D., Paradies, S., Meierhoff, K. and Westhoff, P. (2007) HCF208, a homolog of Chlamydomonas CCB2, is required for accumulation of native cytochrome b6 in Arabidopsis thaliana. *Plant Cell Physiol*, **48**, 1737-1746.

Lyska, D., Schult, K., Meierhoff, K. and Westhoff, P. (2011) pAUL: A Gateway-based vector system for adaptive expression and flexible tagging of proteins in Arabidopsis. In preparation.

Meurer, J., Meierhoff, K. and Westhoff, P. (1996) Isolation of high-chlorophyll-fluorescence mutants of Arabidopsis thaliana and their characterisation by spectroscopy, immunoblotting and northern hybridisation. *Planta*, **198**, 385-396.

Mosser, G., Breyton, C., Olofsson, A., Popot, J.L. and Rigaud, J.L. (1997) Projection map of cytochrome b6 f complex at 8 A resolution. *J Biol Chem*, **272**, 20263-20268.

Murashige, T and Skoog, F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant*, **15**, 473-497.

Pasch, J.C., Nickelsen, J. and Schunemann, D. (2005) The yeast split-ubiquitin system to study chloroplast membrane protein interactions. *Appl Microbiol Biotechnol*, **69**, 440-447.

Robertson, D.E., Farid, R.S., Moser, C.C., Urbauer, J.L., Mulholland, S.E., Pidikiti, R., Lear, J.D., Wand, A.J., DeGrado, W.F. and Dutton, P.L. (1994) Design and synthesis of multi-haem proteins. *Nature*, **368**, 425-432.

Rosso, M.G., Li, Y., Strizhov, N., Reiss, B., Dekker, K. and Weisshaar, B. (2003) An Arabidopsis thaliana T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant Mol Biol*, **53**, 247-259.

Rurek, M. (2008) Proteins involved in maturation pathways of plant mitochondrial and plastid c-type cytochromes. *Acta Biochim Pol*, **55**, 417-433.

Saint-Marcoux, D., Wollman, F.A. and de Vitry, C. (2009) Biogenesis of cytochrome b6 in photosynthetic membranes. *J Cell Biol*, **185**, 1195-1207.

Schagger, H., Cramer, W.A. and von Jagow, G. (1994) Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem*, **217**, 220-230.

Schagger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem*, **166**, 368-379.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*, **1**, 2856-2860.

Shikanai, T. (2007) Cyclic electron transport around photosystem I: genetic approaches. *Annu Rev Plant Biol*, **58**, 199-217.

Stroebel, D., Choquet, Y., Popot, J.L. and Picot, D. (2003) An atypical haem in the cytochrome b(6)f complex. *Nature*, **426**, 413-418.

van Kooten, O. and Sneel, J.F.H. (1990) The use of chlorophyll fluorescence nomenclature in plant stress physiology. *Photosynth Res*, **25,** 147-150.

Yamashita, E., Zhang, H. and Cramer, W.A. (2007) Structure of the cytochrome b6f complex: quinone analogue inhibitors as ligands of heme cn. *J Mol Biol*, **370**, 39-52.

Supplement

Supplemental Table 1:

Accession	Description	Number of Peptides 208 exp.1	208 exp.2	208 exp.3
AT1G59840	CCB4 (COFACTOR ASSEMBLY OF COMPLEX C)	6	10	3
AT5G52110	HCF208, CCB2 (HIGH CHLOROPHYLL FLUORESCENCE 208)	2	2	1
AT2G47610	60S ribosomal protein L7A (RPL7aA)	0	1	0
AT5G13510	ribosomal protein L10 family protein	0	1	0
AT2G19810	zinc finger (CCCH-type) family protein	0	1	0
AT3G21410	F-box family protein (FBW1)	0	1	0
AT5G36110	CYP716A1 electron carrier/ heme binding / iron ion binding / monooxygenase/ oxygen binding	1	1	0
AT5G55500	ATXYLT, XYLT (ARABIDOPSIS THALIANA BETA-1,2-XYLOSYLTRANSFERASE) xylosyltransferase	0	1	0
AT1G23020	ATFRO3, FRO3 ferric-chelate reductase	0	1	0
AT2G31240	tetratricopeptide repeat (TPR)-containing protein	0	1	1
AT5G16750	TOZ (TORMOZEMBRYO DEFECTIVE) nucleotide binding	0	1	0
AT4G19610	RNA binding / nucleic acid binding / nucleotide binding	0	1	0
AT1G12470	Pep3/Vps18/deep orange family protein	0	1	0
AT5G62410	SMC2, TTN3, ATCAP-E1, ATSMC4 | SMC2 (STRUCTURAL MAINTENANCE OF CHROMOSOMES 2) transporter	1	1	0
AT4G39450	unknown protein	0	1	0
AT1G22275	ZYP1b, ZYP1	0	1	0
AT1G46840	F-box family protein	1	1	0
AT1G48300	unknown protein	0	1	0
AT2G19440	glycosyl hydrolase family 17 protein	0	1	0
AT2G42510	molecular function unknown\tab INVOLVED IN: spliceosome assembly, nuclear mRNA splicing, via spliceosome LOCATED IN: cellular_component unknown CONTAINS InterPro DOMAIN: Survival motor neuron interacting protein 1	0	1	0
AT3G14205	phosphoinositide phosphatase family protein	0	1	0
AT3G16540	DegP11 (DegP protease 11) catalytic/ protein binding / serine-type endopeptidase/ serine-type peptidase	1	1	0
AT3G19530	unknown protein	0	1	0

Accession	Description	Number of Peptides		
		208 exp.1	208 exp.2	208 exp.3
AT3G42786	unknown protein	0	1	0
AT3G63450	RNA binding / nucleic acid binding / nucleotide binding	0	1	0
AT4G20740	pentatricopeptide (PPR) repeat-containing protein	1	1	1
AT5G27050	AGL101 transcription factor	0	1	0
AT5G62710	leucine-rich repeat family protein / protein kinase family protein	0	1	0
AT4G34620	SSR16 (SMALL SUBUNIT RIBOSOMAL PROTEIN 16) structural constituent of ribosome	1	0	0
AT4G26210	mitochondrial ATP synthase g subunit family protein	2	0	0
AT3G05690	UNE8, ATHAP2B, HAP2B, NF-YA2 (NUCLEAR FACTOR Y, SUBUNIT A2) transcription factor	1	0	1
AT3G15150	zinc ion binding	0	0	1
AT5G38720	unknown protein	0	0	1
AT5G10710	protein binding	1	0	0
AT4G10340	LHCB5 (LIGHT HARVESTING COMPLEX OF PHOTOSYSTEM II 5) chlorophyll binding	0	0	1
AT2G14520	CBS domain-containing protein	1	0	0
AT1G06190	ATP binding / ATPase, coupled to transmembrane movement of ions, phosphorylative mechanism	0	0	1
AT1G71340	glycerophosphoryl diester phosphodiesterase family protein	0	0	1
AT4G15545	unknown protein	1	0	0
AT1G75300	isoflavone reductase, putative	0	0	1
AT3G09030	potassium channel tetramerisation domain-containing protein	1	0	0
AT4G13670	PTAC5 (PLASTID TRANSCRIPTIONALLY ACTIVE5) heat shock protein binding / unfolded protein binding	0	0	1
AT4G17270	Mo25 family protein	2	0	0
AT5G05640	nucleoprotein-related	0	0	1
AT3G10530	transducin family protein / WD-40 repeat family protein	0	0	1
AT3G52340	SPP2, ATSPP2 | SPP2 (SUCROSE-6F-PHOSPHATE PHOSPHOHYDROLASE 2) catalytic/ magnesium ion binding / phosphatase/ sucrose-phosphatase	1	0	0
AT3G14660	CYP72A13 electron carrier/ heme binding / iron ion binding / monooxygenase/ oxygen binding	1	0	0

Accession	Description	Number of Peptides		
		208 exp.1	208 exp.2	208 exp.3
AT3G48780	ATSPT1, SPT1 | SPT1 (SERINE PALMITOYLTRANSFERASE 1) serine C-palmitoyltransferase	1	0	0
AT5G18820	EMB3007 (embryo defective 3007) ATP binding / protein binding	1	0	1
AT3G07850	exopolygalacturonase / galacturan 1,4-alpha-galacturonidase / pectinase	0	0	1
AT2G42960	protein kinase family protein	1	0	1
AT1G72160	SEC14 cytosolic factor family protein / phosphoglyceride transfer family protein	0	0	1
AT5G05520	outer membrane OMP85 family protein	0	0	1
AT5G14760	AO (L-ASPARTATE OXIDASE) L-aspartate oxidase/ electron carrier/ oxidoreductase	1	0	1
AT3G04820	pseudouridine synthase	1	0	0
AT1G26830	CUL3A, ATCUL3A, ATCUL3, CUL3 | ATCUL3 (ARABIDOPSIS THALIANA CULLIN 3) protein binding / ubiquitin-protein ligase	0	0	1
AT1G68020	ATTPS6 alpha,alpha-trehalose-phosphate synthase (UDP-forming)/ transferase, transferring glycosyl groups / trehalose-phosphatase	1	0	0
AT3G16890	PPR40 | PPR40 (PENTATRICOPEPTIDE (PPR) DOMAIN PROTEIN 40)	1	0	0
AT3G18900	molecular function unknown INVOLVED IN: biological_process unknown LOCATED IN: cellular_component unknown CONTAINS InterPro DOMAIN/s: F-box associated (InterPro:IPR006527), F-box associated type 1 (InterPro:IPR01	1	0	0
AT3G14120	molecular function unknown INVOLVED IN: transport LOCATED IN: nuclear pore EXPRESSED IN: 23 plant structures EXPRESSED DURING: 13 growth stages CONTAINS InterPro DOMAIN/s: Nuclear pore protein 84/107	0	0	1
AT1G53440	leucine-rich repeat family protein / protein kinase family protein	1	0	0
AT5G13590	unknown protein	1	0	1
AT1G68750	ATPPC4 phosphoenolpyruvate carboxylase	1	0	0

Accession	Description	Number of Peptides		
		208 exp.1	208 exp.2	208 exp.3
AT1G17580	MYA1, ATMYA1, XI-1 | MYA1 (MYOSIN 1) motor/ protein binding	0	0	1
AT1G01350	nucleic acid binding / protein binding / zinc ion binding	0	0	1
AT1G05810	ARA, ARA-1, ATRAB11D, ATRABA5E, RABA5E | RABA5E (RAB GTPASE HOMOLOG A5E) GTP binding	0	0	1
AT1G23640	unknown protein	0	0	1
AT1G31390	meprin and TRAF homology domain-containing protein / MATH domain-containing protein	0	0	1
AT1G44094	unknown protein	0	0	1
AT1G60815	RALFL7 (RALF-LIKE 7) signal transducer	0	0	1
AT2G33240	XID, ATXID motor/ protein binding	1	0	1
AT2G35170	MORN (Membrane Occupation and Recognition Nexus) repeat-containing protein /phosphatidylinositol-4-phosphate 5-kinase-related	0	0	1
AT2G42820	HVA22F (HVA22-LIKE PROTEIN F)	1	0	0
AT3G15354	SPA3 (SPA1-RELATED 3) protein binding / signal transducer	1	0	0
AT3G19630	radical SAM domain-containing protein	0	0	1
AT3G25170	RALFL26 (ralf-like 26) signal transducer	1	0	0
AT4G27610	unknown protein	0	0	1
AT4G28580	magnesium transporter CorA-like family protein (MRS2-6)	0	0	1
AT5G49850	jacalin lectin family protein	0	0	1
AT5G60880	unknown protein	1	0	0

Known system IV components are highlighted in red

Supplemental Table 2:

Primer name	Sequence (5´>3´)	Used in
CCB1-*Sfi*I-H	GATCGAGGGCCATTACGGCC**ATG**GCCATGATTGTGGAGCCAC	Yeast Split Ubiquitin analysis
CCB1-*Sfi*I-R2	GATCACGGGCCGAGGCGGCC*CT*TGATGATCTCTCAAAGAG	
CCB2-*Sfi*I-H	GATCGAGGGCCATTACGGCC**ATG**GCCCAAAGGTCAACACGAATC	
CCB2-*Sfi*I-R2	GATCACGGGCCGAGGCGGCC*CC*ACCTCTGAATTTCTCAGC	
CCB3-*Sfi*I-H	GATCGAGGGCCATTACGGCC**ATG**TCAGCAACATTAAGCCAC	
CCB3-*Sfi*I-R2	GATCACGGGCCGAGGCGGCC*CC*ATTAACTTGTTGTTGAGAAAC	
CCB4-*Sfi*I-H	GATCGAGGGCCATTACGGCC**ATG**GCTTCTTCTTCCTCAACCTC	
CCB4-*Sfi*I-R3	GATCACGGGCCGAGGCGGCC*CC*AACTGTTACTCTGGAAATC	
petB*Sfi*I-H	GATCGAGGGCCATTACGGCC**ATG**AGTAAAGTTTATGATTG	
petB*Sfi*I-R2	GATCACGGGCCGAGGCGGCC*CC*TAAGGGACCAGAAATACCTTG	
pBT3-C-H	TGGCATGCATGTGCTCTG	
pBT3-N-H	CAGAAGGAGTCCACCTTAC	
pBT3-C-R	GTAAGGTGGACTCCTTCT	
pBT3-N-R	AAGCGTGACATAACTAATTAC	
ccb1-f	TGACAATGGCGACGAAGCTGATTTCTC	Genotyping
ccb1-r	TGTAATTAGTAAAGTTCGAACC	
ccb2-f2	ATTGTGCAGTCTCAACTTTGACAG	
ccb2-r2	CACTAACAACGTCGAACAGTATCC	
ccb4-f1	ACTTCCTTCATCTTCAATGGAAGC	
ccb4-r1	CAGAGTCCGGTTTGATCAACTTAC	
LB GK045H07 gen-H	GCGTGGACCGCTTGCTGCAACT CATAGAAGAAAGAAACAGAGC	
GK045H07 gen-R	TTGGGAATTGGAAGTGAAATTTAG	
TDNA-LB-H	ATAATAACGCTGCGGACATACATTTT	
208geno*Bcl*I-H	GAATCTTATCTACCGAGATTGATC	
208geno-R	CTCTATCACCAAAAAATTCAC	

Supplemental data 1

Production of antibodies against the AtCCB proteins

Selection of Peptides

AtCCB1

1 MATKLISPPLSCPWVTSREVIIKGLPRRRREWMVTKRNRVSAVTAMIVEP

51 LSVVSSSAIQIHQVWEQNPNSLLLMTEATGGYS**LASYYTSLGLFVISVPG**

101 **LWSLI**KRSVKSKIVRKTFVVNDVKKEPKQVAGEILSFFTRKNFNITDRGE

151 TITFEGK**MVPSRGQAALLTFCTCISLAS**VGLVL**TITVPDFGNNWFFIILL**

201 **SPL**AGVYYWKKASRKEEIKVKMMVGSKGRLDEIVVQGDDVQVEEMRKELQ

251 LNEKGMVYVKGLFERSS

HCF208/ AtCCB2

1 MSIQICNFPFHPKFALQPRAQRSTRIFARTENDSPQSKTSDQQLNLSVLR

51 FTFGIPGFDESYLPR**WIGYGFGSLLLLNHFSASAP**ISESQMR**SEALGLSL**

101 **AAFSIALPYIG**KFLKGSVVEQRSLPEEGEQVFVISSNIGDSLKEDLAWAT

151 YVLLRNTSTIAVLISVQGELCVRGYWNCPDQMSKAQLHDWFKKKVDEIGL

201 ADVKETLYFPQYAGSALSLDILPDGTRSLFVQPLVQNTNEPQKVNGFLLV

251 ASTAGYAYSDKDRAWIGAMAEKFRG

AtCCB3

1 MTTVTTSFVSFSPALMIFQKKSRRSSPNFRNRSTSLPIVSATLSHIEEAA

51 TTTNLIRQTNSISESLRNISLADLDP**GTAKLAIGILGPALSAFGFLFILR**

101 **IVMSW**YPKLPVDKFPYVLAYAPTEPILVQTRK**VIPPLAGVDVTPVVWFGL**

151 **VSFLSEIL**VGPQGLLVLVSQQQVN

AtCCB4

1 MEARIILLRIQIPWSANRQFSHPPLDFPRFIRASSSSTSQKPKTYEGPKP

51 RKNLVADFISKNDDLVRS**LPIYVGGASLLAVLFNRTVSGIAPV**ADASSSQ

101 SR**ADLLALGLAVTNLLTGLVWLSI**RPKSITPVNPKGVECKVVESDLPASM

151 VSELLWAWESLKVATCCKSLVIVYNGICLIQIGMVAESPEDKKTVIVKTD

201 KLMQGSVYRGVMKSKAQSYLANLSLYPGRSELPFLPANTQAVILQPLGDK

251 GIAVIGGNTIRGFTSSDQAWISSIGEKLDATLGRYFVDSDEISRVTV

cTP (WoLF PSORT)

cTP (chloroP)

peptide for antibody production

bold letters transmembrane domain (TMAP)

Figure S1: Protein sequences, predicted domains and selected peptides for antibody production.
Depending on the prediction algorithm, WoLF PSORT or chloroP, two different cTPs are displayed for AtCCB1 and HCF208/ AtCCB2 (black bar and sequences highlighted in gray).

Two peptide sequences were selected for AtCCB1 to AtCCB4 protein according to the following criteria: (i) location in the mature protein (taking account of cTPs), (ii) no overlapping with predicted transmembrane domains (for eventual applications on native proteins), and (iii) prediction of being an applicable antigen. The latter point was analyzed by the company Agrisera AB (Vännäs, Sweden), which was selected for antibody production. The second criterion could not be fulfilled for the AtCCB3 protein since it is very small and mainly consists of transmembrane domains. Localization of selected peptides, predicted transmembrane domains and cTPs are indicated in Supplemental Figure S1.

Characterization of anitsera

A mix of the two peptides selected for each protein was used for the immunization of two rabbits per protein after preimmunsera were tested for unspecific signals. Sera from final bleeds were tested using membrane proteins isolated from *A. thaliana* Col-0 wild type, *hcf208* mutants, HCF208pAUL2, and the T-DNA insertion mutants atccb1_SALK, atccb2_SALK, atccb4_SALK (Lezhneva et al., 2008) and atccb1_GK (Figure 1).

Using sera from rabbits immunized with AtCCB1-peptides a strong signal at a molecular weight of ~20 kDa (Figure S2 A) was detected, which corresponds to the predicted molecular weight of mature AtCCB1 (22.8 kDa). The ~20 kDa signal was absent in *atccb1*_SALK mutants, whereas in *atccb1*_GK mutants a weak signal was detected by the serum of rabbit 2. These results indicate that the ~20 kDa signal is specific to AtCCB1.

Analysis of sera from rabbits immunized with AtCCB3-peptides is shown in Figure S2 B. Since no AtCCB3 mutant lines are available, sera were tested using *Arabidopsis* Col-0 wild type protein extracts only. No band corresponding to the predicted molecular weight of mature AtCCB3 protein (14.9 kDa) was detected with both sera. The signal detected at ~35 kDa is peroxidase activity of cytochrome *f*. Variation of dilutions of sera did not improve the signal and eventually led to even stronger background signals.

The predicted molecular weight of mature AtCCB4 is 29.2 kDa. Testing sera from rabbits immunized with AtCCB4 peptides revealed the presence of a band with appropriate molecular weight detected with the serum from rabbit 1 but not rabbit 2 (Figure S2 C). This band was not detected in the *atccb4*_SALK mutant indicating that the signal is specific to AtCCB4. However, a strong unspecific signal was present below the AtCCB4 signal making it necessary to perform SDS-PAGE using high polyacrylamide concentrations.

Immunoblot analysis using sera from rabbits immunized with HCF208/ AtCCB2 peptides exhibited no band corresponding to the predicted molecular weight of the mature HCF208/ AtCCB2 protein (~27 kDa) in wild type. However, when tested with proteins from the overexpressing line HCF208pAUL2, a slight band could be detected with the serum from

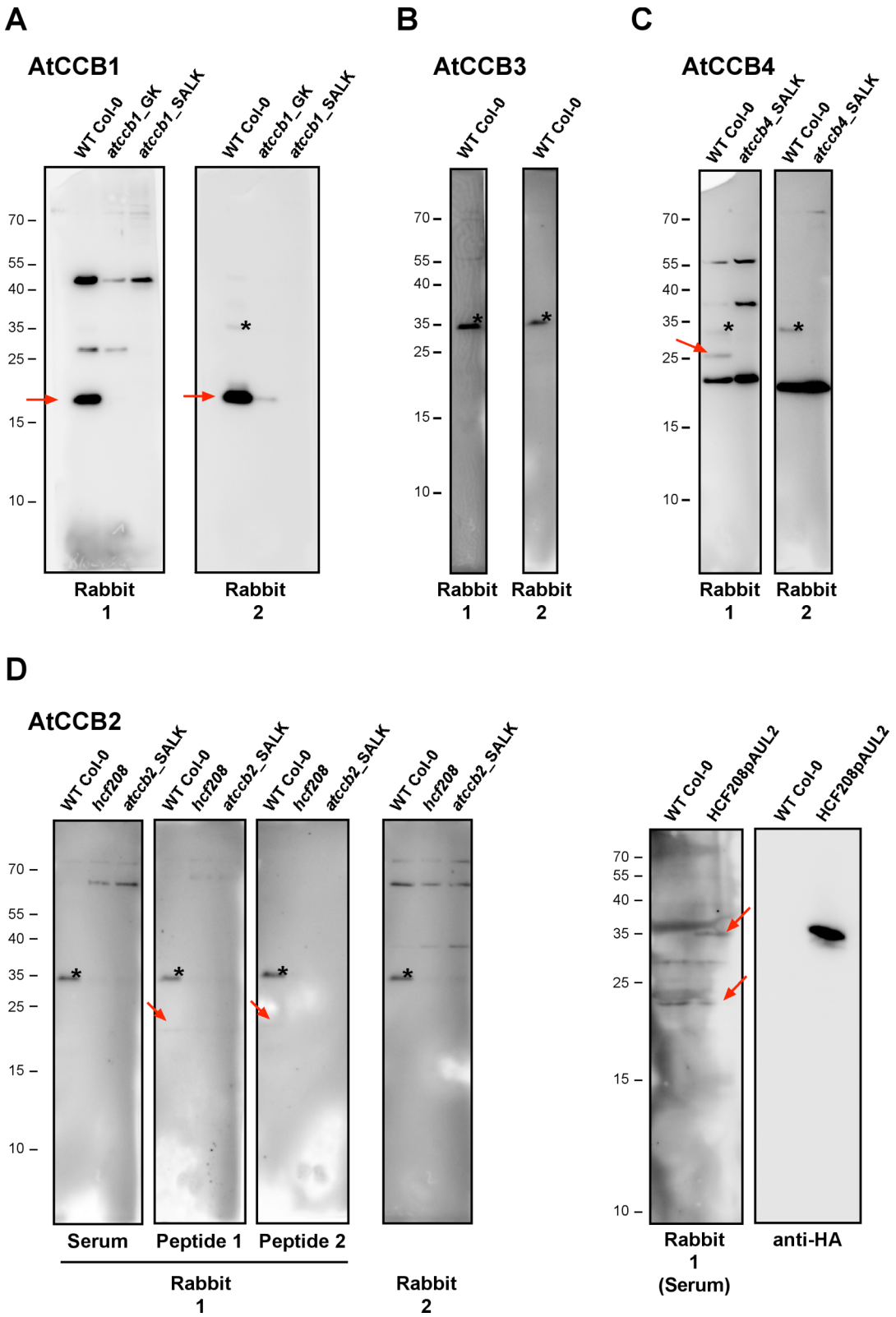


Figure S2: Characterization of sera after immunization.
Membrane proteins were isolated from wild-type (WT) Col-0, *atccb1*_GK, *atccb1*_SALK, *hcf208*, *atccb2*_SALK, *atccb4*_SALK and HCF208pAUL2 plants. 20μg of membrane proteins were analyzed by SDS-PAGE and Western blot using respective sera. Cytochrome *f* peroxidase activity is indicated by an asterisk. Signals corresponding to CCB proteins are indicated by red arrows. Characterization of (A) Anti-AtCCB1, (B) Anti-AtCCB3, and (C) Anti-CCB4 sera. Sera were diluted 1:2000 in TBS-T+5% milk powder. (D) Characterization of Anti-HCF208/ AtCCB2 sera. Sera were diluted 1:1000. Purified sera from rabbit 1 via peptide 1 or peptide 2 were diluted 1:400. Gels for testing of membrane proteins from WT and HCF208pAUL2 plants were loaded with 30μg.

rabbit 1 corresponding to the signal obtained from detection of the HCF208/ AtCCB2 fusion protein by the Anti-HA-Peroxidase antibody (Figure S2 D, right side). This signal was absent in the wild-type sample. Therefore, the serum from rabbit 1 was purified via peptide 1 (QMSKAQLHDWFKKK) and peptide 2 (STAGYAYSDKDRAW) respectively by Agrisera AB (Vännäs, Sweden). Purified antibodies were tested under the same conditions as the sera. For both antibodies (Peptide 1 and Peptide 2) very weak signals were detected at a molecular weight of ~25 kDa, which were not observed in *hcf208* and *atccb2*_SALK and therefore most likely representing the endogenous HCF208/ AtCCB2 protein (Figure S2 D).
Taken together, strongly reacting antibodies for AtCCB1 could be generated in this study. Also, antibodies against AtCCB4 were obtained, although they exhibit background signals. The HCF208/ AtCCB2 antibody reacted only weakly and no AtCCB3 antibody could be generated.

Supplemental Figures

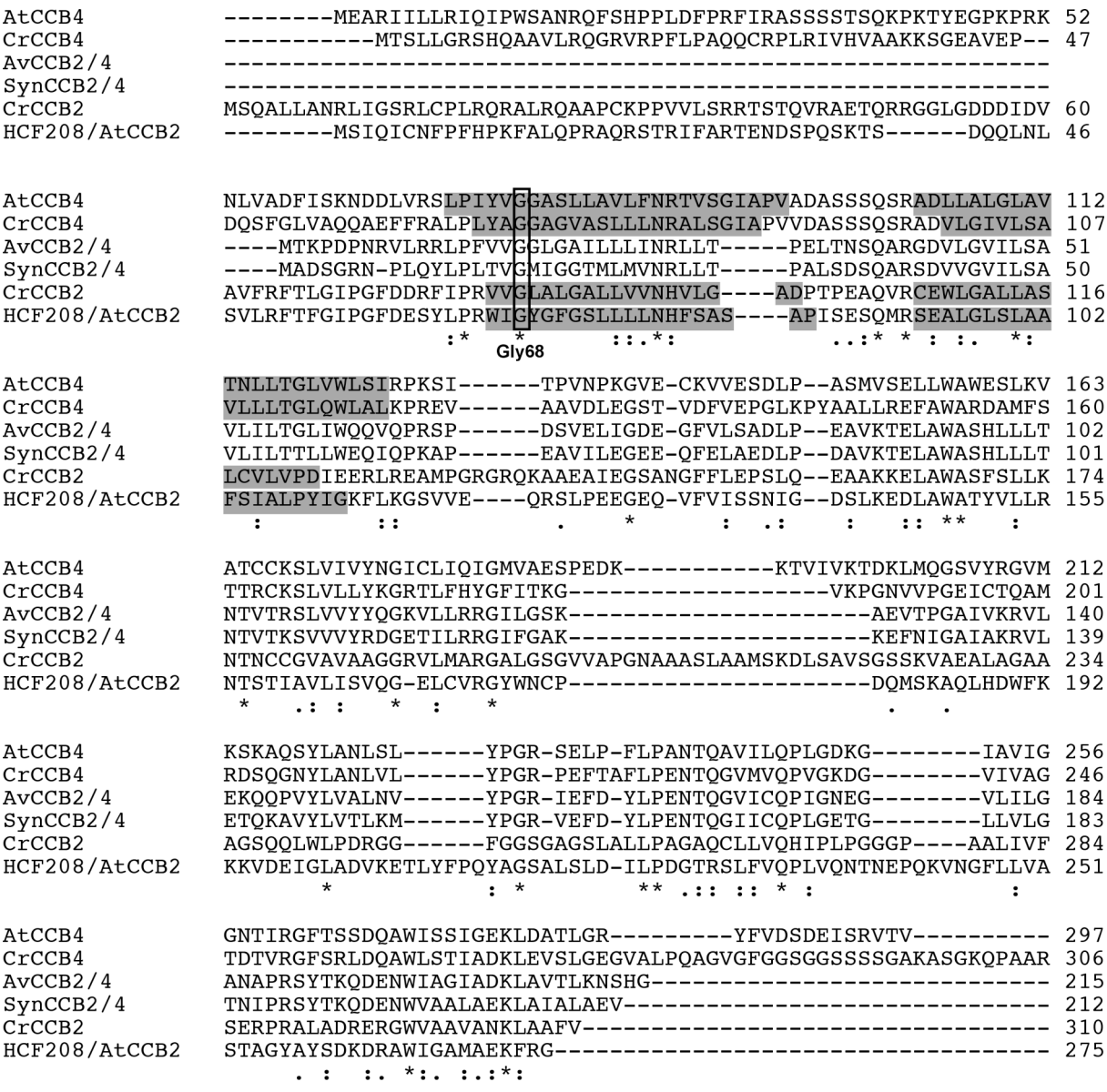


Figure S3: Sequence alignment of CCB2 and CCB4 proteins
Alignment of CCB2 and CCB4 protein sequences from *C. reinhardtii* (CrCCB2, CrCCB4) and *A. thaliana* (HCF208/ AtCCB2, AtCCB4) and of the CCB2/4 protein from the cyanobacteria *Anabena variabilis* and *Synechococcus sp.* (AvCCB2/4, SynCCB2/4). Transmembrane domains are indicated in gray. The glycine residue exchanged in *hcf208* is enframed.

- Part C -

Manuscript 4

Analyses on protein function and complex formation of the PsbH synthesis factor HCF107

Introduction

In the course of evolution from a cyanobacterium-like cell to the chloroplast that is fully integrated into the regulatory network of its eukaryotic host cell the prokaryotic genome has been massively rearranged. Major parts of the ancestral genome of the endosymbiont were lost or transferred to the nucleus but also new factors derived from the host genome. A large number of nucleus-encoded proteins are involved in the regulation of chloroplast gene expression and the nucleus provides structural components of the photosynthetic complexes, photosystems I and II (PS I and II) and the cytochrome b_6f complex (Herrmann et al., 1985; Wollman et al., 1999). Therefore, to ensure proper biogenesis and maintenance of the photosynthetic complexes, coordination of nuclear and chloroplast gene expression in response to developmental and environmental changes is of great importance.

Generally, plastid gene expression is regulated by post-transcriptional mechanisms (Deng and Griussem, 1987). According to their cyanobacterial origin most chloroplast genes are arranged in polycistronic gene clusters, which are transcribed from a single promoter (Herrmann, 1992; Sugiura, 1992). Those transcripts often need to be processed into monocistronic forms via cleavage by endonucleases to allow efficient translation (Barkan et al., 1994; Sturm et al, 1994; Hirose et al, 1997). Further, 5´ and 3´ end trimming of transcripts and binding of specific proteins play important roles in transcript stability and translation. Biochemical and genetic studies in higher plants (*Arabidopsis thaliana*, maize) and the green alga *Chlamydomonas reinhardtii* helped to identify factors involved in these processes. Besides general factors like exo- and enoribonucleases or ribosomal proteins, the nucleus provides a variety of factors that are specific to one or several transcripts (Marin-Navarro et al., 2007; Stern et al., 2010). Many of these factors are members of the TPR- (TetratricoPeptide Repeat) and PPR- (PentatricoPeptide Repeat) protein families, which belong to the helical repeat protein superfamily (Schmitz-Linneweber and Small, 2008; Stern et al., 2010). TPR and PPR proteins are defined by degenerate 34 or 35 amino acid repeats, respectively, forming two antiparallel helices (Helices A and B; Blatch and Lässle, 1999; Small and Peeters, 2000). TPR-tandem arrays are expected to form a right-handed superhelix enclosing a groove mainly formed by helix A, which crates surfaces for protein-protein interactions (Das et al., 1998; Han et al., 2007). PPR proteins are supposed to form similar structures that are rather involved in protein-RNA interactions (Small and Peeters, 2000). However, TPR proteins that are linked to RNA-associated processes exhibit specific sequence properties of TPR-motifs differing from the classical consensus and are therefore referred to as RNA-TPR proteins (RTPR; Ben-Yehuda et al., 2000; Sane et al., 2005).

In the last years an ever-increasing number of helical repeat proteins have been identified to play roles in transcript stabilization and translation, which are tightly coupled processes

(Schmitz-Linneweber and Small, 2008; Stern et al., 2010). The PPR10 protein from maize, which binds to the 5´ region of *atpH* transcripts and protects them from 5´→ 3´ degradation thereby defining their 5´ ends (Pfalz et al., 2009), was found to activate translation at the same time by remodeling of the RNA structure (Prikryl et al., 2010). Another way of translation activation was described for the PPR protein MCA1 from *Chlamydomonas,* which stabilizes the *petA* transcript (encoding for cytochrome *f*) by binding to its 5´ end and recruits another protein that acts as a translational factor, TCA1 (Raynaud et al., 2007; Boulouis et al., 2011). A role in stabilization and translation of the *psbD* and *psbB* transcripts in *Chlamydomonas* was prescribed for the RTPR proteins NAC2 and MBB1, respectively (Nickelsen et al., 1999; Boudreau et al., 2000; Vaistij et al., 2000b).

In former studies the *Arabidopsis* PSII mutant *hcf107*, meanwhile represented by seven mutant alleles, was identified and shown to be incapable of synthesis of the plastome-encoded PSII subunits PsbH and PsbB/CP47 (Meurer et al., 1996; Felder et al., 2001). Both proteins are encoded by the same polycistronic transcription unit, the *psbB*-operon. The mutants do not accumulate transcripts with *psbH* as the leading cistron, whereas the *psbB* transcript encoding for CP47 is stable (Felder et al., 2001). Recent studies point to a secondary effect of the *hcf107* mutation on rapid CP47 turnover rather than lack of synthesis since translation of *psbB* is unaffected and the *hcf107* mutant phenotype can be complemented by a nuclear-encoded, chloroplast-targeted version of the PsbH protein (Hinsen, 2008; Hermanns, 2010). Thus, the primary function of HCF107 is supposed to be processing and/ or stabilization of transcripts with *psbH* as the leading cistron an their translation.

The HCF107 protein is a RTPR family protein possessing 11 tandemly arranged RTPR motifs. It was found to peripherally associate with chloroplast membranes with a small soluble fraction when using tagged versions of the protein (Sane et al., 2005). Membrane-associated HCF107 forms a high molecular weight complex of 600-800 kDa (Sane et al., 2005).

In the present study, extended analyses on the endogenous HCF107 protein are described. Analyses on the seven mutant alleles disclose essential amino acids of the RTPR protein HCF107. Further, the data shows that the function of of the HCF107 complex is stabilization and on-membrane translation of the *psbH* transcript and that these processes are strongly regulated by light.

Results

Analyses of HCF107 protein levels in allelic mutants

During the past years seven alleic mutants of HCF107, *hcf107.1* to *hcf107.7*, have been identified and characterized in our laboratory.

The *hcf107.1* and *hcf107.3-.7* mutant alleles carry point mutations induced by ethylmethansulfonate (EMS; Figure 1 A). *Hcf107.2* was obtained from a T-DNA insertion line, which has lost its T-DNA leading to a replacement of a 17 nt sequence by a 9 nt repeat sequence at the former insertion site (exon 3-intron 3 border; Figure 1 A) and therefore to mis-splicing of the transcript (Sane et al., 2005). The read-through of exon 3 to intron 3 is predicted to result in premature termination of the polypeptide (Figure 1 B; Sane et al., 2005). The mutations in *hcf107.1* and *hcf107.4* cause amino acid exchanges at positions 244 (Ala → Thr) and 451 (Glu → Lys), respectively (Sane et al., 2005; Lyska, 2006). Both mutations are located in RTPR motifs and affect highly conserved residues (Ala8 of RTPR 3 and Glu11 of RTPR 9; Figure 1 B) located in the first of the two antiparallel helices formed by each RTPR motif (helix A). In *hcf107.3*, *hcf107.6* and *hcf107.7* nucleotide exchanges in exons 1, 6 and 7, respectively, result in the generation of stop-codons and therefore presumably expression of truncated proteins (Figure 1 B). In *hcf107.5* the last nucleotide of intron 5 is exchanged leading to mis-splicing, which would result in synthesis of a truncated protein due to a premature stop-codon in intron 5 (Figure 1 B).

To test the predictions on accumulation of HCF107 in the seven mutants, total proteins of all mutant alleles were analyzed by Western blot using an HCF107-specific antibody. As indicated in Figure 1 C in both, *hcf107.1* and *hcf107.4*, full-length proteins were detected according to their mutations displayed by amino acid exchanges. However, the amount of the *hcf107.4* protein was drastically reduced when compared to wild type, whereas *hcf107.1* accumulated to wild-type levels, indicating an effect of the Glu451 → Lys exchange in *hcf107.4* on protein stability. In *hcf107.3* no protein could be detected (Figure 1 C) according to the presence of a premature termination signal after 29 amino acids, which would result in a protein with a molecular weight of ~3 kDa (Figure 1 B). In *hcf107.2*, *hcf107.5*, *hcf107.6* and *hcf107.7* truncated proteins were detected at different molecular weights that are in agreement with the predictions (*hcf107.2*: 351 AA, ~39 kDa; *hcf107.5*: 457 AA, ~50 kDa; *hcf107.6*: 509 AA, ~56 kDa; *hcf107.7*: 451 AA, ~50 kDA; Figures 1 B and C). However, all truncated proteins were reduced significantly indicating either degradation at the mRNA level by mechanisms like nonsense-mediated decay or at the protein level due to misfolding by the proteasome (Cartegni et al., 2002; Goldberg, 2003).

Data presented here confirmed the predicted protein models for each of the seven *hcf107* mutant alleles. Since all mutants do not accumulate transcripts with *psbH* as the leading cistron (Meurer et al., 1996; Felder et al., 2001; Lyska, 2006; Hermanns, 2010), the reduction and/ or the truncation of the HCF107 protein but also the point mutations in the RTPR motifs result in complete loss of protein function.

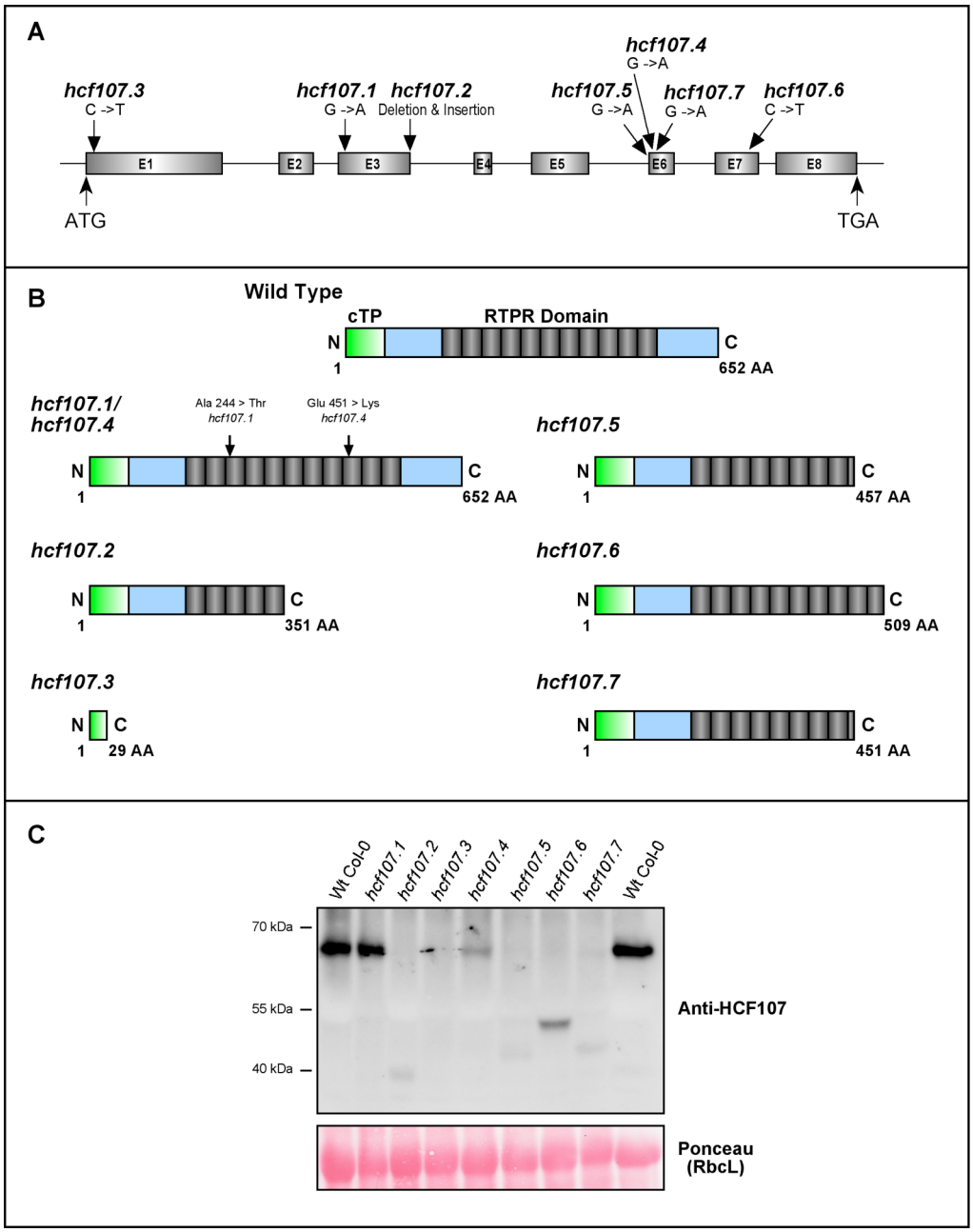


Figure 1: Comparison of *hcf107* mutant alleles.
(A) Schematic illustration of the HCF107 gene (At3g17040). Exons are displayed by gray boxes (E1-E8). The positions of the seven allelic mutations are indicated. (B) Schematic representation of protein structures of HCF107 and proteins encoded by the mutant alleles *hcf107.1* to *hcf107.7*. Amino acid exchanges are indicated by arrows. Predicted sizes of truncated proteins are provided. AA = amino acids. (C) Analysis of proteins from mutant alleles *hcf107.1* to *hcf107.7* and wild type (Col-0). Total proteins were isolated from two-week-old seedlings and analyzed by SDS-PAGE and Western blot using an Anti-HCF107 antibody. Ponceau staining of the membrane after transfer served as a loading control.

HCF107 protein accumulation is independent of light in developed leaves

Sane et al. (2005 and unpublished) showed that the HCF107 transcript is absent in the dark, but accumulates in light-grown seedlings. To test whether HCF107 expression is regulated by light in later stages of plant development after the differentiation of chloroplasts, the accumulation of the HCF107-protein was determined over 24 hours in seedlings grown under long day conditions (8 h darkness/ 16 h light). As shown in Figure 2 A the amount of HCF107 did not change drastically in the course of the day. HCF107 protein levels were slightly increased in the beginning of the day after 0.5 hours of illumination, but decreased again

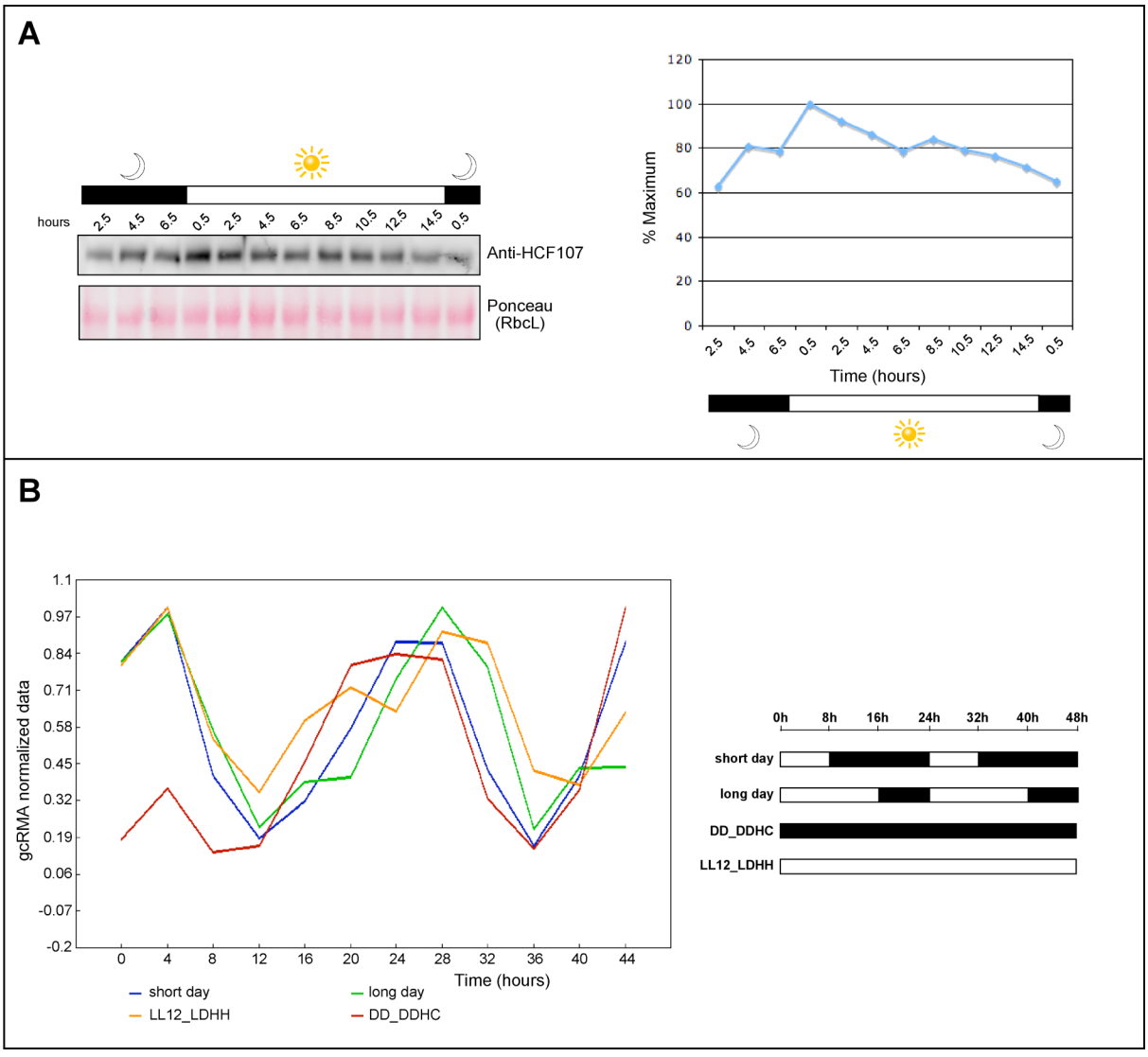


Figure 2: Regulation of HCF107 expression.
(A) Analyses of HCF107 protein accumulation in the course of the day. Plants (wild type Col-0) were grown for one week on 1x MS medium under long day conditions and harvested every two hours for 24 hours. Total proteins were analyzed by SDS-PAGE and Western blot using an Anti-HCF107 antibody. Signals were quantified and intensities are displayed as percentage from the maximal signal intensity. (B) Transcript accumulation under long and short day conditions and in continuous light and darkness. Data was obtained from the DIURNAL search tool (http://diurnal.cgrb.oregonstate.edu; Mockler et al., 2007; Michael et al., 2008).

later in the day and reached their minimum in the 0.5 and 2.5 hours darkness samples (Figure 2 A). However, after 4.5 and 6.5 hours of darkness HCF107 levels slightly increased (Figure 2 A) arguing against a light-regulated expression. Analyses using the DIURNAL search tool (Mockler et al., 2007; Michael et al., 2008), which provides experimental data on transcript accumulation under different light conditions, supported this finding (Figure 2 B). Four different conditions were chosen from the database: (i) long day conditions corresponding to the situation displayed in Figure 2 A, (ii) short day conditions (16 h darkness/ 8 h light), (iii) continuous light and (iv) continuous darkness. As shown in Figure 2 B, HCF107 transcripts de- and increased in an oscillatory manner in a cycle of ~24 hours and independent of a dark-light cycle. Under long day conditions transcript levels already increased before light was switched on, which stands in accordance with the accumulation of HCF107 protein levels (Figure 2 A). Therefore, accumulation of HCF107 during the course of day is light-independent and underlies an endogenously generated self-sustaining circadian rhythm.

Membrane-bound HCF107 forms a high molecular weight complex associated with RNA

Previous experiments concerning the HCF107 protein were performed using a tagged version of the HCF107 protein expressed under the control of a 35S *CaMV* promoter. The majority of this protein was found to be peripherally associated with chloroplast membranes with a small fraction in the stroma (Sane et al., 2005). Sucrose density centrifugation experiments using digitonin-solubilized membrane proteins revealed that the protein forms a high molecular weight complex of 600-800 kDa (Sane et al., 2005). In the present study, the formation of this high molecular weight complex by the endogenous HCF107 protein was confirmed under the same experimental conditions using wild-type plants and an HCF107-specific antibody (Supplemental Figure S1).

However, since solubilization with digitonin was not effective under the chosen buffer conditions, experiments were subsequently performed using n-dodecyl-β-D-maltoside (DM), which increased the yield of solubilized material. Interestingly, when using DM-solubilized samples, the high molecular weight complex migrated to fractions corresponding to the highest sucrose concentration (Figure 3 A upper lane) indicating the presence of a very high molecular weight complex way beyond 800 kDa. This complex presumably at least partially dissociated when samples were solubilized with digitonin. Furthermore, when increasing the sucrose concentration to 50%, the HCF107 complex still migrated to the highest sucrose concentration (Figure 3 B) further sustaining the assumption that the complex is of very high molecular weight.

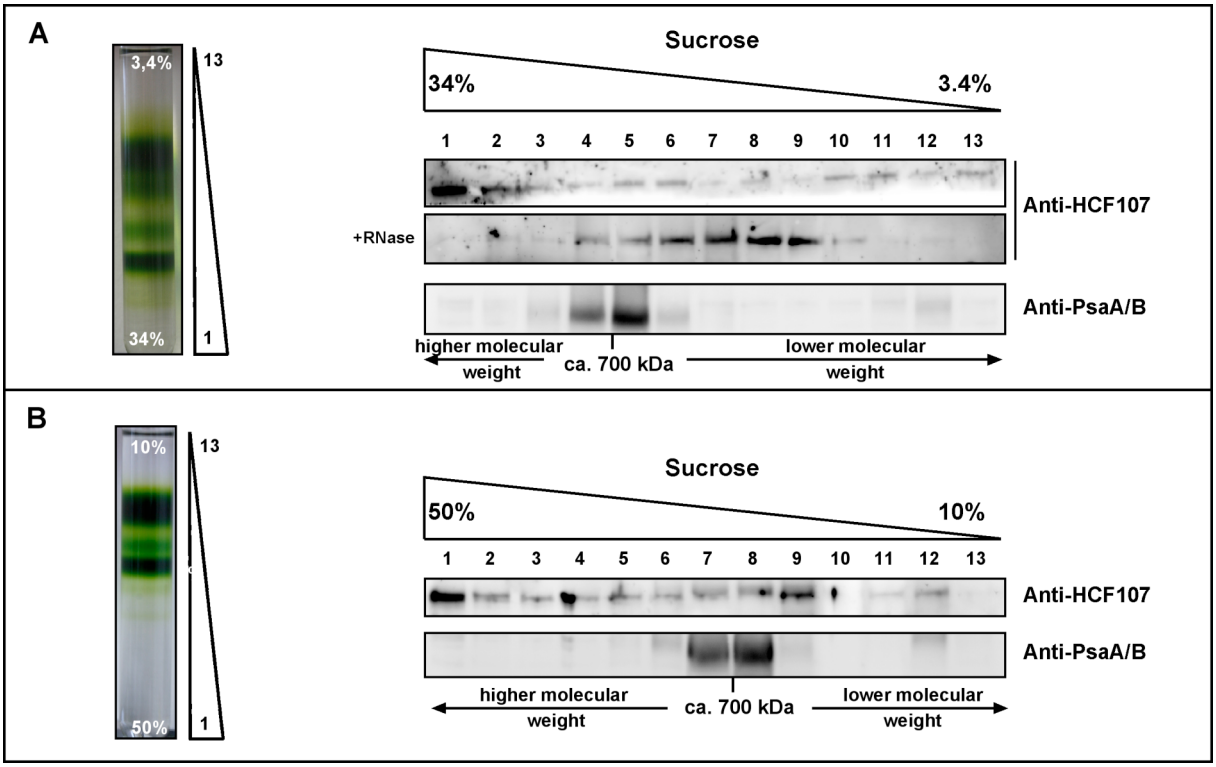


Figure 3: Association of HCF107 with a membrane-bound, RNA-associated high molecular weight complex.
(A) Analyses of HCF107-complexes using 3.4-34% sucrose gradients. Membrane proteins were isolated and solubilized with 1% (w/v) n-dodecyl-β-D-maltoside. After solubilization, samples were incubated with or without RNase A. (B) Analysis of HCF107 complexes using 10-50% sucrose gradients. Membranes were solubilized with 1% (w/v) n-dodecyl-β-D-maltoside.
Gradients were collected in 1 ml fractions and analyzed by SDS-PAGE and Western blot. Membranes were immunodecorated with Anti-HCF107 and Anti-PsaA/B as a control.

Since for HCF107 a function in processing and/or stabilization and translation of *psbH*-transcripts was implicated (Felder et al., 2001), the high molecular weight complex was tested for association with RNA. Solubilized membrane proteins were incubated with RNAse and subjected to sucrose density gradient centrifugation. As shown in Figure 3 A the signal corresponding to the high molecular weight complex was absent from RNase-treated samples. Moreover, smaller complexes of less than 700 kDa (in the range of about 300 kDa) were detected (Figure 3 A). The presented data indicates that the membrane-bound high molecular weight complex formed by HCF107 is associated with, but also adhered by RNA, since it is degraded when the RNA-component is removed. However, it also forms smaller complexes that are stable in the absence of RNA. These complexes should contain the true interactors of HCF107.

Formation of the membrane-bound HCF107-complex is regulated by light

Since it is known that the expression of PSII subunits is regulated by light, its impact on HCF107-complex formation was analyzed.

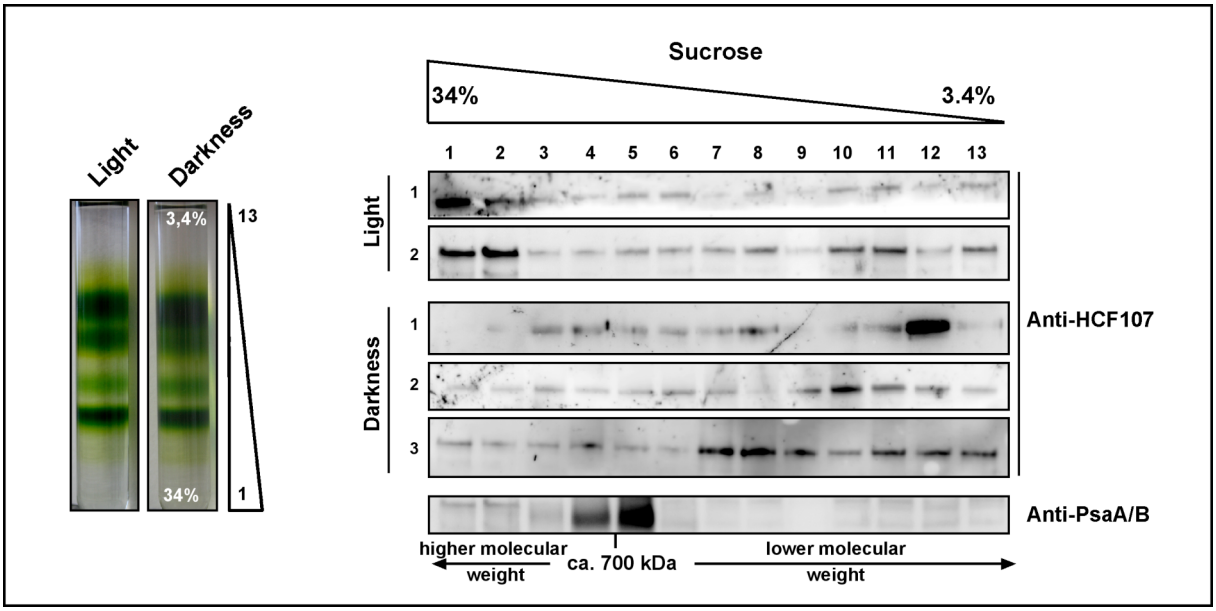


Figure 4: Analyses on light-dependent HCF107 complex formation.
Plant material was harvested at midday. For "Darkness" samples plant material was harvested under a green-light source. Membranes were solubilized with 1% (w/v) n-dodecyl-β-D-maltoside. Complexes were analyzed by 3.4-34% sucrose gradients.
After ultracentrifugation, gradients were collected in 1 ml fractions and analyzed by SDS-PAGE and Western blot. Membranes were immunodecorated with Anti-HCF107 and Anti-PsaA/B as a control.

Solubilized membrane-proteins prepared from plants harvested at midday after four hours of illumination ("Light") or after they were kept in darkness for at least 20 hours ("Darkness") were analyzed by sucrose density gradient centrifugation. The high molecular weight complex migrating to the highest sucrose concentration was only present in light- but not in dark-adapted plants (Figure 4). In fact, the majority of the HCF107 protein was detected in lower molecular weight regions in darkness; depending on the experiment, HCF107 was found in a monomeric form (Figure 4, Darkness, lane 1) or it formed complexes of approximately 100 or 300 kDa (Figure 4, Darkness, lanes 2 and 3, respectively). The presence of those complexes in some samples may be due to minimal light exposure during preparation of membranes, since not all steps could be carried out in complete darkness. Therefore, the monomeric form may display the de facto situation in darkness. The molecular weight of the largest complex formed by HCF107 from dark-adapted plants corresponded to the complex observed after RNase-treatment described above pointing to a similar composition of both complexes. Altogether, this data demonstrates that the association of HCF107 with RNA and with the proteins forming the membrane-bound high molecular weight complex is a light-dependent process.

The fraction of membrane-associated HCF107 protein decreases in light

In experiments described above Western blot signals from membrane fractions of dark-adapted plants appeared to be more intense than signals from light-adapted plants. To confirm this observation, total membrane protein fractions were isolated from light- and dark-adapted plants. As indicated in Figure 5, HCF107 in fact accumulated to lower amounts in membranes of light-adapted plants when compared to dark-adapted plants. Accumulation of ATP-Synthase, which served as a control, was not altered (Figure 5) indicating that the change in HCF107 abundance was not due to a general reduction of protein content of membrane proteins. Since overall levels of the HCF107 protein were not altered by light (Figure 2) and material of light- and dark-adapted plants was harvested simultaneously, the decrease of HCF107 in the membrane-fraction must be due to its dissociation into the soluble fraction. However, further experiments are required for confirmation of this first finding.

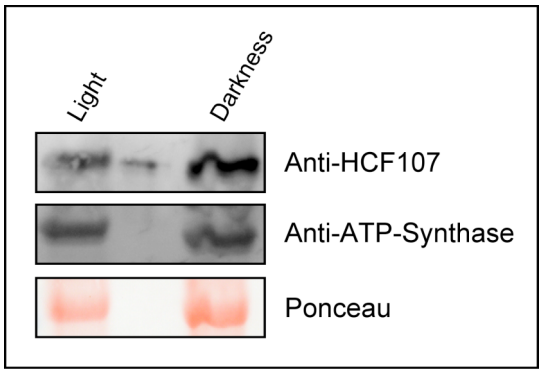


Figure 5: Analyses of HCF107 protein levels in membrane fractions of light- and dark-adapted plants.
Plant material was grown under the same conditions as described for analysis by sucrose density gradients. 20 μg of membrane proteins were resolved by SDS-PAGE and subjected to Western blot analyses. After transfer, nitrocellulose membranes were stained with Ponceau red and then immunodecorated with Anti-HCF107 and Anti-ATP-Synthase as a control.

Stromal HCF107 forms a small complex in light

Since membrane-associated HCF107 formed distinct complexes in light- and dark-adapted plants and the protein was partially released from membranes in light, the stromal fraction of light-adapted plants was analyzed by sucrose density gradient centrifugation. Indeed, the HCF107 protein was detected in the stroma. The majority of the protein migrated in form of a complex of less than 550 kDa (about 300 kDa; Figure 6), whereas only slight bands could be observed in fractions corresponding to higher sucrose concentrations. The molecular weight of the complex again corresponded to the complex observed after RNase-treatment and dark-adaptation, which suggests that the stromal complex the two membrane-bound complexes have a similar composition or are equal. However, these results are preliminary

since this experiment has so far been performed only once. Further experiments confirming this first finding and addressing the composition of the stromal and membrane-associated complex(es) must be performed.

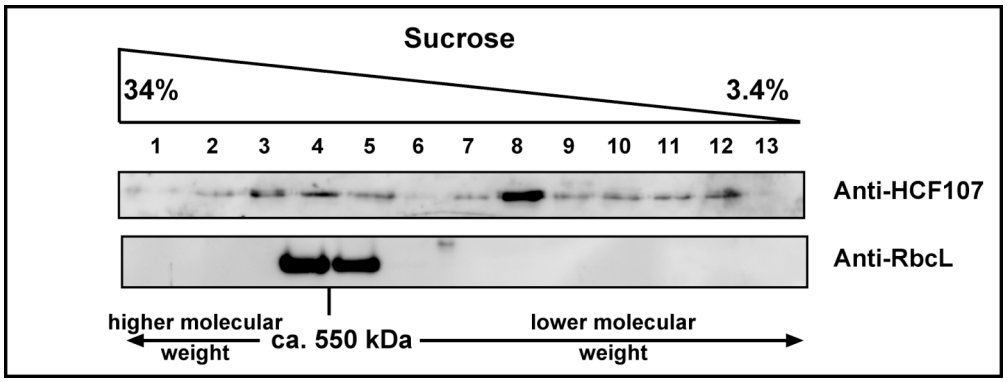


Figure 6: Analysis of stromal HCF107 complexes.
Chloroplasts from light-adapted plants were isolated and lysed. The stromal fraction was loaded onto 3.4-34% sucrose gradients. After ultracentrifugation, gradients were collected in 1 ml fractions and analyzed by SDS-PAGE and Western blot. Membranes were immunodecorated with Anti-HCF107 and Anti- RbcL as a control.

Analyses of sucrose density gradient fractions containing HCF107 by mass spectrometry

Experiments on HCF107 so far did not disclose its molecular function. The predicted RTPR motifs making up about 65% of the HCF107 protein point to both, RNA-association and mediation of protein-protein interactions. Also, no motifs indicating a catalytical function were predicted. Since previous analyses revealed that HCF107 forms complexes the identification of associated proteins could shed light on the precise function of HCF107.

Approaches using affinity purification remained unsuccessful so far. Therefore, light-dependent complex formation was turned to account to get a hint on the composition of the RNA-associated high molecular weight complex formed by HCF107. Solubilized membrane proteins from light- and dark-adapted plants were subjected to sucrose density gradient centrifugation using a maximum sucrose concentration of 50%. Both samples were analyzed by Western blot to confirm the absence of the complex in the first fraction of samples from dark-adapted plants (Figure 7 A). This fraction was not contaminated with highly abundant proteins like those from the photosynthetic complexes since chlorophyll-containing complexes migrated to about half the distance of the gradient (Figure 3 B). Fractions 1 from light- and dark-adapted samples (5 and 4 gradients, respectively) were separated by SDS-PAGE and each lane was analyzed by mass spectrometry (MS). Proteins detected in samples from light-adapted but not from dark-adapted plants should indicate potential candidates for HCF107-associated proteins.

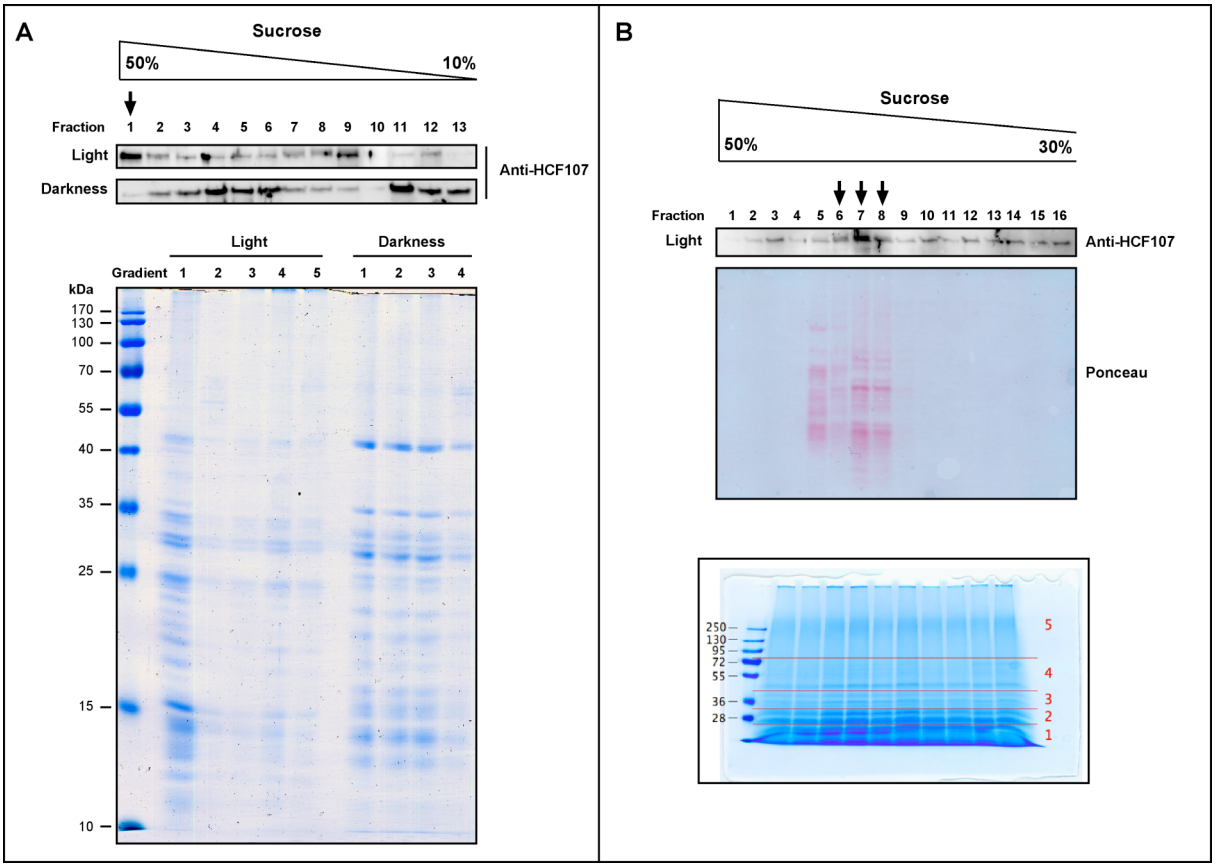


Figure 7: Sample preparation for mass spectrometric analysis of sucrose gradient fractions.
(A) Sample preparation for comparison of fraction 1 from light- and dark-adapted plants. Plant material was harvested at midday. For "Darkness" samples plants were harvested under a green-light source. Membranes were solubilized with 1% (w/v) n-dodecyl-β-D-maltoside and subjected to 10-50% sucrose gradients. Fractions 1 from 5 "light" gradients and 4 "dark" gradients were separated by SDS-PAGE and the gel was stained with Colloidal coomassie. Each lane was separately analyzed by mass spectrometry. (B) Sample preparation for yield increase for mass spectromic analysis of light-adapted samples. Plant material was harvested at midday. Membranes were solubilized with 1% (w/v) n-dodecyl-β-D-maltoside and subjected to 10-50% sucrose gradients. Fractions with the strongest HCF107-signal (7) and surrounding fractions (6 and 8) from five gradients were collected, precipitated and pooled. For mass spectrometry, proteins were separated by SDS-PAGE and the gel was stained with Colloidal coomassie. Red lines and numbers indicate sections according to different molecular weights that were analyzed separately by MS.

Peptides corresponding to overall 334 proteins were recovered by MS. Out of these proteins 83 were absent or clearly reduced in darkness vs. light. Proteins were divided into four groups (listed in Supplemental tables 1 A-D). Group 1 (Supplemental table 1 A) displays ribosomal proteins, which made up about 60% of identified proteins indicating that the high molecular weight fraction contains translational complexes. Group 2 consists of other proteins, which are related to translation or may be associated with RNA (Supplemental table 1 B). Among these proteins were HCF107 and the factors HCF173 and HCF244 that are required for translation of the *psbA* transcript encoding the D1 subunit of PSII (Schult et al., 2007; S. Link, unpublished). Also, the translation initiation factor FUG1 (Miura et al., 2007), translation elongation factors and proteins with RNA-binding motifs were detected. Group 3 represents proteins with known functions but without direct function in translation or RNA-

association (Supplemental table 1 C). Interestingly the ALB3 protein and a SECY homolog, which are supposed to be required for co-translational membrane insertion of photosynthetic proteins, were found among proteins in this group. Finally, group 4 (Supplemental table 1 D) was formed by uncharacterized proteins.

However, in this experiment many proteins were detected by only one peptide and therefore did not provide reliable data. In this regard, HCF107 was only detected by one peptide in one single light-sample. To increase the sensitivity of MS analysis an initial experiment was performed using only samples from light-adapted plants and three times higher amounts of solubilized membrane proteins loaded onto higher volumes of sucrose density gradients. In Western blot analyses of the first 16 fractions obtained from these gradients the bulk of the HCF107 protein was detected in a lower fraction than before (fraction 7; Figure 7 B), which may have resulted from different ultracentrifugation conditions and different diameters of centrifugation tubes necessary for the large-scale experiment. Fractions 6 to 8 of 5 gradients were pooled and separated by SDS-PAGE. To further increase sensitivity, SDS-polyacrylamide gels were divided into five parts according to different molecular weights (Figure 7 B). After MS analysis data from each gel-section was combined. Overall 696 proteins were identified by this experiment. For evaluation, proteins, which were detected by 2 peptides or less and proteins, which were not restricted to chloroplasts were subtracted. Also, the major part of detected proteins was represented by structural components of ribosomes and was not taken into account for further evaluation, since specific factors were to be identified. After all subtractions, 157 proteins remained. These proteins were placed into groups representing proteins with functions in RNA-associated processes and translation (group 1, Supplemental table 2 A) photosynthetic electron transport (including ATP-Synthase; group 2, Supplemental table 2 B), metabolism and transport (group 3, Supplemental table 2 C), protein folding, -transport and -degradation (group 4, Supplemental table 2 D), and others (group 5, supplemental table 2 E). HCF107 this time was identified with nine peptides and also HCF173 and HCF244 were present in these samples. Also, translation initiation, elongation and termination factors as well as putative RNA binding proteins (At3g52150, At1g01080, At2g37230, At5g42310 and At3g23700; Supplemental table 2 A**)** were identified. At3g23700 previously was shown to co-purify with HCF173 (Schult, 2005). Further, three proteins with predicted ribonuclease activity were detected (At4g37510, At5g61090, At5g02250; Supplemental table 2 A).

An interesting finding was the presence of a large number of proteins associated with PSII assembly and repair as well as thlyakoid membrane import proteins in the high molecular weight fractions (e.g. STN8, TLP18.3, CAB proteins, SECY, ALB3, SPR54; Supplemental tables 2 B and D) that could point to the presence of a machinery for co-translational membrane insertion. However, this requires further analyses.

Taken together, results obtained from MS analysis of high molecular weight fractions point to an association of HCF107 with a membrane-bound translation machinery due to the presence of many ribosomal proteins and translational factors (e.g. HCF173, FUG1). However, a direct association with proteins involved in translation needs to be elucidated. Further, several so far uncharacterized proteins with RNA-binding motifs were identified, which present potential components of the HCF107 complex. However, since the overall number of identified proteins is so high, the repertoire of potential interaction partners is large and needs further experiments to narrow down the number of candidate proteins.

Discussion

Identification of essential residues for HCF107 accumulation and/ or function

The classification of HCF107 into the superfamily of helical repeat proteins given by the presence of 11 tandemly arranged TPR-like (RTPR) repeats indicates a function in protein-protein or protein-RNA interactions (Blatch and Lässle, 1999; Sane et al., 2005). Analyses of mutants affected in single residues that are impaired in protein function could help to reveal the function of HCF107.

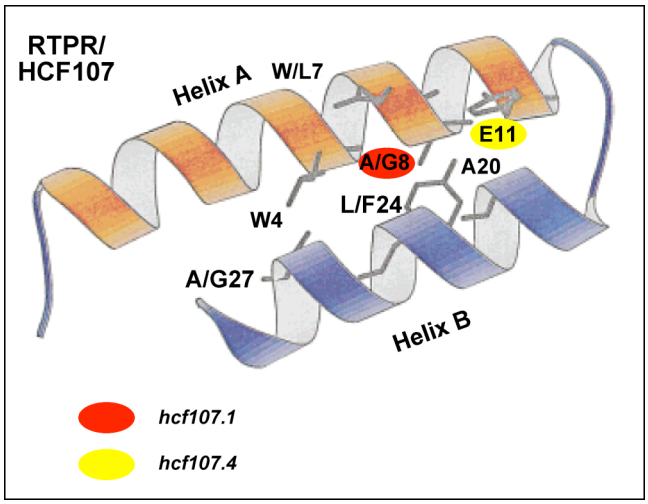


Figure 8: Representation of a RTPR motif forming two antiparallel helices.
Conserved amino acids forming the RTPR consensus sequence are indicated. The aminoacids exchanged in hcf107.1 and hcf107.4 in helix A are indicated (Modified from Batch and Lässle, 1999).

Two allelic loss-of-function mutations were described, which are defined by amino acid exchanges of two conserved amino acids in the A helices of RTPR motifs 3 and 9 (Figure 8; Felder et al., 2001; Lyska, 2006). In *hcf107.1* Ala8 of helix 3 is exchanged by threonine, whereas in *hcf107.4* Glu11 of helix 9 is exchanged to lysine. Analyses on the accumulation of these proteins in *Arabidopsis* revealed that hcf107.4 is strongly reduced when compared to wild type. Based on the fact that in eukaryotic cells misfolded proteins are targeted for degradation by the 26S proteasome (Goldberg, 2003) it is likely that the Glu11 → Lys11 exchange in *hcf107.4* leads to the formation of an alternative structure of the protein. Both amino acids are hydrophilic, but they differ in their charge properties (glutamic acid is negative; lysine is positive), which may alter the folding of RTPR motif 9. In contrast, the amino acid exchange in *hcf107.1* did not affect protein accumulation. Here, a hydrophobic amino acid (Ala8) is exchanged by the neutral threonine. According to structures from TPR proteins Ala8 would be positioned at the interface between helix A and helix B of a RTPR domain (Figure 8). This position was supposed be essential for helix folding since mutations

were previously shown to disrupt protein function of the p62 protein from *Saccharomyces cerevisiae* (Sikorski et al., 1993). However, since the theonine is also a small amino acid without large sidechains, the Ala8 → Thr8 exchange is not likely to influence protein folding, which is supported by the accumulation of the protein to wild-type amounts *in hcf107.1*. Therefore, the Ala8 residue in RTPR motif 3 appears to be required for protein function rather than folding. Since it is located in helix A, which forms a binding groove together with other A helices from tandem arrays of motifs, it may be involved in binding of HCF107 to either another protein or even RNA. Therefore, further analyses using the protein encoded by the *hcf107.1* mutant allele will help understanding the molecular function of the HCF107 protein.

Indications for a role of HCF107 in translation and co-translational membrane insertion of PsbH

HCF107 was formerly shown to form a membrane-associated high molecular weight complex of 600-800 kDa (Sane et al., 2005). Based on the phenotype of mutants affected in HCF107 lacking transcripts with *psbH* as the leading cistron and on the structure of the protein, HCF107 was supposed to be associated with RNA either directly or indirectly via other HCF107 complex components. Analyses presented in this study confirmed the presence of an HCF107-containing high molecular weight complex in membrane fractions. However, using different solubilization conditions, the size of this complex was corrected from 600-800 kDa to way beyond 800 kDa. Further, it could be shown that this high molecular weight complex is degraded by RNase treatment and therefore, indeed, is associated with RNA. A smaller HCF107-complex (~300 kDa) occurred after deletion of RNA indicating the high molecular weight complex displays an aggregate of this ~300 kDa complex, RNA and other factors associating in an RNA-dependent manner. The extremely large size of the RNA-containing complex indicates that it may be associated with polysomes, which also fractionate to about 50% sucrose (Barkan, 1993). This is supported by mass spectromic analyses of fractions containing this complex since the majority of identified factors were ribosomal proteins and factors for translation initiation, elongation and termination. Also, the presence of the D1 translation initiation protein HCF173 (Schult et al., 2007) supports this hypothesis. Therefore, this data provides evidence that the high molecular weight complex formed by HCF107 may be associated with ribisomes and therefore involved in translation.

The two structural homologs of HCF107 from *Chlamydomonas*, MBB1 and NAC2, exhibit similar properties in terms of complex formation. MBB1, which is involved in 5´ end processing, stabilization, and/ or translation of the *psbB* transcript, was also identified as part of a large complex of up to 2000 kDa that is destabilized by RNase treatment (Vaistij et al.,

2000a; Vaistij et al., 2000b). The RNA-free MBB1-complex has a size of about 300 kDa (Vaistij et al., 2000a) standing in agreement with the size of the HCF107-complex that occurs after RNA-degradation. Similarly, the *psbD* transcript stabilizing factor NAC2 is part of a 550 kDa complex, but less abundant complexes of up to 2000 kDa that are supposed to contain ribosomes were also observed (Boudreau et al., 2000). A translational activator, RBP40, was found to be part of the 550 kDa NAC2-complex and bind to a U-rich site in the *psbD* 5´UTR in a NAC2-dependent fashion (Ossenbuehl and Nickelsen, 2000; Schwarz et al., 2007). Like NAC2 and MBB1 many other factors, mainly PPR proteins, were described that play a role in stabilization of transcripts by association with 5´ ends and concomitant initiation of translation (Stern et al., 2010). According to the model provided by Pfalz et al. (2009) and Prikryl et al. (2010) maturation of transcripts encoded by polycistronic transcription units, like the *psbB*-operon, is regulated in the same way. Therefore, he absence of transcripts with *psbH* as the leading cistron in *hcf107* mutants and the formation of a HCF107-high molecular weight complex that is likely to be involved in translation emphasize a role of HCF107 in stabilization and translation of *psbH*.

Interestingly, unlike NAC2 and MBB1, the HCF107-high molecular weight complex is associated with membranes. In fact, chloroplast membranes are generally assumed to be the sites of translation for integral membrane proteins, especially for the subunits of PSII (Zerges, 2000; Marin-Navarro, 2007). In *Chlamydomonas* several RNA-binding proteins (RBPs), inter alia the *psbA* (D1) translation regulator RB47, were found to be associated with low density membranes (LDMs), which are supposed to be sites for *de novo* PSII synthesis (Zerges and Rochaix, 1998). In stromal thylakoid membranes, which are the sites of PSII repair (Uniacke and Zerges, 2007; Mulo et al., 2008), the *psbA* binding protein RBP63 was found to be involved in targeting *psbA* expression to membranes (Ossenbühl et al., 2002). A similar role was discussed for HCF173 from *Arabidopsis*, which, interestingly, was found in the same fractions like the HCF107-high molecular weight complex in the present study. Therefore, the association of the HCF107-high molecular weight complex with chloroplast membranes links *psbH* translation to this compartment and supports the general assumption that PSII subunits are co-translationally inserted into membranes.

Light regulates HCF107-complex formation and membrane-association

Previous studies on transcript levels in dark- vs. light-grown seedlings revealed a light-induced expression of HCF107 at the seedling stage (Sane et al., 2005 and unpublished). In later developmental stages, as shown in this study, the expression of HCF107 is regulated by a circadian rhythm like it was formerly shown for HCF173 (Schult, 2005). Since photosynthesis is subjected to circadian rhythms (Hennessey and Field, 1991), expression of

regulatory proteins might be adjusted to this way of regulation. Thus, light does not play a role in accumulation of the HCF107 protein during the course of day.

Moreover, formation of HCF107-containing complexes was found to be regulated by light and preliminary data suggests that light-regulation of HCF107 also includes its association with chloroplast membranes. In darkness, the HCF107 protein was found to accumulate on chloroplast membranes in a monomeric form and as part of small complexes of up to 300 kDa, which was ascribed to minimal light exposition during preparation of samples. In light, the amount of membrane-associated HCF107 decreased and a complex of again ~300 kDa was found in the stroma fraction, whereas the complete membrane fraction of HCF107 was part of the high molecular weight complex that is involved in translation of the *psbH* transcript. Therefore, the association of HCF107 with RNA and ribosomes is a light-dependent process.

Light-induced regulation of translation has been previously described in *Chlamydomonas*. The abovementioned RBPs that were found associated with LDMs and are supposed to be involved in translation were shown to exhibit enhanced RNA-binding activity in light (Zerges et al., 2002). This process apparently is regulated by rising ATP concentrations induced by photosynthetic electron transport in light (Zerges et al., 2002). For RB47, which forms a complex with three other proteins (RB60, RB38 and RB55), a mechanism for light- and redox state-dependent binding of the psbA 5´ end has been revealed (Danon and Mayfield, 1991; Trebitsh et al., 2000; Kim and Mayfield, 2002; Somanchi et al., 2005). The association of RB47 with the *psbA* 5´ end was found to be regulated by the RB60 protein, a protein disulfide isomerase homolog (cPDI), and by TBA1, an oxidoreductase homolog, in an opponent manner in response to the redox status of the chloroplast, which changes in the course of photosynthetic electron transport (Kim and Mayfield, 1997; Somanchi et al., 2005; Algerand et al., 2006). Also, ADP-dependent phosphorylation of RB60 plays a role in light-dependent regulation of *psbA* translation and the presence of a kinase in the *psbA* transcript-binding complex was postulated (Danon and Mayfield, 1994a).

The HCF107 protein may be part of a similar, regulatory complex that is involved in light-dependent RNA-binding and subsequent initiation of translation. The identification of a ~300kDa, that was found in (i) membranes of dark-adapted plants but was supposed to occur due to minimal light exposure, (iii) the stroma, and (iii) membranes after RNAse-treatment emphasizes the presence of such a complex. According to the *psbA* translation complex in *Chlamydomonas* this complex may contain an RNA-binding protein, a kinase or a redox-sensitive protein and other regulatory proteins. The HCF107 structure as a RTPR protein points to involvement in both, direct RNA-binding or light-activated recruitment of a RNA-binding protein to facilitate association with the *psbH* transcript. However, several proteins with RNA-binding motifs were identified by mass spectrometry in HCF107-high

molecular weight fractions and analysis of these candidates may help identifying the composition of the HCF107 complex and the identity of the RNA-binding protein.
Interestingly, experimental mass spectromic data published on the PhosPhAt 3.0 database (phosphat.mpimp-golm.mpg.de; Haezlewood et al., 2008; Durek et al., 2010) revealed two phosphorylation sites for HCF107 (Figure 9) further sustaining the assumption that HCF107 function may be modulated by light-induced phosphorylation.

AVVDRSS**S**GVFSPQKESANGEGEESNTEEGVLVVRRPLLENSDKESSEEEGKKYPARIDAGLSN
IAKKMPIFEPERSESSSSSSAAAAARAQERPLAVNLDLSLYKAKVLARNFRYKDAEKILEKCIAYW
PEDGRPYVALGKILSKQSKLAEARILYEKGCQSTQGENSYIWQCWAVLENRLGNVRRARELFDA
ATVADKKHVAAWHGWANLEIKQGNISKARNLLAKGLKFCGRNEYIYQTLALLEAKAGRYEQARYL
FKQATICNSRSCASWLAWAQLEIQQERYPAARKLFEKAVQASPKNRFAWHVWGVFEAGVGNVE
RGRKLLKIGHALNPRDPVLLQSLGLLEYKHSSANLARALLRRASELDPRHQPVWIAWGWMEWK
EGNTTTARELYQRAL**S**IDANTESASRCLQAWGVLEQRAGNLSAARRLFRSSLNINSQSYVTWMT
WAQLEEDQGDTERAEEIRNLYFQQRTEVVDDASWVTGFLDIIDPALDTVKRLLNFGQNNDNNRL
TTTLRNMNRTKDSQSNQQPESSAGREDIETGSGFNLDVFLRSKLSLDPLKLDVNLDSKRLERFT
RGRINGA

Figure 9: Phosohorylation sites of the mature HCF107 identified by mass spectrometry.
The mature protein without predicted cTP (68 amino acids) is shown. Phosphorylated peptides as indicated by the PhosPhAt 3.0 database are indicated in red, the phophorylated residue is bold. RTPR domains are underlined.

However, regulation by the redox state of the chloroplast including thioredoxin as it was described for *psbA* translation (Danon and Mayfield, 1994b) may also play a role. Future experiments should address this question.

Tentative model for light-regulated PsbH expression by HCF107

The function of HCF107 in stabilization of *psbH* transcripts and recruitment of the translation machinery indicate the following, tentative model for light-dependent regulation of HCF107 function (Figure 10): In darkness, HCF107 is a membrane-associated monomer (Figure 10 A), which upon illumination and therefore induction of photosynthesis leading and of the chloroplast redox potential and ATP levels, associates with other proteins to form a ~300 kDa complex (Figure 10 B). This process appears to happen very fast like it was also shown for a light-regulated stromal complex containing Calvin cycle enzymes that de- and re-associated within minutes when shifted to light or darkness (Howard et al., 2008). The presence of a ~300 kDa complex in the stroma in light-adapted plants and its absence from membranes indicates that the complex originally formed at membranes dissociates into the stroma. This is supported by the finding that the overall content of membrane-associated HCF107 is decreased in light. The role of the stromal HCF107 complex might be binding of transcripts

with *psbH* as the leading cistron, presumably via the 5´ end, and therefore protection from degradation by 5´→ 3´exoribonuclease (Figure 10 C). Formation of the HCF107-RNA-complex would enable subsequent association of the translation machinery either in the stroma or directly on the membrane (Figure 10 D). The lack of the high molecular weight complex in the stroma indicates association of the complex with membrane-bound ribosomes, but this has to be further confirmed since the data is only preliminary. The high molecular weight complex that is now associated with RNA and ribosomes facilitates co-translational membrane insertion of the PsbH protein (Figure 10 E).

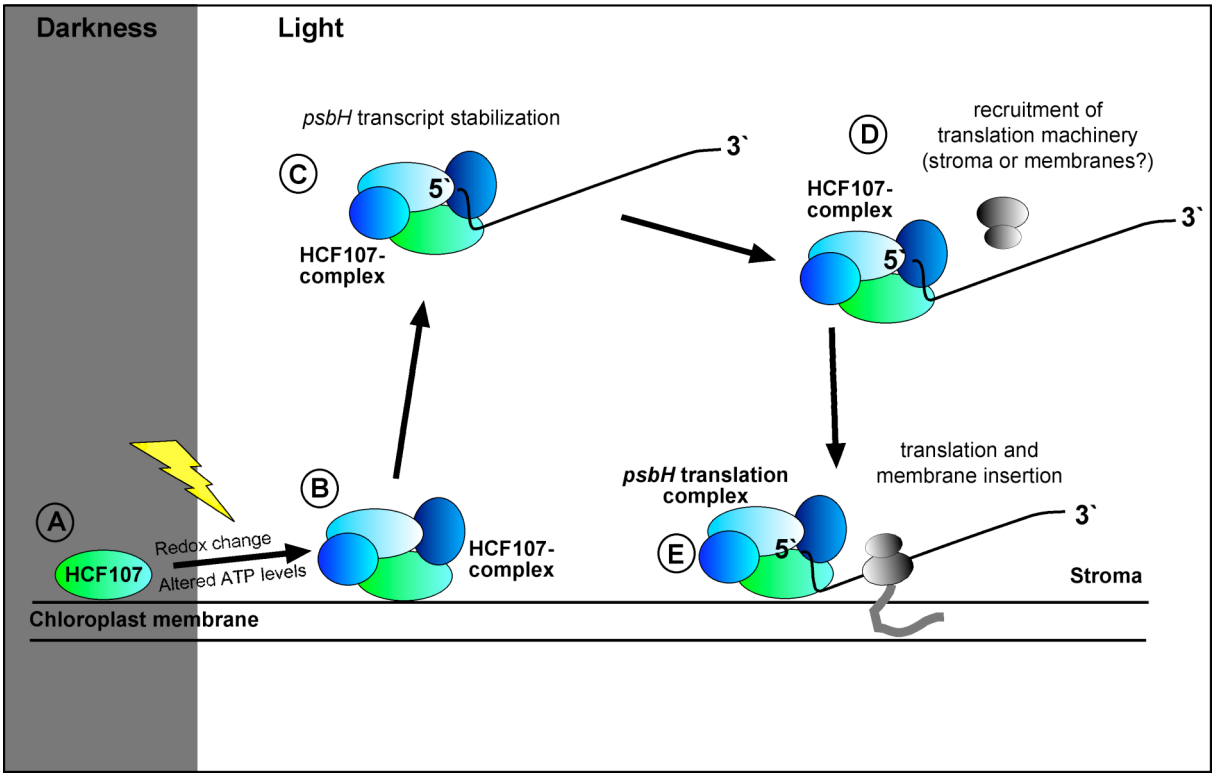


Figure 10: Tentative model for light-dependent regulation of HCF107 function.

The regulation of PsbH synthesis by light in developmental stages that harbor differentiated chloroplasts as it was shown in this study indicates that the PsbH protein needs to be synthesized consistently. It is known that PSII is not only synthesized *de novo*, but it is constantly renewed due to photooxidative damage and that the D1 protein is the major target (Mattoo et al., 1981; Ohad et al., 1990; Aro et al., 1993). However, recent studies on PSII assembly steps in spinach revealed that the PsbH protein was vigorously synthesized and incorporated into PSII core monomers together with the D1 protein therefore suggesting a role for PsbH in the photoinhibition repair cycle of PSII (Rokka et al., 2005).

In *Chlamydomonas*, PSII repair and *de novo* assembly are located in distinct membrane compartments and D1 synthesis is regulated by different regulatory mechanisms (Minai et al., 2006; Uniacke and Zerges, 2007, 2009). Whereas PSII *de novo* synthesis occurs at low

density membranes or T zones, D1 translation for PSII repair was found to be located stromal thylakoid membranes (Zerges and Rochaix, 1998; Uniacke and Zerges, 2007, 2009). It would be interesting to find whether *psbH* translation is also allocated to distinct compartments and if HCF107 indeed plays a role in PsbH synthesis for PSII repair.

Taken together data presented in this study, emphasizes a role of the RTPR protein HCF107 in stabilization and translation of the *psbH* transcript. This process was found to be regulated in a light-dependent manner by modulation of HCF107 complex formation and presumably also its localization in the chloroplast. The role of light-regulated PsbH translation in the synthesis and assembly of PSII remains to be addressed by future studies.

Materials and Methods

Growth conditions and mutant selection

Mutant lines are listed in table 1.1.1. Surface sterilized seeds were plated on 0.5x MS medium (Murashige and Skoog, 1962) supplemented with 1% (w/v) sucrose for mutant selection or on 1x MS medium for analyses of protein accumulation. For all other purposes seeds were sown on Floraton I-soil (Florogard, Oldenburg, Germany).
Seeds were stratified in 4°C for two to five days an then grown in a growth chamber at a constant temperature of 21°C and a PFD of 50-70 μmol s^{-1} m^{-2}. For mutant selection, total protein extraction and seed production plants were grown under long day conditions with a 16 h light/ 8 h darkness period. Plants used for sucrose density centrifugation and membrane protein isolation were grown under short day conditions (8 h light/ 16 h darkness).
Selection of mutant plants exhibiting high chlorophyll fluorescence phenotype was performed in the dark under UV light (Meurer et al., 1996).

Table 1.1.1: Mutant lines used in this study

Mutant allele	Mutagen	Ecotype	Reference/ Source
hcf107.1	EMS	Col-0	Meurer et al. (1996)
hcf107.2	T-DNA	Ws	Felder et al. (2001)
hcf107.3	EMS	Col-0	HHU, Düsseldorf, Germany
hcf107.4	EMS	Col-0	Lyska (2006)
hcf107.5	EMS	Col-0	Lyska (2006)
hcf107.6	EMS	Col-0	Hermanns (2010)
hcf107.7	EMS	Col-0	Hermanns (2010)

SDS-PAGE of proteins, Coomassie staining of gels, immunoblotting and heme staining

Proteins were separated in discontinuous gel systems after Schägger (Schägger and von Jagow, 1987) or Laemmli (1970). Immunodecoration of electroblotted proteins followed the methods described by Meurer et al. (1996). The fusion proteins HCF208-HA/Strep-tag*III*, HCF208-HA/Strep-tag*III*/ProteinA and HCF107-HA/Strep-tag*III*/ProteinA were detected by anti-HA-peroxidase (high affinity (3F10); Roche, Mannheim, Germany). SDS-polyacrylamide gels were stained using PageBlue™ Protein Staining Solution (Fermentas, St. Leon-Rot, Germany) based on the Coomassie Brilliant Blue G-250 dye.

Extraction of total membrane proteins from *Arabidopsis* tissue

Total membrane proteins were isolated from light- and dark-adapted 4-6 week old wild type plants. Material was frozen in liquid nitrogen and ground to fine powder, which was transferred to homogenization buffer (10 mM EDTA; 2 mM EGTA; 50 mM Tris/ HCl, pH 8; 10 mM DTE) and incubated on ice for 10 minutes. After filtration through two layers of Miracloth, membranes were separated from soluble fractions by centrifugation for 10 minutes at 10000 rpm (SS34 rotor) and 4°C. Sedimented membranes were resuspended in carbonate buffer (100 mM Na_2CO_3; 10% (w/v) sucrose; 50 mM DTE) and analyzedby SDS-PAGE and immunoblot.

Extraction of total proteins from *Arabidopsis* tissue

Proteins were extracted according to Shen et al. (2007). Material was pestled in liquid nitrogen and immediately transferred to extraction buffer (10mM Tris/ HCl, pH 7.8; 4 M urea; 5% (w/v) SDS; 15% (w/v) glycerol; 10mM β-mercaptoethanol). The samples were boiled for 4 minutes and cleared by centrifugation at 15,000 x *g* for 5 minutes. After determination of protein concentrations, 50 µg of proteins were loaded on SDS-Polyacrylamide gels without addition of loading buffers.

Isolation and fractionation of crude chloroplasts

Plant material for isolation of intact chloroplasts was grown under short day conditions for four to six weeks. Plant material was homogenized in 10 volumes of homogenization buffer (450 mM sorbitol; 20 mM Tricine/ KOH, pH 8.4; 10 mM EDTA; 10 mM $NaHCO_3$; 0.1% (w/v) BSA)/ gram (fresh weight) in a Waring blender. After filtration through two layers of Miracoth, chloroplasts were centrifuged at 4°C until a speed of 5,900 x *g* were reached. The pellet was resuspended in a small volume in lysis buffer (10 mM Hepes/ KOH (pH 7.8); 10 mM $MgCl_2$; 25 mM KCl) and incubated on ice for 10 minutes. Lysed chloroplasts were centrifuged for 15 minutes at 4°C and 15,000 x *g*. Supernatants were collected and analyzed by sucrose density gradient centrifugation.

Membrane extraction from *Arabidopsis,* solubilization and RNase treatment

Membranes were isolated from plants grown under short day conditions for four to six weeks. If required, plants were kept in darkness for at least 20 hours. Plant material was homogenized in 10 volumes of lysis buffer (10 mM Hepes/ KOH, pH 7.8; 10 mM $MgCl_2$; 25

mM KCl)/ gram (fresh weight) in a Waring blender. After filtration through two layers of Miracoth, membranes were centrifuged at 5900 x *g* and 4°C for two minutes. The pellet was resuspended in a small volume of lysis buffer. The membranes were solubilized with 1% (w/v) digitonin or 1% (w/v) n-dodecyl-β-D-maltoside in lysis buffer (10 mM Hepes/ KOH, pH 7.8; 10 mM $MgCl_2$; 25 mM KCl) for 30 minutes at 4°C and a chlorophyll concentration of 1 mg/ml. Unsolubilized material was removed by centrifugation (20 min at 4°C and 15,000 x *g*). Supernatants from solubilization with digitonin were supplemented with 0.5% (w/v) n-dodecyl-β-D-maltoside.

For analyses of RNA-association of complexes, 400 µl of supernatants were incubated with 2U of RNase A (Roche, Mannheim, Germany) for 30 minutes at 4°C and a rotary shaker. In parallel, samples were incubated without RNase.

Supernatants were loaded on sucrose density gradients.

Sucrose density gradient centrifugation of protein complexes

Linear sucrose density gradients (12 or 36 ml) had concentrations of either 3.4 – 34% (w/v) or 10 – 50% (w/v) sucrose in lysis buffer (10 mM Hepes/ KOH, pH 7.8; 10 mM $MgCl_2$; 25 mM KCl) supplemented with 0.6% (w/v) n-dodecyl-β-D-maltoside. 400 µl of solubilized membrane proteins or stroma proteins were loaded on 12 ml-gradients; for 36 ml-gradients 1.2 ml of solubilized membranes were applied. Gradients were centrifuged at 4°C for 16-17 hours at 38,000 rpm (~ 180,000 x *g*) in a SW40Ti rotor (12 ml gradients) or at 28,000 rpm (~ 104,000 x *g*) in a SW28 rotor (36 ml gradients). After ultracentrifugation 1 ml fractions were collected using a gradient fractionator (Beckmann) and precipitated with 15% (w/v) trichloroacetic acid and analyzed by SDS-PAGE. Gels were subjected to Coomassie staining or immunoblot.

Sample preparation for Mass spectrometry

Protein samples obtained from sucrose density gradient centrifugation were subjected to SDS-PAGE according to Laemmli (1970) and gels were subsequently sliced into several fractions. Each gel slice was diced into small pieces. In gel digestion was performed according to a modified protocol from Shevchenko et al. (2006). After digestion, dried peptides were resuspended in 3% acetonitrile 0.2% trifluoretic acid and cleaned up using Sepak Cartridges (Waters, Milford, Massachusetts, USA). Clean samples were dried and resuspended in 12µl 3% acetonitrile 0.2% formic acid for Mass spectrometry.

Analysis by LC-ESI-MS/MS (Liquid Chromatography Electrospray Ionization Tandem Mass Spectrometry) and data mining

Dried peptides were resuspended in 3% acetonitrile, 0.2% formic acid and analyzed on a LTQ Orbitrap mass spectrometer (Thermo Fischer Scientific, Bremen, Germany) coupled to an Eksigent-Nano-HPLC system (Eksigent Technologies, Dublin (CA), USA). Peptide mixtures were loaded onto laboratory made capillary columns (75 µm inner diameter (BGB Analytik, Boeckten, Switzerland), 8 cm length, packed with Magic C18 AQ beads, 3 µm, 100 Å (Michrom BioResources, Auburn, CA, USA). Peptides were eluted from the column by an increased acetonitrile concentration in the mobile phase from 5% acetonitrile, 0.2% formic acid to 40% acetonitrile, 0.2% formic acid over 74 minutes, followed by a 10 minutes wash step at 5% acetonitrile, 0.2% formic acid. Full-scan MS spectra (300−2000 m/z) were acquired with a resolution of 60000 at 400 m/z after accumulation to a target value of 500000. Collision induced dissociation (CID) MS/MS spectra were recorded in data dependent manner in the ion trap from the six most intense signals above a threshold of 500, using a normalized collision energy of 35% and an activation time of 30 ms. Charge state screening was enabled and singly charge states were rejected. Precursor masses already selected for MS/MS were excluded for further selection for 120 s and the exclusion window was set to 20 ppm. The size of the exclusion list was set to a maximum of 500 entries.
MS/MS spectra were searched with Mascot (Matrix Science, London, UK) version 2.2.04 against the *Arabidopsis thaliana* TAIR9 protein database (download on June 29th, 2009) with a concatenated decoy database supplemented with contaminants (67'079 entries). The search parameters were: requirement for tryptic ends, one missed cleavage allowed, mass tolerance = ± 5 ppm. Beside carbamidomethylation of cysteines as fixed modification, oxidation of methionine was included as variable modification. Individual ions scores higher than 22 indicated identity or extensive homology. Peptide identification was accepted with a minimal Mascot ion score of 23 and a Mascot expectation value below 0.05. The spectrum false discovery rate was calculated by dividing the number of decoy database spectrum assignments by the number of spectrum assignments in the final dataset and was below 1 % for all measured experiments.

Literature

Alergand, T., Peled-Zehavi, H., Katz, Y. and Danon, A. (2006) The chloroplast protein disulfide isomerase RB60 reacts with a regulatory disulfide of the RNA-binding protein RB47. *Plant Cell Physiol*, **47**, 540-548.

Aro, E.M., Virgin, I. and Andersson, B. (1993) Photoinhibition of Photosystem II. Inactivation, protein damage and turnover. *Biochim Biophys Acta*, **1143**, 113-134.

Barkan, A. (1993) Nuclear Mutants of Maize with Defects in Chloroplast Polysome Assembly Have Altered Chloroplast RNA Metabolism. *Plant Cell*, **5**, 389-402.

Barkan, A., Walker, M., Nolasco, M. and Johnson, D. (1994) A nuclear mutation in maize blocks the processing and translation of several chloroplast mRNAs and provides evidence for the differential translation of alternative mRNA forms. *EMBO J*, **13**, 3170-3181.

Ben-Yehuda, S., Dix, I., Russell, C.S., McGarvey, M., Beggs, J.D. and Kupiec, M. (2000) Genetic and physical interactions between factors involved in both cell cycle progression and pre-mRNA splicing in Saccharomyces cerevisiae. *Genetics*, **156**, 1503-1517.

Blatch, G.L. and Lassle, M. (1999) The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *Bioessays*, **21**, 932-939.

Boudreau, E., Nickelsen, J., Lemaire, S.D., Ossenbuhl, F. and Rochaix, J.D. (2000) The Nac2 gene of Chlamydomonas encodes a chloroplast TPR-like protein involved in psbD mRNA stability. *EMBO J*, **19**, 3366-3376.

Boulouis, A., Raynaud, C., Bujaldon, S., Aznar, A., Wollman, F.A. and Choquet, Y. (2011) The nucleus-encoded trans-acting factor MCA1 plays a critical role in the regulation of cytochrome f synthesis in Chlamydomonas chloroplasts. *Plant Cell*, **23**, 333-349.

Cartegni, L., Chew, S.L. and Krainer, A.R. (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet*, **3**, 285-298.

Danon, A. and Mayfield, S.P. (1991) Light regulated translational activators: identification of chloroplast gene specific mRNA binding proteins. *EMBO J*, **10**, 3993-4001.

Danon, A. and Mayfield, S.P. (1994) ADP-dependent phosphorylation regulates RNA-binding in vitro: implications in light-modulated translation. *EMBO J*, **13**, 2227-2235.

Das, A.K., Cohen, P.W. and Barford, D. (1998) The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *EMBO J*, **17**, 1192-1199.

Deng, X.W. and Gruissem, W. (1987) Control of plastid gene expression during development: the limited role of transcriptional regulation. *Cell*, **49**, 379-387.

Durek, P., Schmidt, R., Heazlewood, J.L., Jones, A., MacLean, D., Nagel, A., Kersten, B. and Schulze, W.X. (2010) PhosPhAt: the Arabidopsis thaliana phosphorylation site database. An update. *Nucleic Acids Res*, **38**, D828-834.

Felder, S., Meierhoff, K., Sane, A.P., Meurer, J., Driemel, C., Plucken, H., Klaff, P., Stein, B., Bechtold, N. and Westhoff, P. (2001) The nucleus-encoded HCF107 gene of Arabidopsis provides a link between intercistronic RNA processing and the accumulation of translation-competent psbH transcripts in chloroplasts. *Plant Cell*, **13**, 2127-2141.

Goldberg, A.L. (2003) Protein degradation and protection against misfolded or damaged proteins. *Nature*, **426**, 895-899.

Han, D., Oh, J., Kim, K., Lim, H. and Kim, Y. (2007) Crystal structure of YrrB: a TPR protein with an unusual peptide-binding site. *Biochem Biophys Res Commun*, **360**, 784-790.

Heazlewood, J.L., Durek, P., Hummel, J., Selbig, J., Weckwerth, W., Walther, D. and Schulze, W.X. (2008) PhosPhAt: a database of phosphorylation sites in Arabidopsis thaliana and a plant-specific phosphorylation site predictor. *Nucleic Acids Res*, **36**, D1015-1021.

Hennessey, T.L. and Field, C.B. (1991) Circadian Rhythms in Photosynthesis : Oscillations in Carbon Assimilation and Stomatal Conductance under Constant Conditions. *Plant Physiol*, **96**, 831-836.

Hermanns, T. (2010) Molekularbiologische und biochemische Untersuchungen zur Aufklärung der Funktion des TPR-Proteins HCF107 im Chloroplasten. Diplomarbeit, HHU Düsseldorf.

Herrmann, R.G., Westhoff, P., Alt, J., Tittgen, J. and Nelson, N. (1985) Thylakoid membrane proteins and their genes. In van Vloten-Doting, L., Groot, G.S.P and Hall, T.C. (EDs) Molecular form and function of the plant genome. *Plenum Publishing Corporation, New York*, 233-256.

Herrmann, R.G., Westhoff, P. and Link, G. (1992) Biogenesis of plastids in higher plants. In Plant Gene Research: Cell Organelles, R.G. Herrmann, ed (Vienna: Springer Verlag), 275-349.

Hinsen, A. (2008) Molekularbiologische Untersuchung der RNA-Prozessierungsmutante hcf107 von Arabidopsis thaliana. Diplomarbeit, HHU Düsseldorf.

Hirose, T. and Sugiura, M. (1997) Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast ndhD mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. *EMBO J*, **16**, 6804-6811.

Howard, T.P., Metodiev, M., Lloyd, J.C. and Raines, C.A. (2008) Thioredoxin-mediated reversible dissociation of a stromal multiprotein complex in response to changes in light availability. *Proc Natl Acad Sci U S A*, **105**, 4056-4061.

Kim, J. and Mayfield, S.P. (1997) Protein disulfide isomerase as a regulator of chloroplast translational activation. *Science*, **278**, 1954-1957.

Kim, J. and Mayfield, S.P. (2002) The active site of the thioredoxin-like domain of chloroplast protein disulfide isomerase, RB60, catalyzes the redox-regulated binding of chloroplast poly(A)-binding protein, RB47, to the 5' untranslated region of psbA mRNA. *Plant Cell Physiol*, **43**, 1238-1243.

Laemmli, U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680-685.

Lyska, D. (2006) Molekularbiologische Charakterisierung ethylmethansulfonat-induzierter Photosynthesemutanten von Arabidopsis thaliana. Diplomarbeit, HHU Düsseldorf.Molekularbiologische und biochemische Untersuchungen zur Aufklärung der Funktion des TPR-Proteins HCF107 im Chloroplasten.

Marin-Navarro, J., Manuell, A.L., Wu, J. and S, P.M. (2007) Chloroplast translation regulation. *Photosynth Res*, **94**, 359-374.

Mattoo, A.K., Pick, U., Hoffman-Falk, H. and Edelman, M. (1981) The rapidly metabolized 32,000-dalton polypeptide of the chloroplast is the "proteinaceous shield" regulating photosystem II electron transport and mediating diuron herbicide sensitivity. *Proc Natl Acad Sci U S A*, **78**, 1572-1576.

Meurer, J., Meierhoff, K. and Westhoff, P. (1996) Isolation of high-chlorophyll-fluorescence mutants of Arabidopsis thaliana and their characterisation by spectroscopy, immunoblotting and northern hybridisation. *Planta*, **198**, 385-396.

Michael, T.P., Mockler, T.C., Breton, G., McEntee, C., Byer, A., Trout, J.D., Hazen, S.P., Shen, R., Priest, H.D., Sullivan, C.M., Givan, S.A., Yanovsky, M., Hong, F., Kay, S.A. and Chory, J. (2008) Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genet*, **4**, e14.

Minai, L., Wostrikoff, K., Wollman, F.A. and Choquet, Y. (2006) Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. *Plant Cell*, **18**, 159-175.

Miura, E., Kato, Y., Matsushima, R., Albrecht, V., Laalami, S. and Sakamoto, W. (2007) The balance between protein synthesis and degradation in chloroplasts determines leaf variegation in Arabidopsis yellow variegated mutants. *Plant Cell*, **19**, 1313-1328.

Mockler, T.C., Michael, T.P., Priest, H.D., Shen, R., Sullivan, C.M., Givan, S.A., McEntee, C., Kay, S.A. and Chory, J. (2007) THE DIURNAL PROJECT: Diurnal and circadian

expression profiling, model-based pattern matching and promoter analysis. Cold Spring Harb Symp Quant Biol, **72**, 353-363.

Mulo, P., Sirpio, S., Suorsa, M. and Aro, E.M. (2008) Auxiliary proteins involved in the assembly and sustenance of photosystem II. *Photosynth Res*, **98**, 489-501.

Murashige, T and Skoog, F. (1962) A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol Plant*, **15**, 473-497.

Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M. and Rochaix, J.D. (1999) Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in Chlamydomonas. *Plant Cell*, **11**, 957-970.

Ohad, I., Adir, N., Koike, H., Kyle, D.J. and Inoue, Y. (1990) Mechanism of photoinhibition in vivo. A reversible light-induced conformational change of reaction center II is related to an irreversible modification of the D1 protein. *J Biol Chem*, **265**, 1972-1979.

Ossenbuhl, F., Hartmann, K. and Nickelsen, J. (2002) A chloroplast RNA binding protein from stromal thylakoid membranes specifically binds to the 5' untranslated region of the psbA mRNA. *Eur J Biochem*, **269**, 3912-3919.

Ossenbuhl, F. and Nickelsen, J. (2000) cis- and trans-Acting determinants for translation of psbD mRNA in Chlamydomonas reinhardtii. *Mol Cell Biol*, **20**, 8134-8142.

Pfalz, J., Bayraktar, O.A., Prikryl, J. and Barkan, A. (2009) Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *EMBO J*, **28**, 2042-2052.

Prikryl, J., Rojas, M., Schuster, G. and Barkan, A. (2010) Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proc Natl Acad Sci U S A*, **108**, 415-420.

Raynaud, C., Loiselay, C., Wostrikoff, K., Kuras, R., Girard-Bascou, J., Wollman, F.A. and Choquet, Y. (2007) Evidence for regulatory function of nucleus-encoded factors on mRNA stabilization and translation in the chloroplast. *Proc Natl Acad Sci U S A*, **104**, 9093-9098.

Rokka, A., Suorsa, M., Saleem, A., Battchikova, N. and Aro, E.M. (2005) Synthesis and assembly of thylakoid protein complexes: multiple assembly steps of photosystem II. *Biochem J*, **388**, 159-168.

Sane, A.P., Stein, B. and Westhoff, P. (2005) The nuclear gene HCF107 encodes a membrane-associated R-TPR (RNA tetratricopeptide repeat)-containing protein involved in expression of the plastidial psbH gene in Arabidopsis. *Plant J*, **42**, 720-730.

Schagger, H. and von Jagow, G. (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem*, **166**, 368-379.

Schmitz-Linneweber, C. and Small, I. (2008) Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends Plant Sci*, **13**, 663-670.

Schult, K. (2005) Molekularbiologische und biochemische Analysen zur Funktion des nukleären Hilfsfaktors HCF173 bei der Biogenese des Photosystems II in Arabidopsis thaliana. Dissertation, HHU Düsseldorf.The nuclear-encoded factor HCF173 is involved in the initiation of translation of the psbA mRNA in Arabidopsis thaliana. *Plant Cell*, **19**, 1329-1346

Schult, K., Meierhoff, K., Paradies, S., Toller, T., Wolff, P. and Westhoff, P. (2007) The nuclear-encoded factor HCF173 is involved in the initiation of translation of the psbA mRNA in Arabidopsis thaliana. *Plant Cell*, **19**, 1329-1346.

Schwarz, C., Elles, I., Kortmann, J., Piotrowski, M. and Nickelsen, J. (2007) Synthesis of the D2 protein of photosystem II in Chlamydomonas is controlled by a high molecular mass complex containing the RNA stabilization factor Nac2 and the translational activator RBP40. *Plant Cell*, **19**, 3627-3639.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. and Mann, M. (2006) In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc*, **1**, 2856-2860.

Sikorski, R.S., Michaud, W.A. and Hieter, P. (1993) p62cdc23 of Saccharomyces cerevisiae: a nuclear tetratricopeptide repeat protein with two mutable domains. *Mol Cell Biol*, **13**, 1212-1221.

Small, I.D. and Peeters, N. (2000) The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci*, **25**, 46-47.

Somanchi, A., Barnes, D. and Mayfield, S.P. (2005) A nuclear gene of Chlamydomonas reinhardtii, Tba1, encodes a putative oxidoreductase required for translation of the chloroplast psbA mRNA. *Plant J*, **42**, 341-352.

Stern, D.B., Goldschmidt-Clermont, M. and Hanson, M.R. (2010) Chloroplast RNA metabolism. *Annu Rev Plant Biol*, **61**, 125-155.

Sturm, N.R., Kuras, R., Buschlen, S., Sakamoto, W., Kindle, K.L., Stern, D.B. and Wollman, F.A. (1994) The petD gene is transcribed by functionally redundant promoters in Chlamydomonas reinhardtii chloroplasts. *Mol Cell Biol*, **14**, 6171-6179.

Sugiura, M. (1992) The chloroplast genome. *Plant Mol Biol*, **19**, 149-168.

Trebitsh, T., Levitan, A., Sofer and Danon, A. (2000) Translation of chloroplast psbA mRNA is modulated in the light by counteracting oxidizing and reducing activities. *Mol Cell Biol*, **20**, 1116-1123.

Uniacke, J. and Zerges, W. (2007) Photosystem II assembly and repair are differentially localized in Chlamydomonas. *Plant Cell*, **19**, 3640-3654.

Uniacke, J. and Zerges, W. (2009) Chloroplast protein targeting involves localized translation in Chlamydomonas. *Proc Natl Acad Sci U S A*, **106**, 1439-1444.

Vaistij, F.E., Boudreau, E., Lemaire, S.D., Goldschmidt-Clermont, M. and Rochaix, J.D. (2000a) Characterization of Mbb1, a nucleus-encoded tetratricopeptide-like repeat protein required for expression of the chloroplast psbB/psbT/psbH gene cluster in Chlamydomonas reinhardtii. *Proc Natl Acad Sci U S A*, **97**, 14813-14818.

Vaistij, F.E., Goldschmidt-Clermont, M., Wostrikoff, K. and Rochaix, J.D. (2000b) Stability determinants in the chloroplast psbB/T/H mRNAs of Chlamydomonas reinhardtii. *Plant J*, **21**, 469-482.

Wollman, F.A., Minai, L. and Nechushtai, R. (1999) The biogenesis and assembly of photosynthetic proteins in thylakoid membranes1. *Biochim Biophys Acta*, **1411**, 21-85.

Zerges, W. (2000) Translation in chloroplasts. *Biochimie*, **82**, 583-601.

Zerges, W. and Rochaix, J.D. (1998) Low density membranes are associated with RNA-binding proteins and thylakoids in the chloroplast of Chlamydomonas reinhardtii. *J Cell Biol*, **140**, 101-110.

Zerges, W., Wang, S. and Rochaix, J.D. (2002) Light activates binding of membrane proteins to chloroplast RNAs in Chlamydomonas reinhardtii. *Plant Mol Biol*, **50**, 573-585.

Supplemetal Figures

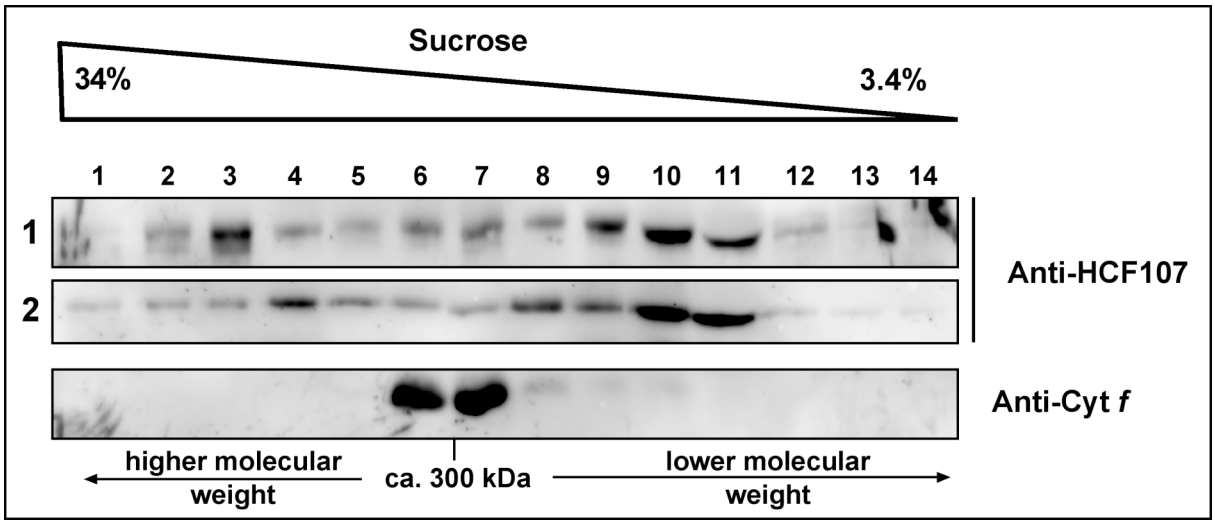


Figure S1: Analysis of HCF107 complexes by sucrose density gradient centrifugation.
Analyses of HCF107-complexes using 3.4-34% sucrose gradients. Membrane proteins were isolated and solubilized with 1% (w/v) digitonin.
Gradients were collected in 1 ml fractions and analyzed by SDS-PAGE and Western blot. Membranes were immunodecorated with Anti-HCF107 and Anti-Cytochrome *f* as a control.

Supplemental Data

Supplemental Table 1 A: Ribosomal proteins

Accession	Description
AT1G05190	emb2394 (Ribosomal protein L6 family)
AT1G07320	RPL4 (ribosomal protein L4)
AT1G32990	PRPL11 (plastid ribosomal protein l11)
AT1G35680	Ribosomal protein L21
AT1G48350	Ribosomal L18p/L5e family protein
AT1G68590	Ribosomal protein PSRP-3/Ycf65
AT1G74970	RPS9, TWN3 (ribosomal protein S9)
AT1G75350	emb2184 (Ribosomal protein L31)
AT1G79850	RPS17, CS17, PRPS17 (ribosomal protein S17)
AT2G24090	Ribosomal protein L35
AT2G33450	Ribosomal L28 family
AT2G33800	Ribosomal protein S5 family protein
AT2G38140	PSRP4 (plastid-specific ribosomal protein 4)
AT2G43030	Ribosomal protein L3 family protein
AT3G13120	Ribosomal protein S10p/S20e family protein
AT3G15190	chloroplast 30S ribosomal protein S20, putative
AT3G25920	RPL15 (ribosomal protein L15)
AT3G27160	GHS1 (Ribosomal protein S21 family protein)
AT3G44890	RPL9 (ribosomal protein L9)
AT3G54210	Ribosomal protein L17 family protein
AT3G56910	PSRP5 (plastid-specific 50S ribosomal protein 5)
AT3G63490	Ribosomal protein L1p/L10e family
AT4G01310	Ribosomal L5P family protein
AT4G17560	Ribosomal protein L19 family protein
AT5G13510	Ribosomal protein L10 family protein
AT5G14320	Ribosomal protein S13/S18 family
AT5G30510	RPS1, ARRPS1 (ribosomal protein S1)
AT5G40950	RPL27 (ribosomal protein large subunit 27)
AT5G47190	Ribosomal protein L19 family protein
AT5G65220	Ribosomal L29 family protein
ATCG00065	RPS12A, RPS12 (ribosomal protein S12A)
ATCG00160	RPS2 (ribosomal protein S2)
ATCG00330	RPS14 (chloroplast ribosomal protein S14)
ATCG00380	RPS4 (chloroplast ribosomal protein S4)
ATCG00650	RPS18 (ribosomal protein S18)
ATCG00660	RPL20 (ribosomal protein L20)
ATCG00750	RPS11 (ribosomal protein S11)
ATCG00760	RPL36 (ribosomal protein L36)
ATCG00770	RPS8 (ribosomal protein S8)
ATCG00780	RPL14 (ribosomal protein L14)

Accession	Description
ATCG00790	RPL16 (ribosomal protein L16)
ATCG00800	structural constituent of ribosome
ATCG00810	RPL22 (ribosomal protein L22)
ATCG00830	RPL2.1 (ribosomal protein L2)
ATCG00840	RPL23.1, RPL23 (ribosomal protein L23.1)
ATCG00900	RPS7.1, RPS7 (Ribosomal protein S7p/S5e family protein)
ATCG01020	RPL32 (ribosomal protein L32)
ATCG01120	RPS15 (chloroplast ribosomal protein S15)

Supplemental Table 1 B: Translation/ RNA-associated

Accession	Description	Chloroplast?	Function
AT1G01080	RNA-binding (RRM/RBD/RNP motifs) family protein	yes	Unknown
AT1G16720	HCF173 (high chlorophyll fluorescence phenotype 173)	yes	D1 translation
AT1G17220	FUG1 (Translation initiation factor 2, small GTP-binding protein)	yes	Translation initiation factor
AT1G64510	Translation elongation factor EF1B/ribosomal protein S6 family protein	yes	Structural constituent of ribisome
AT3G17040	HCF107 (high chlorophyll fluorescent 107)	yes	RNA processing/ stability
AT3G52150	RNA-binding (RRM/RBD/RNP motifs) family protein	yes	Unknown
AT4G20360	ATRAB8D, ATRABE1B, RABE1b (RAB GTPase homolog E1B)	yes	Translation elongation factor activity
AT4G29060	emb2726 (elongation factor Ts family protein)	yes	Translation elongation factor
AT4G35250	HCF244 (NAD(P)-binding Rossmann-fold superfamily protein)	yes	D1 translation
AT5G26742	emb1138 (DEAD box RNA helicase (RH3))	yes	ribosomal rRNA biogenesis
AT5G54600	Translation protein SH3-like family protein	yes	Structural constituent of ribisome

Supplemental Table 1 C: Known proteins/ no RNA-association

Accession	Description	Chloroplast?	Function
AT1G01090	PDH-E1 ALPHA (pyruvate dehydrogenase E1 alpha)	yes	Pyruvat dehydrogenase; glycolysis,
AT1G65260	PTAC4, VIPP1 (plastid transcriptionally active 4)	yes	Thylakoid membrane organization, vesicle organization
AT2G05620	PGR5 (proton gradient regulation 5)	yes	Cyclic electron transport
AT2G18710	SCY1 (SECY homolog 1)	yes	SEC translocase; protein import into/ through chloroplast thylakoid membrane
AT2G28800	ALB3 (63 kDa inner membrane family protein)	yes	Protein import into chloroplast thylakoid membrane
AT2G38040	CAC3 (acetyl Co-enzyme a carboxylase carboxyltransferase alpha subunit)	yes	Lipid metabolism
AT3G44310	NIT1, ATNIT1, NITI (nitrilase 1)	yes	Carbon nitrogen hydrolase
AT5G22830	ATMGT10, GMN10, MGT10, MRS2-11 (magnesium (Mg) transporter 10)	yes	Magnesium transporter
AT5G23060	CaS (calcium sensing receptor)	yes	Modulation of cytoplasmic Ca2+ concentration; stomatal regulation in response to elevations of external Ca2+
ATCG00280	PSBC (photosystem II reaction center protein C)	yes	Photosynthetic electron transport
ATCG00350	PSAA (Photosystem I, PsaA/PsaB protein)	yes	Photosynthetic electron transport
ATCG00520	YCF4 (unfolded protein binding)	yes	Photosystem I assembly
ATCG00680	PSBB (photosystem II reaction center protein B)	yes	Photosynthetic electron transport

Supplemental Table 1 D: Uncharacterized proteins

Accession	Description	Chloroplast?	Function
AT1G08150	ATCHX5, CHX5 (Cation/hydrogen exchanger family protein)	no: cytoplasm	
AT1G34430	EMB3003 (2-oxoacid dehydrogenases acyltransferase family protein)	yes	Acyl-lipid metabolism
AT1G74470	Pyridine nucleotide-disulphide oxidoreductase family protein	yes	Chlorophyll biosynthesis; geranylgeranyl reductase activity
AT2G10940	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	yes	Lipid transport
AT3G45760	Nucleotidyltransferase family protein	no: cytoplasm	
AT4G10490	2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase superfamily protein	no:cytoplasm	
AT4G29654	Protein kinase superfamily protein	unknown cellular component	Protein phosphorylation
AT5G14910	Heavy metal transport/detoxification superfamily protein	yes	Metal ion transport
AT5G42610	Protein of unknown function (DUF607)	no:mitochondria	
AT5G55220	Trigger factor type chaperone family protein	yes	Protein folding, protein transport

Supplemental Table 2 A: RNA-associated processes/ translation

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT3G46780	plastid transcriptionally active 16	97	23	yes	NAD(P) rossman fold
AT3G52150	RNA-binding (RRM/RBD/RNP motifs) family protein	97	20	unknown	RNA binding
AT5G54600	Translation protein SH3-like family protein	71	15	yes	Structural constituent of ribosome
AT1G32990	PRPL11 | plastid ribosomal protein l11	63	14	yes	Mutant has Decreased effective quantum yield of photosystem II
AT1G16720	HCF173 (high chlorophyll fluorescence phenotype 173)	53	23	yes	D1 translation
AT5G26742	DEAD box RNA helicase (RH3)	51	14	yes	Translation initiation factors have this domain
AT5G63420	emb2746 (RNA-metabolising metallo-beta-lactamase family protein)	50	20	yes	Unknown function
AT4G20360	ATRAB8D, ATRABE1B, RABE1b (RAB GTPase homolog E1B)	42	12	yes	Translation elongation factor activity
AT1G09590	Translation protein SH3-like family protein	37	17	yes	Structural constituent ribosome
AT1G57660	Translation protein SH3-like family protein	34	15	?	Structural constituent of ribosome
AT4G35250	HCF244 (NAD(P)-binding Rossmann-fold superfamily protein)	32	6	yes	D1 translation
AT4G29060	Symbols: emb2726 | elongation factor Ts family protein | chr4:14317744-14321315 FORWARD LENGTH=953	26	12	yes	RNA binding; translation elongation
AT3G49910	Translation protein SH3-like family protein	25	12	yes	Structural constituent ribosome
AT1G07920	GTP binding Elongation factor Tu family protein	20	7	yes	Translation elongation
AT1G64510	Translation elongation factor EF1B/ribosomal protein S6 family protein	13	3	yes	Structural constituent of ribosome
AT1G01080	RNA-binding (RRM/RBD/RNP motifs) family protein	10	7	yes	Unknown function
AT1G71720	Nucleic acid-binding proteins superfamily	10	5	yes	RNA binding
AT2G37230	Tetratricopeptide repeat (TPR)-like superfamily protein	10	6	yes	PPR
AT3G62910	APG3 (Peptide chain release factor 1) chr3:23257661-23260386 REVERSE LENGTH=422	10	5	yes	Translation termination
AT5G41520	RNA binding Plectin/S10 domain-containing protein	10	5	yes	Structual constituent of ribosome
AT3G17040	HCF107 (high chlorophyll fluorescent 107)	9	5	yes	RNA processing/ stability
AT5G42310	Pentatricopeptide repeat (PPR-like) superfamily protein	8	4	yes	Unknown function
AT1G17220	FUG1 (Translation initiation factor 2, small GTP-binding protein)	6	3	yes	Translation initiation
AT4G37510	Ribonuclease III family protein	6	3	yes	RNC1; RNA processing

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT5G61090	Polynucleotidyl transferase, ribonuclease H-like superfamily protein	5	1	yes	Nucleic acid binding; unknown function
AT5G02250	EMB2730, RNR1, ATMTRNASEII (Ribonuclease II/R family protein)	4	2	yes	RNR1: rRNA maturation
AT1G11290	CRR22 (Pentatricopeptide repeat (PPR) superfamily protein)	3	1	yes	Editing, endonuclease activity
AT3G09650	HCF152, CRM3 (Tetratricopeptide repeat (TPR)-like superfamily protein)	3	1	yes	RNA processing/ stability
AT3G23700	Nucleic acid-binding proteins superfamily	3	2	yes	S1 RNA binding motif; AtHIP

Proteins that are also listed in tables 1 A-D are highlighted in turquoise

Supplemental Table 2 B: Photosynthetic electron transport/ ATP-Synthase

Accession	Description	Total Nb peptides	Unique Nb peptides	Function
ATCG00480	ATPB, PB (ATP synthase subunit beta)	182	42	
ATCG00120	ATPA (ATP synthase subunit alpha)	127	36	
AT4G10340	LHCB5 (light harvesting complex of photosystem II 5)	123	19	
AT1G31330	PSAF (photosystem I subunit F)	77	9	
AT5G01530	LHCB4.1 (light harvesting complex photosystem II)	58	9	
AT3G08940	LHCB4.2 (light harvesting complex photosystem II)	51	9	
AT4G04640	ATPC1 (ATPase, F1 complex, gamma subunit protein)	51	16	
AT1G61520	LHCA3 (photosystem I light harvesting complex gene 3)	49	5	
AT1G15820	LHCB6, CP24 (light harvesting complex photosystem II subunit 6)	40	7	
AT1G68830	STN7 (STT7 homolog STN7)	38	18	State transitions
AT1G79040	PSBR (photosystem II subunit R)	34	5	
AT1G44575	NPQ4, PSBS (Chlorophyll A-B binding family protein)	32	9	
AT1G29910	CAB3, AB180, LHCB1.2 (chlorophyll A/B binding protein 3)	28	5	
AT2G07698	ATPase, F1 complex, alpha subunit protein	25	11	
AT5G08670	ATP synthase alpha/beta family protein	25	13	
ATCG00280	PSBC (photosystem II reaction center protein C)	25	8	
ATCG00020	PSBA (photosystem II reaction center protein A)	23	5	
ATCG00350	PSAA (Photosystem I, PsaA/PsaB protein)	20	10	
ATCG00270	PSBD (photosystem II reaction center protein D)	19	3	
AT4G22890	PGR5-LIKE A (PGR5-LIKE A)	18	5	PGRL1, cyclic electron transport
ATCG00540	PETA (photosynthetic electron transfer A)	16	6	
ATCG00680	PSBB (photosystem II reaction center protein B)	16	8	
ATCG00340	PSAB (Photosystem I, PsaA/PsaB protein)	15	5	
AT1G30380	PSAK (photosystem I subunit K)	13	4	
AT1G03130	PSAD-2 (photosystem I subunit D-2)	11	8	
AT4G12800	PSAL (photosystem I subunit l)	11	3	

Accession	Description	Total Nb peptides	Unique Nb peptides	Function
AT2G05620	PGR5 (proton gradient regulation 5)	10	1	Cyclic electron transport
AT4G32260	ATPase, F0 complex, subunit B/B~, bacterial/chloroplast	10	6	
ATCG00720	PETB (photosynthetic electron transfer B)	10	3	
AT3G47470	LHCA4, CAB4 (light-harvesting chlorophyll-protein complex I subunit A4)	8	2	
AT5G64040	PSAN (photosystem I reaction center subunit PSI-N, chloroplast, putative / PSI-N, putative (PSAN))	8	1	
ATCG00130	ATPF (ATPase, F0 complex, subunit B/B~, bacterial/chloroplast)	7	4	
AT2G46820	PTAC8, TMP14, PSAP, PSI-P (photosystem I P subunit)	6	3	
AT1G55670	PSAG (photosystem I subunit G)	5	2	
AT3G54890	LHCA1 (photosystem I light harvesting complex gene 1)	4	2	
AT4G03280	PETC, PGR1 (photosynthetic electron transfer C)	4	1	
AT4G28750	PSAE-1 (Photosystem I reaction centre subunit IV / PsaE protein)	4	3	
AT4G09650	ATPD (ATP synthase delta-subunit gene)	3	2	

Proteins that are also listed in tables 1 A-D are highlighted in turquoise

Supplemental Table 2 C: Metabolism and transport

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT3G25860	LTA2, PLE2 (2-oxoacid dehydrogenases acyltransferase family protein)	99	16	yes	Subunit of pyruvat dehydrogenase, lipid metabolism
AT3G08580	AAC1 (ADP/ATP carrier 1)	96	22	yes	ATP:ADP antiporter
AT1G74470	Pyridine nucleotide-disulphide oxidoreductase family protein	83	24	yes	Chlorophyll biosynthesis
AT3G16950	LPD1, ptlpd1 (lipoamide dehydrogenase 1)	53	26	yes	Acetyl-CoA biosynthetic process from pyruvate
AT3G45140	(LOX2, ATLOX2 | lipoxygenase 2)	51	22	yes	Pyrovate dehydrogenase complex
AT4G27440	PORB (protochlorophyllide oxidoreductase B)	51	11	yes	Chlorophyll biosynthesis
AT3G26650	GAPA, GAPA-1 (glyceraldehyde 3-phosphate dehydrogenase A subunit)	49	13	yes	Glyceraldehyde-3-phosphate dehydrogenase activity, glycolysis,
AT1G12900	GAPA-2 (glyceraldehyde 3-phosphate dehydrogenase A subunit 2)	44	12	yes	Glyceraldehyde-3-phosphate dehydrogenase activity, glycolysis,
ATCG00500	ACCD (acetyl-CoA carboxylase carboxyl transferase subunit beta)	31	11	yes	Encodes the carboxyltransferase beta subunit of acetyl-CoA carboxylase, involved in de novo fatty acid biosynthesis
AT2G38040	CAC3 (acetyl Co-enzyme a carboxylase carboxyltransferase alpha subunit)	27	18	yes	Encodes the carboxyltransferase alpha subunit of acetyl-CoA carboxylase, involved in de novo fatty acid biosynthesis

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT1G03630	POR C, PORC (protochlorophyllide oxidoreductase C)	24	8	yes	Chlorophyll biosynthesis
AT3G56940	CRD1, CHL27, ACSF (dicarboxylate diiron protein, putative (Crd1))	19	9	yes	Protochlorophyllide synthesis (publications)
AT1G13440	GAPC-2, GAPC2 (glyceraldehyde-3-phosphate dehydrogenase C2)	15	5	yes	Glyceraldehyde-3-phosphate dehydrogenase activity, glycolysis
AT1G30120	PDH-E1 BETA (pyruvate dehydrogenase E1 beta)	14	4	yes	Putative plastid pyruvate dehydrogenase
ATCG00490	RBCL (ribulose-bisphosphate carboxylases)	14	8	yes	Calvin cycle
AT1G01090	PDH-E1 ALPHA (pyruvate dehydrogenase E1 alpha)	10	5	yes	Pyruvat dehydrogenase; glycolysis, metabolic process, oxidation-reduction process
AT3G13930	Dihydrolipoamide acetyltransferase, long form protein	10	5	yes	Metabolic process
AT5G22830	ATMGT10, GMN10, MGT10, MRS2-11 (magnesium (Mg) transporter 10)	10	6	yes	Magnesium transporter
AT4G25080	CHLM (magnesium-protoporphyrin IX methyltransferase)	8	4	yes	Chlorophyll biosynthesis
AT3G14110	FLU (Tetratricopeptide repeat (TPR)-like superfamily protein)	7	2	yes	Chlorophyll biosynthesis
AT3G63410	APG1, VTE3, IEP37, E37 (S-adenosyl-L-methionine-dependent methyltransferases superfamily protein)	6	3	yes	PQ biosynthesis
AT5G18660	PCB2 (NAD(P)-binding Rossmann-fold superfamily protein)	5	3	yes	Chlorophyll biosynthesis
AT5G46110	APE2, TPT (Glucose-6-phosphate/phosphate translocator-related)	5	2	yes	TRIOSE-PHOSPHATE／PHOSPHATE TRANSLOCATOR
AT1G50030	TOR (target of rapamycin)	4	1	?	1-phosphatidylinositol-3-kinase activity
AT3G55190	alpha/beta-Hydrolases superfamily protein	4	1	?	Glycerol biosynthetic process
AT5G14910	Heavy metal transport/detoxification superfamily protein	4	2	yes	Metal ion transport

Proteins that are also listed in tables 1 A-D are highlighted in turquoise

Supplemental Table 2 D: Protein folding, transport, degradation

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT5G42270	VAR1, FTSH5 (FtsH extracellular protease family)	37	18	yes	Metallopeptidase ; D1 degradation
AT5G01920	STN8 (Protein kinase superfamily protein)	35	10	yes	Phosphorylation; PSII repair
AT2G30950	VAR2, FTSH2 (FtsH extracellular protease family)	28	15	yes	Metallopeptidase ; D1 degradation
AT2G18710	SCY1 (SECY homolog 1)	24	9	yes	Protein import into chloroplast thylakoid membrane
AT1G06950	ATTIC110, TIC110 (translocon at the inner envelope membrane of chloroplasts 110)	23	14	yes	Plastid import
AT5G55220	trigger factor type chaperone family protein	22	9	yes	Protein folding, transport
AT1G06430	FTSH8 (FTSH protease 8)	20	10	yes	Protein degradation
AT5G50920	CLPC, ATHSP93-V, HSP93-V, DCA1, CLPC1 (CLPC homologue 1)	20	8	yes	Plastid import, TIC complex
AT4G01800	AGY1, AtcpSecA, SECA1 (Albino or Glassy Yellow 1)	17	9	yes	Protein import into thylakoid membrane; ATPase of SEC-translocon
AT1G29310	SecY protein transport family protein	13	7	yes	Protein import into thylakoid membrane
AT2G25140	HSP98.7, CLPB-M, CLPB4 (casein lytic proteinase B4)	13	3	yes	Chaperonin
AT5G02500	HSC70-1, HSP70-1, AT-HSC70-1, HSC70 (heat shock cognate protein 70-1)	13	8	yes	Protein folding
AT5G03940	FFC, 54CP, CPSRP54, SRP54CP (chloroplast signal recognition particle 54 kDa subunit)	13	5	yes	Protein import into thylakoid membrane, SRP
AT1G80030	Molecular chaperone Hsp40/DnaJ family protein	12	5	yes	Protein folding
AT4G13670	PTAC5 (plastid transcriptionally active 5)	10	3	yes	Protein folding, transport
AT1G65260	PTAC4, VIPP1 (plastid transcriptionally active 4)	7	4	yes	Thylakoid membrane organization, vesicle organization
AT3G08530	Clathrin, heavy chain	7	6	yes	Intracellular protein transport; vesicle-mediated transport
AT2G22360	DNAJ heat shock family protein	6	4	yes	Unfolded protein binding
AT2G28900	OEP16, ATOEP16-L, ATOEP16-1, OEP16-1 (outer plastid envelope protein 16-1)	5	1	yes	Plastid import
AT4G20850	TPP2 (tripeptidyl peptidase ii)	5	3	yes	Proteolysis
AT2G28800	ALB3 (63 kDa inner membrane family protein)	4	3	yes	Protein import into chloroplast thylakoid membrane
AT3G13470	TCP-1/cpn60 chaperonin family protein	4	3	yes	Protein folding
AT4G34450	coatomer gamma-2 subunit, putative / gamma-2 coat protein, putative / gamma-2 COP, putative	4	2	yes	Vesicle mediated transport

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT1G02910	LPA1 (tetratricopeptide repeat (TPR)-containing protein)	3	2	yes	PSII assembly; D1 interaction
AT1G54780	TLP18.3 (thylakoid lumen 18.3 kDa protein)	5	3	yes	PSII repair cycle
AT1G79030	Chaperone DnaJ-domain superfamily protein	3	1	?	Unfolded protein binding
AT4G02510	TOC159, TOC86, PPI2, TOC160, ATTOC159 (translocon at the outer envelope membrane of chloroplasts 159)	3	2	yes	Plastid import

Proteins that are also listed in tables 1 A-D are highlighted in turquoise

Supplemental Table 2 E: Others

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT5G23060	CaS (calcium sensing receptor)	124	23	yes	Modulation of cytoplasmic Ca2+ concentration; stomatal regulation in response to elevations of external Ca2+
AT1G34430	EMB3003 (2-oxoacid dehydrogenases acyltransferase family protein)	92	22	yes	Acyltransferase activity
AT5G01600	ATFER1, FER1 (ferretin 1)	64	15	yes	Iron ion homeostasis
AT4G16155	dihydrolipoyl dehydrogenases	50	27	yes	Cell redox homeostasis
AT2G40300	ATFER4, FER4 (ferritin 4)	39	10	yes	Iron ion homeostasis
AT1G71500	Rieske (2Fe-2S) domain-containing protein	34	8	yes	Chloroplast, envelope, thylakoid, redox
AT1G42970	GAPB (glyceraldehyde-3-phosphate dehydrogenase B subunit)	29	8	yes	Glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphorylating) activity
AT5G35970	P-loop containing nucleoside triphosphate hydrolases superfamily protein	29	16	yes	DNA binding; ATP binding; hydrolase; chloroplast
AT3G56090	ATFER3, FER3 (ferritin 3)	23	7	yes	Iron ion homeostasis
AT3G20820	Leucine-rich repeat (LRR) family protein	19	10	yes	Defense, signalling
AT4G01050	TROL (thylakoid rhodanese-like)	14	4	yes	Tethering of ferredoxin:NADP(+) oxidoreductase to thylakoid membranes
AT1G12150	Plant protein of unknown function (DUF827)	12	2	?	Unknown function
AT1G79600	Protein kinase superfamily protein	12	7	yes	Kinase
AT5G03900	Iron-sulphur cluster biosynthesis family protein	11	4	yes	Unknown function
AT5G12470	Protein of unknown function (DUF3411)	11	3	yes	Unknown function

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
AT5G19940	Plastid-lipid associated protein PAP / fibrillin family protein	11	2	yes	Unknown function
AT4G27700	Rhodanese/Cell cycle control phosphatase superfamily protein	9	2	yes	Transfer of sulfane atom of a thiosulfate to cyanide
AT5G17170	ENH1 (rubredoxin family protein)	9	4	yes	Electron carrier protein
AT5G64940	ATATH13, ATH13, ATOSA1, OSA1 (ABC2 homolog 13)	9	5	yes	Encodes a member of ATH subfamily of ATP-binding cassette (ABC) proteins.
AT1G18990	Protein of unknown function, DUF593	8	1	?	Unknown function
AT5G28150	Plant protein of unknown function (DUF868)	8	1	?	Unknown function
AT1G52510	alpha/beta-Hydrolases superfamily protein	7	2	yes	Hydrolase activity
AT1G27930	Protein of unknown function (DUF579)	6	3	?	Unknown function
AT5G21222	protein kinase family protein	6	3	?	Unknown function
AT5G38660	APE1 (acclimation of photosynthesis to environment)	6	4	yes	Unknown function
AT1G71810	Protein kinase superfamily protein	5	4	yes	Phosphorylation
AT1G74730	Protein of unknown function (DUF1118)	5	1	yes	Unknown function
AT4G06634	zinc finger (C2H2 type) family protein	5	2	?	Nucleic acid binding; zinc ion binding
AT1G43760	DNAse I-like superfamily protein	4	1	?	Endo-/ exonuclease domain
AT1G77490	TAPX (thylakoidal ascorbate peroxidase)	4	3	yes	Ascorbate peroxidases are enzymes that scavenge hydrogen peroxide in plant cells
AT2G01210	Leucine-rich repeat protein kinase family protein	4	1	yes	Kinase activity; phosphorylation
AT2G10940	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	4	3	yes	Lipid transport
AT2G26100	Galactosyltransferase family protein	4	1	yes	Protein glycosylation
AT2G34630	GPPS, GPS1 (geranyl diphosphate synthase 1)	4	1	yes	Geranyl diphosphate synthase
AT3G02060	DEAD/DEAH box helicase, putative	4	1	yes	DNA repair; regulation of transcription
AT4G02990	Mitochondrial transcription termination factor family protein	4	2	?	Transcription termination
AT4G05400	copper ion binding	4	3	?	Copper ion binding
AT4G09010	APX4, TL29 (ascorbate peroxidase 4)	4	2	yes	Oxydation-reduction process; heme binding
AT4G39795	Protein of unknown function (DUF581)	4	1	?	Unknown function

Accession	Description	Total Nb peptides	Unique Nb peptides	Chloroplast?	Function
ATCG00170	RPOC2 (DNA-directed RNA polymerase family protein)	4	3	yes	Transcription
AT3G02450	cell division protein ftsH, putative	3	2	yes	Metalloendopeptidase activity, nucleoside-triphosphatase activity, nucleotide binding, zinc ion binding
AT3G02650	Tetratricopeptide repeat (TPR)-like superfamily protein	3	3	yes	Unknown function
AT3G23820	GAE6 (UDP-D-glucuronate 4-epimerase 6)	3	1	?	NAD-dependent epimerase/ dehydratase
AT4G15110	CYP97B3 (cytochrome P450, family 97, subfamily B, polypeptide 3)	3	1	yes	Electron carrier activity, heme binding, iron ion binding, monooxygenase activity, oxidation-reduction process, oxygen binding
AT5G08740	NDC1 (NAD(P)H dehydrogenase C1)	3	1	yes	NADH dehydrogenase activity,
AT5G38520	alpha/beta-Hydrolases superfamily protein	3	1	yes	Hydrolase activity
ATCG00520	YCF4 (unfolded protein binding)	3	1	yes	PSI assembly

Proteins that are also listed in tables 1 A-D are highlighted in turquoise

Danksagung/ Acknowledgements

Ich danke ganz herzlich…

…Prof. Dr. Peter Westhoff für die Betreuung dieser Arbeit, die vielen hilfreichen Diskussionen und die Möglichkeit diese Arbeit in weitgehender Selbständigkeit durchführen zu können.

…Prof. Dr. Peter Jahns für die Übernahme des Koreferates.

…Dr. Karin Meierhoff für viele Ratschläge und Diskussionen.

…Prof. Dr. Christian Schmitz-Linneweber für seine Einladung nach Berlin und die Einweihung in die Geheimnisse des RIP-chip-Assays.

…Sylvain Bischof für seinen Einsatz im Namen der LC-ESI-MS/MS.

…Susanne Paradies für die häufige Pflege meiner Pflänzeken und ihrer tatkräftigen Unterstützung bei der Venichtung des ein oder anderen Cockails, Bierchens, Weinchens, Sektchens, …

…allen ehemaligen und aktuellen Mitgliedern unserer kleinen aber feinen HCF-Gruppe.

…der Besetzung des Hauptlabors II für die Tolerierung meiner “ausschweifenden” Arbeitsweise.

…dem ganzen Rest der Botanik IV für eine angenehme Arbeitsatmosphäre, ihre Ratschläge und Diskussionen.

…den Mitarbeitern des Dachgewächshauses für die Pflege meiner Pflanzen.

…meiner Familie und meinen Freunden für ihre Unterstützung.

…und ganz besonders Alexander Maier. Danke, dass du mir auch (und gerade) in schwierigen Lagen zur Seite gestanden hast und Danke für Deinen Glauben an mich :*