



Title	Functional analysis of N-terminal domains of Arabidopsis chlorophyllide a oxygenase
Author(s)	Sakuraba, Yasuhito; Yamasato, Akihiro; Tanaka, Ryouichi; Tanaka, Ayumi
Citation	Plant Physiology and Biochemistry, 45(10-11), 740-749 https://doi.org/10.1016/j.plaphy.2007.07.016
Issue Date	2007-10
Doc URL	http://hdl.handle.net/2115/33905
Type	article (author version)
File Information	tanaka2.pdf



[Instructions for use](#)

Title: Functional analysis of N-terminal domains of Arabidopsis chlorophyllide *a* oxygenase.

Article Type: Research Paper

Keywords: Chlorophyllide *a* oxygenase, Chlorophyll cycle, Arabidopsis.

Corresponding Author: Yasuhito Sakuraba,

Corresponding Author's Institution: Insitute of Low Temperature Science

First Author: Yasuhito Sakuraba

Order of Authors: Yasuhito Sakuraba; Akihiro Yamasato; Ryouichi Tanaka, Doctor; Ayumi Tanaka, Doctor

Title

Functional analysis of N-terminal domains of Arabidopsis chlorophyllide *a* oxygenase.

^aYasuhito Sakuraba, ^bAkihiro Yamasato, ^{a,b}Ryouichi Tanaka, ^{a,b}Ayumi Tanaka.

^a: Institute of Low Temperature Science, Hokkaido University, N19 W8 Kita-Ku, Sapporo, 060-0819, Japan

^b: Core Research for Evolutional Science and Technology, Japan Science and Technology Agency, Kawaguchi, 332-0012 Japan

Abstract: Higher plants acclimate to various light environments by changing the antenna size of a light harvesting photosystem. The antenna size of a photosystem is partly determined by the amount of chlorophyll *b* in the light-harvesting complexes. Chlorophyllide *a* oxygenase (CAO) converts chlorophyll *a* to chlorophyll *b* in a two-step oxygenation reaction. In our previous study, we demonstrated that the cellular level of the CAO protein controls accumulation of chlorophyll *b*. We found that the amino acids sequences of CAO in higher plants consist of three domains (A, B, and C domains). The C domain exhibits a catalytic function, and we demonstrated that the combination of the A and B domains regulates the cellular level of CAO. However, the individual function of each of A and B domain has not been determined yet. Therefore, in the present study we constructed a series of deleted CAO sequences that were fused with green fluorescent protein and overexpressed in a chlorophyll *b*-less mutant of *Arabidopsis thaliana*, *chl1-1*, to further dissect functions of A and B domains. Subsequent comparative analyses of the transgenic plants overexpressing B-domain containing proteins and those lacking the B domain determined that there was no significant difference in CAO protein levels. These results indicate that the B domain is not involved in the regulation of the CAO protein levels. Taken together, we concluded that the A domain alone is involved in the regulatory mechanism of the CAO protein levels.

Key word

Chlorophyllide *a* oxygenase, Chlorophyll cycle, *Arabidopsis*.

Introduction

Photosynthetic organisms capture light energy using the light-harvesting antenna system [14]. In higher plants and green algae, the core antenna complexes of the photosystem II (PS II) consist of CP43 and CP47, and those of photosystem I (PSI) consist of P700 chlorophyll *b*-protein complexes (CP1) [13, 14]. These core antenna complexes contain chlorophyll *a* and β -carotene as the photosynthetic pigments [4, 8]. On the other hand, PSI and PSII have light-harvesting chlorophyll *a/b*-protein complexes (LHC) as the peripheral antenna complexes. These complexes consist of chlorophyll *a*, chlorophyll *b*, carotenoids and proteins encoded by the LHC gene family [13, 17, 18]. The antenna size of the photosystems changes drastically depending on the light conditions [33, 36]. This alteration of antenna size is mainly accomplished by the changes in the amounts of LHC per photosystem. The accumulation of LHC is partly determined by the amount of chlorophyll *b* [7, 20, 22]. Chlorophyll *b* is synthesized by chlorophyllide *a* oxygenase (CAO). CAO is a Rieske-type monooxygenase which converts chlorophyll *a* via 7-hydroxymethyl chlorophyll to chlorophyll *b* in a two step oxygenation reaction [27, 34]. A CAO gene was isolated from *Chlamydomonas reinhardtii* for the first time (*Cr*CAO) [34]. Since the first identification of *Cr*CAO, other CAO genes have been isolated from *Arabidopsis thaliana* (*At*CAO) [12], *Prochlorothrix hollandica* (*Ph*CAO) [24], *Oryza sativa* (*Os*CAO) [21], and *Prochlorococcus marinus* (*Pm*CAO) [30]. Nagata et al. [24] found that the amino acid sequence of *Ph*CAO is

approximately 200 residues shorter than those of higher plants' CAO (*At*CAO and *Os*CAO). Higher plants' CAO were found to contain the four domains within the CAO sequences of higher plants, based on the motif prediction and the degree of sequence conservation among the CAO sequences. The N-terminal region of *At*CAO corresponding to the 56 residues (Figure 1) is predicted to be a transit peptide that facilitates the protein import into plastids [11, 19]. The second part consists of the A domain (residues 57 - 170) and is well conserved among higher plants sequences (Figure 1). The third part of *At*CAO consists of the B domain (residues 171 - 200 in the *At*CAO sequence) and is less conserved than A domain (Figure 1). Interestingly, the A and B domains are absent in the prokaryotic sequences. The fourth domain contains a Rieske cluster and an iron-binding motif [35], and was called the C domain. It has been demonstrated that C domain alone is sufficient to catalyze the conversion of chlorophyll *a* to *b* [24]. The deletion of the A domain drastically increased the CAO protein level and decreased the chlorophyll *a/b* ratio in transgenic *Arabidopsis* plants [37]. Similarly, when a *Ph*CAO sequence which only contains C domain was overexpressed in *Arabidopsis thaliana*, the chlorophyll *a/b* ratio drastically reduced [16]. These reports suggest that the A domain is involved in the regulation of the CAO protein levels. Interestingly, this regulatory mechanism did not operate when the chlorophyll-*b* synthesizing activity was deficient [37]. This observation indicates that the level of CAO protein is regulated according to the amount of chlorophyll *b*. Although the functions of the A and C domains are well understood, the function of B domain is not clear. However, an interesting sequence motif, named a PEST motif was identified in the B domain. A PEST motif functions as a target sequence for degradation by the ubiquitin/26S proteasome system [28]. By its definition, PEST sequences contain many proline (P), glutamic acid (E), serine (S), threonine (T), and they are flanked by lysine (K), arginine(R) or histidine (H) residues [29]. Since the ubiquitin system is not found in chloroplasts [6], it is unlikely that the PEST sequence within the B domain is a target of ubiquitinylation. However, since a PEST sequence is hydrophilic as it is obvious from its definition, it is possible that the PEST-like sequence in the B domain is a target of certain proteases within chloroplasts. There are some proteases in chloroplasts [1, 2]. This scenario is plausible because hydrophilic sequences such as the KFERQ motif [10] have been reported as targets of proteolytic mechanisms. If this hypothesis is true, the presence of B domain may be required for the destabilization of CAO. Alternatively, it is also possible that the B domain is a linker that stabilizes the CAO protein. In general, a hydrophilic sequence renders flexibility to its protein structure in aqueous environments [9]. By inserting hydrophilic amino acid sequences between two domains, it can often function to stabilize these domains. If this hypothesis is true, the removal of the B domain would affect negatively the function of the A and/or the C domain. To reveal the functions of each domains of CAO, we examined the protein levels and the *in vivo* chlorophyll *b* biosynthesis in transgenic plants overexpressing various combinations of GFP fusions of each domain in *Arabidopsis*. We found little differences for the CAO protein levels despite the presence or absence of the B domain in these transgenic plants. Therefore, we concluded that the B domain is not involved in the regulation of the CAO protein levels, and that the A domain alone is involved in the regulatory mechanism. We hypothesize that the B domain functions as a linker which stabilizes the CAO protein.

Material & Methods

Plant materials

Arabidopsis thaliana plants were grown at 23°C under continuous light conditions (white fluorescent light at 40 – 60 mmol m⁻² s⁻¹). Seeds from Columbia wild-type (WT) and the *chl1-1* mutant were germinated in soil and grown for 4 weeks. Seeds from the transgenic *Arabidopsis* plants were germinated on plates containing half-strength Murashige-Skoog medium (Wako Pure Chemical Industries, Osaka, Japan), 0.7% (w/v) agar, and 50 mg/l kanamycin. One week old transgenic plants were used for observation by confocal laser scanning microscopy (CLSM), immunoblot analysis, and pigment analysis.

Plasmid construction and plant transformation

In order to overexpress CAO-GFP fusion proteins, we modified the plasmid vector that was constructed in our previous study [37]. The vector contained the CaMV promoters, the tobacco mosaic virus ω sequence, the GFP (S65T) sequence (kindly provided by Dr. Niwa, University of Shizuoka) and the nopaline terminator in the backbone of pGreenII-0029 (kindly provided from R.P. Hellens and P. Mullineaux, John Innes Centre) [15]. The sGFP (S65T) gene in this plasmid was fused with transgenes (full length or truncated *At*CAO cDNA sequences) at the *S*all and *Not*I sites. The DNA sequence integrity of the entire coding region was verified by sequencing. Verified plasmids were subsequently transformed into an *Agrobacterium tumefaciens* strain C58 (strain GV2260) by electroporation using a cuvette with a 1 mm gap under 25 μ F and 1.8 kV. *Arabidopsis* plants wild type and *chl1-1* mutants were transformed by infiltration with *Agrobacterium* cells [5]. Primary *Arabidopsis* transformants were selected on a plate containing 50 mg/l kanamycin as described above.

Immunoblot analysis

Ten milligrams of rosette leaf tissue was homogenized with 100 μ l of extraction buffer (50 mM Tris, pH 6.8, 2 mM EDTA, 10% [w/v] glycerol, 2% [w/v] SDS, and 6% [v/v] 2-mercaptoethanol). Homogenates were centrifuged at 10,000g for 3 min and supernatants (25 μ l) were subjected to 12.5% (w/v) polyacrylamide SDS-PAGE separation. Subsequently, the resolved proteins were electroblotted onto a Hybond-P membrane (GE Healthcare, Buckinghamshire, UK). Anti-GFP rabbit primary antibodies (diluted 1:5000; Invitrogen, Carlsbad, CA) were used to detect CAO-GFP transferred proteins. Anti-rabbit IgG linked to horseradish peroxidase (GE Healthcare) was used as a secondary antibody. The horseradish peroxidase activity was detected with the Enhanced Chemiluminescent plus (ECL plus) western blotting detection system (GE Healthcare) according to the manufacturer-supplied protocol.

Analysis of GFP expression by confocal laser scanning microscopy

Fluorescence images were recorded at the Nikon Imaging Centre in Hokkaido University on a C1si Spectral Imaging Confocal Laser Scanning Microscope (CLSM) System with TE2000-E Inverted microscope (Nikon Corporation, Chiyoda-ku, Tokyo, Japan). The microscope was equipped with a Nikon CFI60 Objective Lens series Plan Achromat 100x N.A. 1.40 Oil Immersion Type lens (Nikon Corporation, Chiyoda-ku, Tokyo, Japan). An argon laser (25 mW) was used to generate an excitation source at 488 nm, and GFP and chlorophyll fluorescence were recorded at 500 – 550 nm and 600 - 680 nm, respectively. The images were processed with EZ-C1 Viewer 3.20 (Nikon Corporation, Tokyo, Japan).

Chlorophyll pigment analysis

Chlorophyll was extracted with acetone from rosette leaves of *Arabidopsis thaliana*. The extracts were centrifuged at 15000g for 10 min at 20 °C. The supernatant was diluted with water to the final acetone concentration of 80% and subjected to HPLC analysis. Pigments were separated on an octadecyl column (Shim-pack CLC-ODS column, 6.0 \times 150mm; Shimadzu, Kyoto, Japan) and eluted with methanol at a flow rate of 1.7 ml/min. Elution profiles were monitored by measuring absorbance at 650 nm. Chlorophyll contents were quantified from the chromatographic peak area.

Results

Constructs for transformation of wild type and *ch1-1* mutant

To reveal the function of each domain of CAO, we made a series of constructs that combine a domain or domains of CAO and GFP, and subsequently overexpressed them in *Arabidopsis* (Figure 2). A sequence encoding GFP was incorporated into the construct to monitor the accumulation of the transgene products *in vivo* by CLSM [26]. The sequence for the CAO-GFP fusion protein was placed downstream from the Cauliflower mosaic virus 35S promoter for constitutive overexpression of the transgene. In addition, a DNA sequence encoding the putative transit sequence of CAO was inserted between the promoter sequence and the protein encoding sequence to localize the overexpressed protein into chloroplasts. We used a WT *Arabidopsis* strain as well as a CAO-deficient mutant, *ch1-1* for transformation. The *ch1-1* mutant contains a 31-bp deletion in its native CAO gene, so that it is unable to synthesize even a trace amount of chlorophyll *b* [27]. The *ch1-1* mutant was used to do comparative analysis of fusion protein levels in the WT background to those in the *ch1-1* background enabled us to examine the effects of chlorophyll *b* on the accumulation of each overexpressed protein.

Effects of B domain on the accumulation of the full-length CAO protein: Comparison GFP-ABC and GFP-AC overexpressors

Since we focused on the analysis of the function of A domain in our previous study [37], we examined all possible combinations of CAO domains (Figure 2) with an emphasis on the function of the B domain in this study. At first, in order to examine the effects of B domain on the accumulation of the full-length CAO protein, we compared the phenotypes of GFP-ABC and GFP-AC overexpressors. Since the GFP-ABC and GFP-AC overexpressors in the WT background essentially exhibited the same results with those of the *ch1-1* background, we only show results with those from the *ch1-1* background. These plants showed similar chlorophyll *a/b* ratio of 3.16 (GFP-ABC) and 3.11 (GFP-AC), and similar chlorophyll contents of 2.01 nmole/mg (GFP-ABC) and 1.95 nmole/mg (GFP-AC) respectively (Table 1). The GFP fluorescence was under a detectable level in both transgenic plants (Figure 3A). Similarly, immunoblot analysis using anti-GFP antibody failed to detect the transgene products (Figure 3B). These results

suggest that B-domain is not required for the regulation of the full-length CAO protein level.

Effects of B-domain on the accumulation of C domain in the absence of A domain: Comparison of GFP-BC and GFP-C overexpressors

In contrast to the results with GFP-ABC and GFP-AC overexpressors, the GFP-BC and GFP-C overexpressors accumulated a substantial amount of the fusion proteins. Strong GFP fluorescence was observed in the chloroplasts of both GFP-BC and GFP-C overexpressors in the *chl1-1* background. It should be noted that GFP-BC and GFP-C overexpressors of both the WT (data not shown) and *chl1-1* backgrounds (Figure 4) showed similar results. The fluorescence of GFP-C overexpressors was slightly weaker than that of GFP-BC overexpressors. The reason for this observation is discussed below. The same high protein levels of transgene products were detected by immunoblot analysis with anti-GFP antibodies in both transgenic plants (Figure 4B). Chlorophyll *a/b* ratios and chlorophyll contents were also at the same level in GFP-C and GFP-BC overexpressing plants (Table 1). Collectively, these data indicate that the B domain alone or the C domain itself is not sufficient to induce degradation of the C domain.

Effects of B domain on the accumulation of A domain: Comparison AB-GFP and A-GFP overexpressors

We subsequently analyzed the effects of the B domain on the accumulation and the function of the A domain. Confocal microscopic observations and immunoblot detection of fusion proteins determined that both AB-GFP and A-GFP overexpressors did not accumulate the transgene products in the WT background (Figure 5A and 5B). The chlorophyll *a/b* ratios and the chlorophyll contents in these transgenic plants were not distinguishable from those of WT. This is likely due to the fact that both AB-GFP and A-GFP fusion proteins lack CAO enzymatic activities. These results indicate that the A domain alone is sufficient to induce protein de-stabilization when chlorophyll *b* is supplied. We emphasize that the A domain functions in trans, in other words, the synthesis of chlorophyll *b* within the same protein is not necessary for the function of the A domain. In the *chl1-1* background, both AB-GFP and A-GFP overexpressors significantly accumulated the fusion proteins. Although, chlorophyll *b* did not accumulate in these plants, total chlorophyll contents were at the same level to those of WT (Table 1). Interestingly, both confocal microscopic and immunoblot analyses demonstrated that the protein level of AB-GFP was higher than that of A-GFP. The reason for this particular observation is discussed below.

Effects of B-domain on the GFP prote in levels: Comparison B-GFP and GFP overexpressors

Finally, we examined the effects of the B domain on the accumulation of the GFP reporter protein, by comparing the phenotypes of B-GFP and GFP overexpressors. Both the B-GFP and GFP overexpressors in the WT background show the same chlorophyll *a/b* ratios and the total chlorophyll contents as WT (Table 1). Likewise the B-GFP and GFP overexpressors in the *chl1-1* background showed the same chlorophyll *a/b* ratios and the total chlorophyll contents as *chl1-1* (data not shown). These results indicate that the overexpression of either B-GFP or GFP did not affect chlorophyll biosynthesis. The failure to affect chlorophyll biosynthesis is likely attributable to the fact that B-GFP and GFP do not have a catalytic function. CLSM observation revealed that the accumulation of B-GFP was at the same level as GFP in the transgenic plants of the WT background (Figure 6A). The fusion protein levels in the B-GFP and GFP overexpressors of the *chl1-1* background were indistinguishable from those in the WT background (data not shown). Similar results were obtained with immunoblot analysis. Protein levels for B-GFP (expected molecular size; 30kD) and GFP (expected molecular size; 27kD) were detected at similarly high levels in the transgenic plants (Figure 6B). These results collectively indicated that the B-domain alone has no effect on the reporter protein level.

Discussion

In this study we were interested to understand the individual functions of all domains comprising the CAO protein. In order to meet this goal, we made several types of transgenic plants that overexpressed all combinations of each domain of CAO. We are particularly interested to analyze the function of the B domain, and we speculated that it is either a protease target sequence or a linker sequence that ensures the correct folding of A and C domains by making a space between A and C domains. We found that the GFP-ABC and GFP-AC overexpressors accumulated the fusion proteins at the same levels, and that these transgenic plants showed similar chlorophyll *a/b* ratios (Figure 3). This data indicated that the B-domain does not promote the de-stabilization of the A and C domains. Furthermore, the B domain alone did not show any de-stabilizing effects on the GFP level (Figure 6A and 6B). Taken together, the B domain is not involved in the regulatory mechanism in which the A domain plays a central role and it is unlikely that the B domain functions as a protease target sequence. Instead, several results imply that the B domain enhances the correct folding of the adjacent domains. For example, GFP fluorescence of GFP-BC was stronger than that of GFP-C (Figure 4A), though the protein levels for GFP-BC and GFP-C were similar in the transgenic plants (Figure 4B). We speculated that the B domain made some space between GFP and the C domain, and this aided in the proper folding

of GFP. This idea seems plausible since the enhancement of GFP fluorescence by the addition of a short hydrophilic sequence has been previously observed with other fusion proteins [23]. It is also possible that the B domain positively contributed to their stability. When we compared the protein levels of AB-GFP and A-GFP in the ch1-1 background, this latter hypothesis seemed to be case with AB-GFP (Figure 5B). At the present time, the exact mechanism for the increase in the fusion protein level is not clearly understood. It is possible that the A-GFP protein is rather unstable due to its incorrect folding. These results supported the idea that the B-domain functions as a linker which stabilizes the CAO protein. Importantly, it is now clear that the chlorophyll-*b* dependent regulatory mechanism is solely attributed to the function of the A domain.

We recently found that a ClpC1 subunit of the chloroplast Clp protease is involved in the regulation of CAO [25]. Since the ClpC1 subunit is a chaperone that recognizes the protease substrate sequences [31, 32], it is tempting to speculate that the ClpC1 recognizes the A domain directly in a chlorophyll-*b* dependent manner. Nevertheless, the involvement of other types of proteases or other factors should not be excluded since mutation in the ClpC1 gene did not abolish the full regulation of the CAO protein levels [25]. Studies from the protease side, as well as those from the substrate side, may greatly contribute to our further understanding of the pigment-dependent regulatory mechanism.

Acknowledgement

We acknowledge financial support from the Grant-in-Aid for Scientific Research (17770027 to R.T.) and the Grant-in-Aid for Creative Scientific Research (no. 17GS0314 to A.T.) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. We would like to thank Dr. Kobayashi and Dr. Saito, Nikon Imaging Center, Hokkaido University, for technical assistance in the use of Confocal Laser Scanning Microscopy.

References

- [1] Z. Adam, I. Adamska, K. Nakabayashi, O. Ostersetzer, K Haussuhi, A. Manuell, B. Zheng, O. Vallon, S.R. Rodermel, K. Shinozaki, and A.K. Clarke, Chloroplast and mitochondrial Protease in Arabidopsis. A Proposed Nomenclature, *Plant Physiology* 125 (2001) 1912-1918
- [2] Z. Adam, and A.K. Clarke, Cutting edge of chloroplast proteolysis, *Trends Plant Sci.* 7 (2002) 451-456.
- [3] B. Albert, A. Matzen, H.L. Hyyrylainen, V.P. Kontinen, R. Meima, J. Chapuis, G. Venema, S. Bron, R. Freudl, J.M. van Dijl, Signal Peptide Peptidase- and Clp-like proteins of *Bacillus subtilis* Required for Efficient Translocation and Processing of Secretory Proteins, *J. Biol. Chem.* 274 (1999)24585-24592
- [4] M. Alfonso, G. Montoya, R. Cases, R. Rodriguez, and R. Picorel, Core antenna complexes, CP43 and CP47, of higher plant photosystem II: Spectral properties, pigment stoichiometry, and amino acid composition, *Biochemistry* 33 (1994) 10494–10500.
- [5] N. Bechtold, J. Ellis, and G. Pelletier, In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *C. R. Adac. Sci. Paris, Life Sci.* 316 (1993) 1194-1199.
- [6] E.P. Beers, T.N. Moreno, and J. Callis, Subcellular localization of ubiquitin and ubiquitinated proteins in *Arabidopsis thaliana*, *J. Biol. Chem.* 267 (1992) 15432–15439.
- [7] G. Bellemare, S. G. Bartlett, and N.H. Chual, Biosynthesis of Chlorophyll *a/b*-binding Polypeptides in Wild Type and the Chlorina *f2* Mutant of Barley, *J. Biol. Chem.* 257 (1982) 7762-7767
- [8] A. Ben-Shem, F. Frolow, N. Nelson, Crystal structure of plant photosystem I, *Nature* 426 (2003) 630–635.
- [9] P.Y. Chou, and G.D. Fasman, Prediction of the secondary structure of proteins from their amino acid sequence, *Advan. Enzymol.* 47 (1978) 45-148
- [10] J.F. Dice, Selective degradation of cytosolic proteins by lysosomes. *Ann. N. Y Acad. Sci.* 674 (1992) 58-64
- [11] O. Emanuelsson, H. Nielsen, S. Brunak, and G. von Heijne, Predicting subcellular localization of proteins based on their N-terminal amino acid sequence, *J. Mol. Biol.* 300 (2000) 1005–1016.
- [12] C.E. Espineda, A.S. Linford, D. Devine, J.A. Brusslan, The AtCAO gene, encoding chlorophyll *a* oxygenase, is required for chlorophyll *b* synthesis in *Arabidopsis thaliana*, *Proc. Natl. Acad. Sci. USA.*31(96) (1999) 10507-10511

- [13] B.R. Green, D.G. Durnford, The chlorophyll-carotenoid proteins of oxygenic photosynthesis, *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47 (1996) 685-714
- [14] A.R. Grossman, D. Bhaya, K.E. Apt, D.M. Kehoe, Light-harvesting complexes in oxygenic photosynthesis : diversity, control, and evolution, *Annu. Rev. Genet.* 29 (1995) 231-288
- [15] R.P. Hellens, E.A. Edwards, N.R. Leyland, S. Bean, and P.M. Mullineaux, pGreen: A versatile and flexible binary Ti vector for *Agrobacterium*-mediated plant transformation, *Plant Mol. Biol.* 42 (2000) 819–832.
- [16] M. Hirashima, S. Satoh, R. Tanaka, A. Tanaka, Pigment Shuffling in Antenna Systems Achieved by Expressing Prokaryotic Chlorophyllide *a* Oxygenase in *Arabidopsis*, *J. Biol. Chem.* 281 (2006) 15385-15393
- [17] S. Jansson, The light-harvesting chlorophyll *a/b*-binding proteins, *Biochim. Biophys. Acta.* 1184 (1994) 1–19.
- [18] S. Jansson, A guide to the Lhc genes and their relatives in *Arabidopsis*, *Trends Plant Sci.* 4 (1999) 236-240.
- [19] T. Konishi, Y. Maruta, K. Shinohara, and A. Watanabe, Transit peptides of thylakoid luminal proteins: The sites of stromal processing are conserved among higher plants, *Plant Cell Physiol.* 34 (1993) 1081–1087.
- [20] A. Kuttkat, I Edhofer, L.A. Eichacker, and H. Paulsen, Light-harvesting Chlorophyll *a/b*-binding Protein Stably Inserts into Etioplast Membranes Supplemented with Zn-pheophytin *a/b*, *J. Biol. Chem.* 272 (1997) 20451-20455
- [21] S. Lee, J-H Kim, ES. You, C-H Lee, H. Hirochika, G. An, Differential regulation of chlorophyll *a* oxygenase genes in rice, *Plant Mol. Biol.* 57:805-818
- [22] A. Melis, Dynamics of photosynthetic membrane composition and function, *Biochimica et Biophysica Acta*, 1058 (1991) 87-106
- [23] D.J. Michael, X. Geng, N.X. Cawley, Y.P. Loh, C.J. Rhodes, P. Drain, R.H. Chow, Fluorescent Cargo Protein in Pancreatic β -Cells: Design Determines Secretion Kinetics at Exocytosis, *Biophysical J.* 87(6) (2004) L3-L5
- [24] N. Nagata, S. Satoh, R. Tanaka, and A. Tanaka, Domain structure of chlorophyllide *a* oxygenase of green plants and *Prochlorothrix hollandica* in relation to catalytic function, *Planta* 218 (2004) 1019-1025
- [25] E. Nakagawara, Y. Sakuraba, A. Yamasato, R. Tanaka, A. Tanaka, Clp protease controls chlorophyll *b* synthesis by regulating the level of chlorophyllide *a* oxygenase, *The Plant J.* 49 (2007) 800-809
- [26] Y. Niwa, T. Hirano, K. Yoshimoto, M. Shimizu, and H. Kobayashi, Non-invasive quantitative detection and applications of nontoxic, S65T-type green fluorescent protein in living plants, *Plant J.* 18 (1999) 455–463.
- [27] U. Oster, R. Tanaka, A. Tanaka, and W. Rudiger, Cloning and functional expression of the gene encoding the key enzyme for chlorophyll *b* biosynthesis (CAO) from *Arabidopsis thaliana*, *Plant J.* 21 (2000) 305–310.
- [28] M. Rechsteiner, and S.W. Rogers, PEST sequences and regulation by proteolysis, *Trends Biochem.Sci.* 21 (1996) 267-271.
- [29] S.W. Rogers, R. Wells, M. Rechsteiner, Amino acid sequences common to rapidly degraded proteins: The PEST hypothesis, *Science* 234 (1986) 364-368.
- [30] S. Satoh, and A. Tanaka, Identification of Chlorophyllide *a* Oxygenase in the *Prochlorococcus* Genome by a Comparative Genome Approach, *Plant Cell Physiol.* 47 (2006) 1622-1629
- [31] J. Shanklin, N.D. Dewitt, J.M. Flanagan, The stroma of higher plant plastids contain ClpP and ClpC, functional homologs of *Escherichia coli* ClpP and ClpA: an archetypal two-component ATP-dependent protease, *Plant Cell.* 7 (1995) 1713–1722
- [32] L.L.E. Sjögren, T.M. MacDonald, S. Sutinen, and A.K. Clarke, Inactivation of the *clpC1* Gene Encoding a

Chloroplast Hsp100 Molecular Chaperone Causes Growth Retardation, Leaf Chlorosis, Lower Photosynthetic Activity, and a Specific Reduction in Photosystem Content, *Plant Physiol.* 136 (2004) 4114-4126

[33] A. Tanaka, and A. Melis, Irradiance-Dependent Changes in the Size and Composition of the Chlorophyll a-b Light-Harvesting Complex in the Green Algae *Dunaliella salina*, *Plant Cell Physiol.* 38 (1997) 17-24

[34] A. Tanaka, H. Ito, R. Tanaka, N. Tanaka, K. Yosida, and K. Okada, Chlorophyll *a* oxygenase (CAO) is involved in chlorophyll *b* formation from chlorophyll *a*, *Proc. Natl. Acad. Sci. U S A* 95 (1998)12719-12723

[35] A. Tomitani, K. Okada, H. Miyashita, H.C.P. Matthijs, T. Ohno, and A. Tanaka, Chlorophyll *b* and phycobilins in the common ancestor of cyanobacteria and chloroplasts, *Nature* 400 (1999) 159–162.

[36] R.G. Walters, P. Horton, Acclimation of *Arabidopsis thaliana* to the environment: changes in composition of the photosynthetic apparatus, *Planta* 195 (1994) 248-256

[37] A. Yamasato, N. Nagata, R. Tanaka, and A. Tanaka, The N-terminal domain of chlorophyllide *a* oxygenase confers protein instability in response to chlorophyll *b* accumulation in *Arabidopsis*. *Plant Cell* 17 (2005) 1585-1597.

Figure Legends

Figure 1 Multi-sequence alignment of the CAO proteins in higher plants (*Arabidopsis thaliana*, *Oryza sativa*, *Lycopersicon esculentum*, *Populus trichocarpa*). A Conserved residues in all organisms are indicated by a black background, and conserved residues in three organisms are indicated by a grey background. Dashes indicate absent residues. Black boxes indicate the region predicted to be a transit peptide. Red, blue and green boxes indicate A, B and C domains, respectively. Purple under bars indicate predicted PEST sequences.

Figure.2 Domain structures of the AtCAO and GFP Proteins and the CAO-GFP Fusion Proteins. A, B, C, t and GFP denote the A,B,C domains of CAO, the predicted transit peptide sequence, and sGFP (S65T), respectively. **A** Domain structures of the AtCAO and GFP proteins. **B** Structures of the CAO-GFP fusion proteins. To analyze the role of the B-domain, several types of transgenic plants were constructed which individually overexpressed various combinations of each domain of CAO and GFP. Four pairs of transgenic plants were compared, each of which overexpressed a B-domain containing protein or a B-domain-less protein.

Figure.3 A & B Comparison of protein levels in transgenic plants overexpressing GFP-AC with GFP-ABC.

A Observations of protein levels of the CAO-GFP fusion proteins in transgenic plants. GFP fluorescence was collected from 500 - 550 nm (middle panels) while the autofluorescence of chlorophyll was collected from 600 - 680 nm (lower panels) in cotyledons using CLSM. The merged images of GFP and chlorophyll fluorescence were shown in the upper panels. GFP fluorescence and chlorophyll fluorescence are respectively indicated by green and red. Scale bar = 20µm **B** Immunoblot analysis of the CAO-GFP fusion protein in transgenic plants. Equal fresh weights (2.5 mg) of rosette leaves were homogenized with the protein extraction buffer (see Materials and Methods), and were subjected to SDS-PAGE. The CAO-GFP fusion or CAO proteins were detected using anti-GFP antibody. Black arrowheads indicate predicted molecular sizes of CAO-fusion proteins (GFP-AC; 78kD, GFP-ABC; 81kD) respectively. Grey arrowheads indicate the predicted molecular size of full-length CAO protein (56kD).

Figure.4 A & B Comparison of protein levels in transgenic plants of GFP-C with GFP-BC.

A Observations of CAO –GFP fusion protein levels in transgenic plants. GFP fluorescence was collected from 500 550 nm (middle panels) while the autofluorescence of chlorophyll was collected from 600 680nm (lower panels) in cotyledons using CLSM. Merged images of GFP and chlorophyll fluorescence are shown in the upper panels and GFP and chlorophyll fluorescence are respectively indicated by green and red. Scale bar = 20 µm **B** Immunoblot analysis of the CAO-GFP fusion protein in transgenic plants. Equal fresh weights (2.5 mg) of rosette leaves were homogenized with the protein extraction buffer (see Materials and Methods), and subjected to SDS-PAGE. The CAO-GFP fusion or CAO proteins were detected using anti-GFP antibody. Black arrowheads indicate predicted molecular sizes of CAO-fusion proteins (GFP-C; 65kD, GFP-BC; 68kD) respectively. Grey arrowheads indicate the predicted molecular size of full-length CAO (56kD).

Figure.5 A & B Comparison of protein levels in transgenic plants of A-GFP with AB-GFP.

A Observations of CAO-GFP fusion protein levels in transgenic plants. GFP fluorescence was collected from 500 550 nm (middle panels) while the autofluorescence of chlorophyll was collected from 600 680 nm (lower panels) in

cotyledons using CLSM. The merged images of GFP and chlorophyll fluorescence were shown in the upper panels. GFP fluorescence and chlorophyll fluorescence are respectively indicated by green and red. Scale bar = 20 μm **B** Immunoblot Analysis of the CAO-GFP fusion protein in transgenic plants. Equal fresh weights (2.5 mg) of rosette leaves were homogenized with the protein extraction buffer (see Materials and Methods), and subjected to SDS-PAGE. and the CAO-GFP fusion or CAO proteins were detected using anti-GFP antibody. Black arrowheads indicate the predicted molecular sizes of CAO-fusion proteins (A-GFP; 42kD, AB-GFP; 45kD) respectively. Grey arrowheads indicate predicted molecular size of full-length CAO (56kD).

Figure.6 A & B Comparison of protein levels in transgenic plants of B-GFP with GFP.

A Observations of CAO-GFP fusion protein levels in transgenic plants. GFP fluorescence was collected from 500 - 550 nm (middle panels) while the autofluorescence of chlorophyll was collected from 600 – 680 nm (lower panels) in cotyledons using CLSM. The merged images of GFP and chlorophyll fluorescence were shown in the upper panels. GFP fluorescence and chlorophyll fluorescence are indicated by green and red respectively. Scale bar = 20 μm **B** Immunoblot analysis of the CAO-GFP fusion protein in transgenic plants. Equal fresh weights (2.5 mg) of rosette leaves were homogenized with the protein extraction buffer (see Materials and Methods), and subjected to SDS-PAGE. and the CAO-GFP fusion or CAO proteins were detected using anti-GFP antibody. Black arrowheads indicate the predicted molecular size of CAO-fusion proteins (B-GFP; 30kD, GFP; 27kD) respectively. Grey arrowheads indicate the predicted molecular size of full-length CAO (56kD).

Table .1 Chlotrophyll accumulation in the transformants

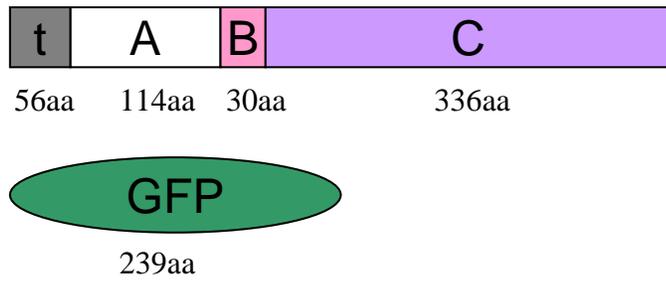
The names of the transgenic plants are described in the legend to Figure 2. na, not analyzed.

Figure 1

	transit peptide	A-domain	B-domain	C-domain	
A. thaliana.seq	1 MNAAVFSPSALS L PIS FSKTRSSF LSRKGV - K - G E F R V F A V F G D E S G L V E K K S Q M R P F D V E D P R S K A P P Y K G K F L D V N Q A I E V A R F D I Q Y L D W R A R Q D L L T I M I L H D K V V D V L N P L A R				118
O. sativa.seq	1 ----MTTVASLS L L P H L L I K P S F R C C S R K G V G R Y G G I K V Y A V L G D D G A D Y A K N N A E A L F H W D D P G P R V P I A K G K F L D V N Q A L E V V R F D I Q Y C D W R A R Q D L L T I M V L H N K V V E V L N P L A R				116
L. esculentum.seq	1 ----ATATAIS L S L P F S L C R S N K S C T R K F V - K - G S F G V F A V Y G E E G G I P D K K S W L T L F N V E D P R T K V P Q S K G K F L D A N Q A L E V A R Y D L Q Y C D W R A R Q D V L T I M L L H E K V V E V L N P L A R				113
P. trichocarpa.seq	1 -----LNFYGLNGD L C V N S Q G V - R - G F G V F A V L G E E G G L D K K S T M G P L F D V E D P R S K M P Q F K G K F L D F Y Q A L E V A R Y D I Q Y C D W R A R Q D L L T I M I L H E K V V E V L N P L A R				104
A. thaliana.seq	119 EYKSIGTVMKELAGLQELSKAHQQVHISEARVSTALDKLAHMEELVNDRLLPGRVVTLDKPPSSITTAFAVELDRE-KTNTGAKSLVMSGVPPYPSPHLKNFWYPVAFTAADLKHDTMVP				237
O. sativa.seq	117 EFKSIGTLRKELAELOEELAKAHNQVHISETRVSSALDKLAQMETLVNDRLLDGGSSASTAECTSLAPSTSSASRVVNKKPPFRSLVMSGVPPYPNPSLKNFWYPVAFSSDLKDDTMVP				236
L. esculentum.seq	114 DYKSIGTMKELAELOEELSAHQVHISEARVSAALDKLAYMEELVNDRLDQERSTVSECTSSASTSTIGLIDTP-KSKQPRTLVMSGVPPQYSSRLKNFWYPVAFSADLKNDTMLP				232
P. trichocarpa.seq	105 DYKSIGTMKELAELODELQAH--RHISEARVSTALDKLAYMEELVNDRLLDQRNPAESDQASPSSTSTOSLDTV-KNK--SKNLKMSGVPPQYHPHLKNFWYPVAFSTDLDKDGTMVP				219
A. thaliana.seq	238 T E C F E Q P W V I F R G E D G K P G C V R N T C A H R A C P L D L G T V N E G R I O C P Y H G W E Y S T D G E C K K M P S T K L L K V I K S L P C L E Q E G M I W I W P G D E P P A P I L P S L Q P P S G F L I H A E - L V M D L P V E H G				356
O. sativa.seq	237 I D C F E Q Q W V I F R G K D G R P G C V R N T C A H R A C P L H L G S V N E G R I O C P Y H G W E Y S T D G K C E K M P S T K M L N V R I R S L P C F E Q E G M V W I W P G N D P P K S T I P S L L P P S G F T I H A E - I V M E L P V E H G				355
L. esculentum.seq	233 I D C F E Q S W V I F R G A D G K P G C V R N T C A H R A C P L D L G S V N E G R I O C P Y H G W E Y S T D G K C E K M P S T R L L N V K I K A L P C F E Q E G M I W I W P G N D P P A T L P S L L P P S G F Q I H A E - I V M E L P V E H G				351
P. trichocarpa.seq	220 I D C F E P W V I F R G M D G K P G C V R N T C A H R A C P L H L G S V N E G R V R C P Y H G W E Y S T D G K C E K M P S T R L L N V K I K S L P C F E Q E G M L V W P G S D P P A A T L P S L Q P P P G F Q V H A E V I V M E L P V E H G				339
A. thaliana.seq	357 L L L D N L L D L A H A P F T H T S T F A K G W S V P S L V K F L T P S G L Q G Y W D P Y P I D M E F R P P C I V L S T I G I S K P G K L E G K S T Q Q C A T H L H Q L H V C L P S S R N K T R L L Y R M S L D F A P I L K N L P F M E H L W				476
O. sativa.seq	356 L L L D N L L D L A H A P F T H T S T F A K G W S V P S L V K F L T P S S G L Q G Y W D P Y P I D M E F R P P C M V L S T I G I S K P G K L E G K S T K Q C S T H L H Q L H I C L P S S R N K T R L L Y R M S L D F A P W I K H V P F M H I L W				475
L. esculentum.seq	352 L L L D N L L D L A H A P F T H T S T F A K G W S V P S L V K F L T P A S G L Q G Y W D P Y P I D M E F R P P C M V L S T I G I S K P G K L E G K S T K Q C S T H L H Q L H V C L P A S R Q K T R L L Y R M S L D F A P L K H I P F M Q Y V W				471
P. trichocarpa.seq	340 L L L D N L L D L A H A P F T H T S T F A K G W S V P R F V K F L T P A S G L Q G Y W D P Y P I D M E F R P P C M V L S T I G I S K P G K L E G Q S T R E C A T H L H Q L H V C L P S S R H K T R L L Y R M S L D F A P V L K H V P F M H Y L W				459
A. thaliana.seq	477 R H F A E Q V L N E D L R L V L G Q Q E R M I N G A N I W N L P V A Y D K L G V R Y R L W R N A V D R G D D K L P F S G -----M S L D F A P I L K N L P F M E H L W R H F A E Q V L N E D L R L V L G Q Q E R M I N G A N I W N L P V A Y				590
O. sativa.seq	476 S H F A E Q V L N E D L R L V L G Q Q E R M I N G A N I W N V P V S Y D K L G I R Y R L W R D A I E R G V D R L P F S N Q S E S G S M S L D F A P W I K H V P F M H I W S H F A E Q V L N E D L R L V L G Q Q E R M I N G A N I W N V P V S Y				595
L. esculentum.seq	472 R H F A E Q V L N E D L R L V L G Q Q D R M I N G A N I W N L P V S Y D K L G V R Y R I W R D A V D S G E K E L P F S K -----M S L D F A P L L K H I P F M Q Y V W R H F A E Q V L N E D L R L V L G Q Q D R M I N G A N I W N L P V S Y				585
P. trichocarpa.seq	460 R H F A E Q V L N E D L R L V L G Q Q E R M I N G A N I W N V P V S Y D K L G V R Y R L W R D A V E K G A Q L P F -----M S L D F A P V L K H V P F M H Y L W R H F A E Q V L N E D L R L V L G Q Q E R M I N G A N I W N V P V S Y				571
A. thaliana.seq	591 D K L G V R Y R L W R N A V D R G D D K L P F S G -----				615
O. sativa.seq	596 D K L G I R Y R L W R D A I E R G V D R L P F S N Q S E S G S -----				626
L. esculentum.seq	586 D K L G V R Y R I W R D A V D S G E K E L P F S K -----				610
P. trichocarpa.seq	572 D K L G V R Y R L W R D A V E K G A Q L P F -----				594

Figure 2

A



B

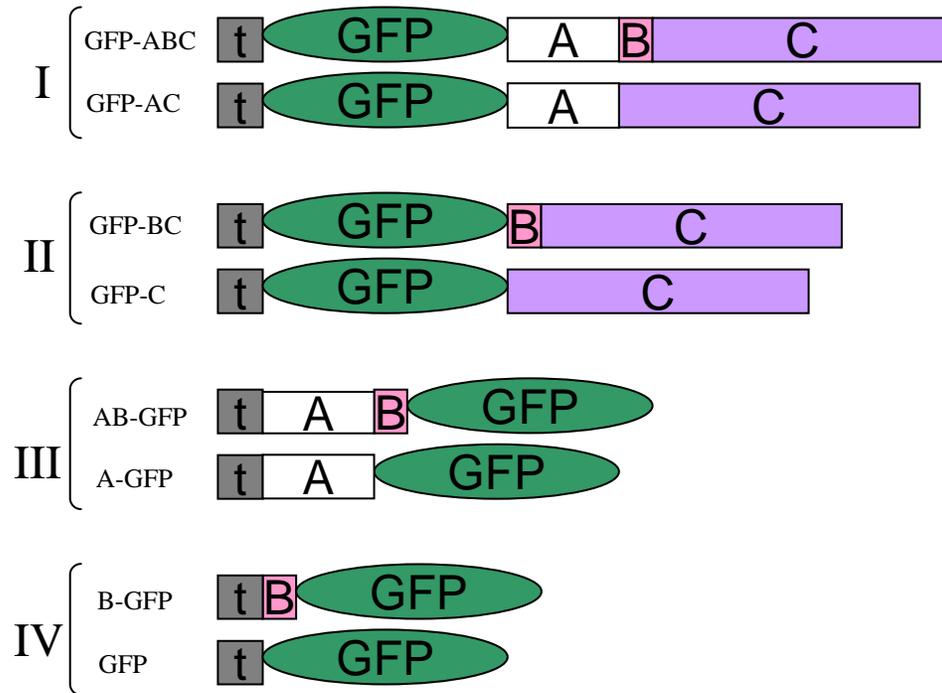


Table 1 Chlorophyll accumulation in the transformants

Arabidopsis transformants	Sample number	Chlorophyll a/b Ratio		Total Chlorophyll (nmole/mg Fresh Weight)	
		Avarage	SD	Average	SD
Colombia WT (nontransformant)	15	3.20	0.23	1.99	0.23
<i>chl-1</i> (nontransformant)	8	na	na	0.68	0.20
tp-GFP-ABC/ <i>chl-1</i>	10	3.16	0.16	2.01	0.23
tp-GFP-AC/ <i>chl-1</i>	9	3.11	0.22	1.95	0.22
tp-GFP-BC/ <i>chl-1</i>	10	2.27	0.32	2.16	0.43
tp-GFP-C/ <i>chl-1</i>	12	2.25	0.21	2.10	0.23
tp-AB-GFP/WT	8	3.13	0.23	2.00	0.11
tp-AB-GFP/ <i>chl-1</i>	13	na	na	0.74	0.33
tp-A-GFP/WT	8	3.18	0.32	2.01	0.23
tp-A-GFP/ <i>chl-1</i>	9	na	na	0.78	0.33
tp-B-GFP/WT	7	3.23	0.40	2.01	0.35
tp-GFP/WT	9	3.09	0.35	1.96	0.32

The names of transformants are described in Figure 4. na= not analyzed

Figure 3

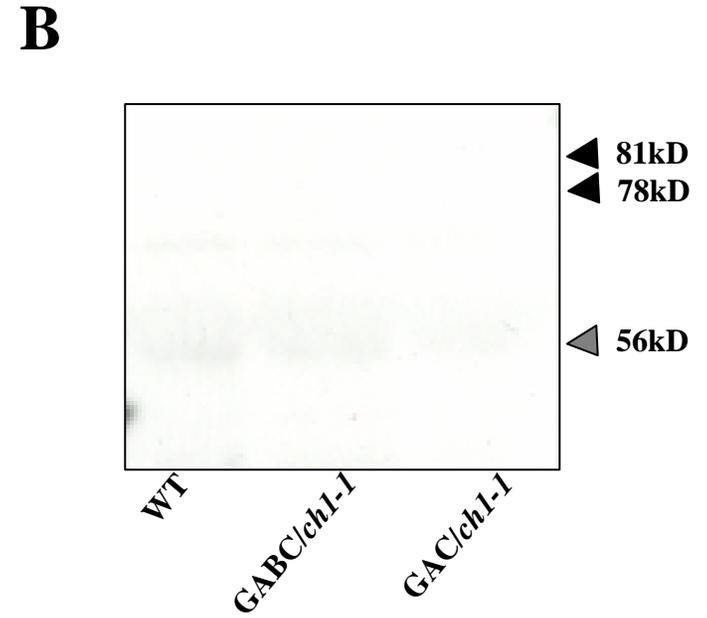
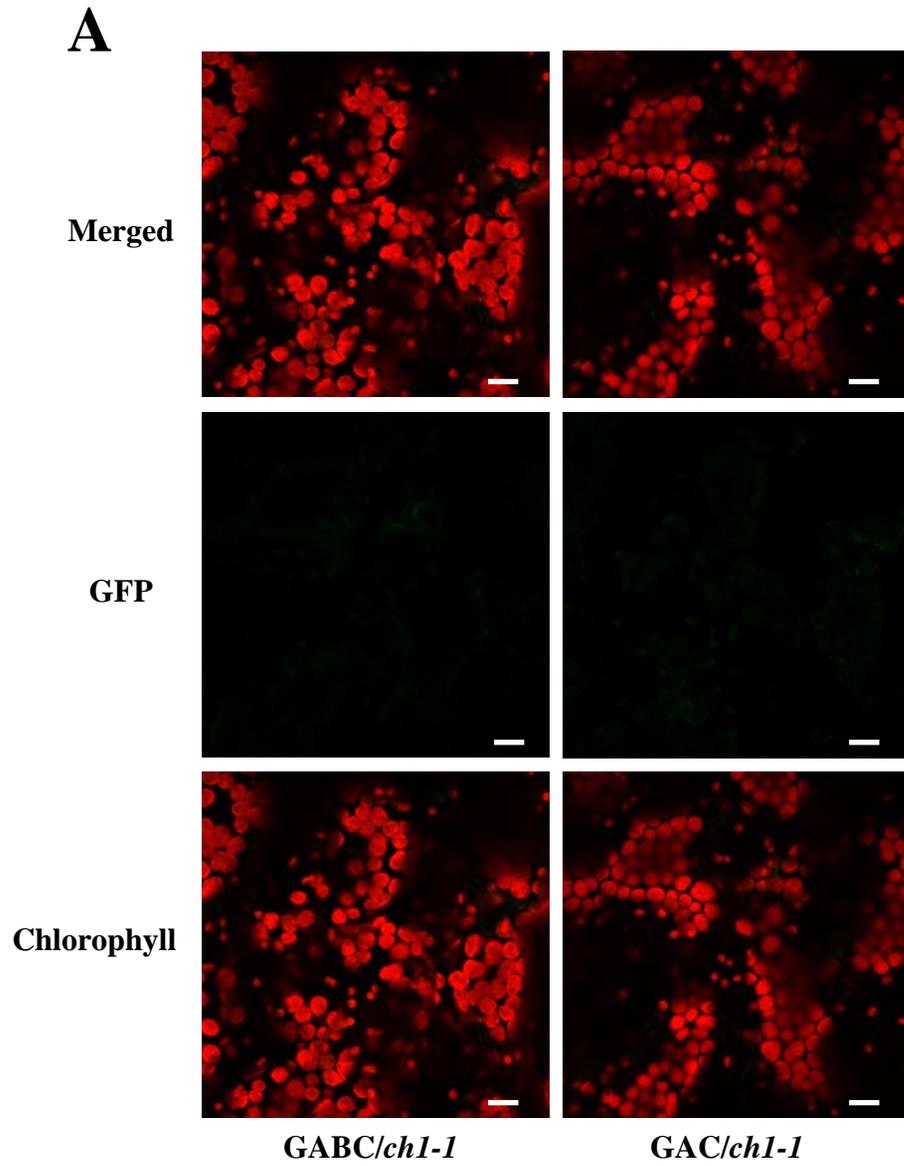
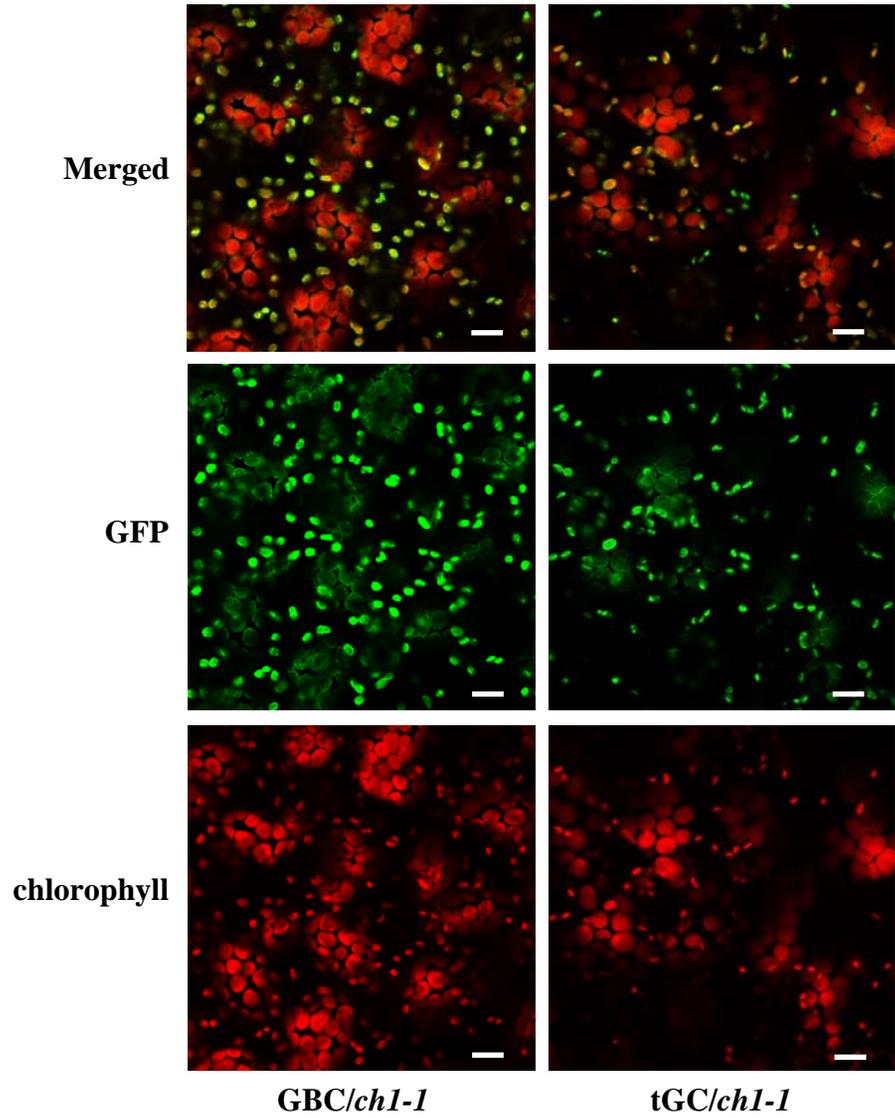


Figure 4

A



B

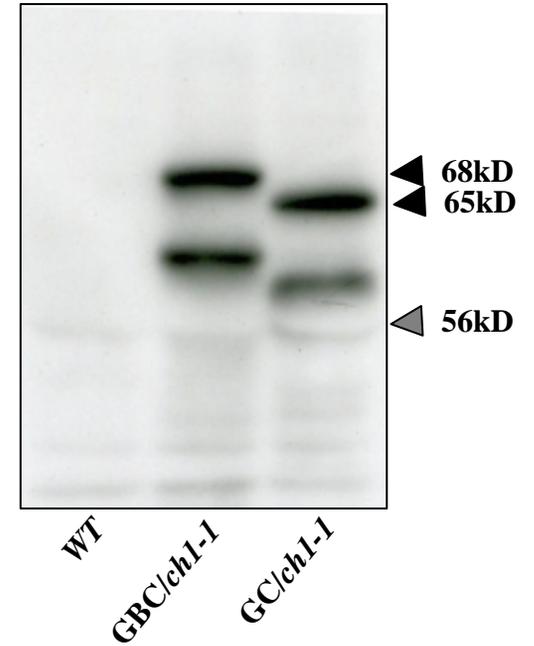


Figure 5

A

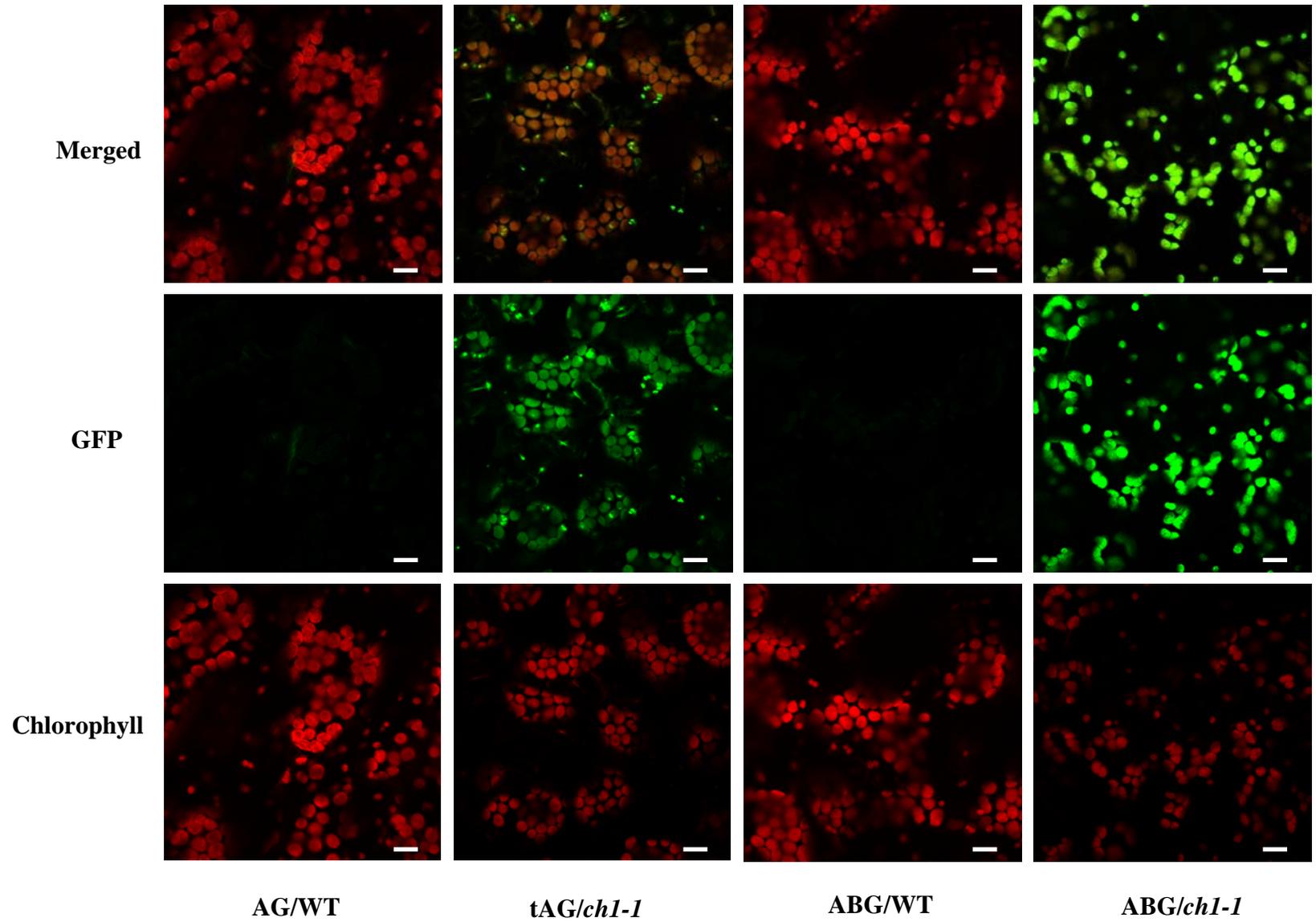


Figure 5

B

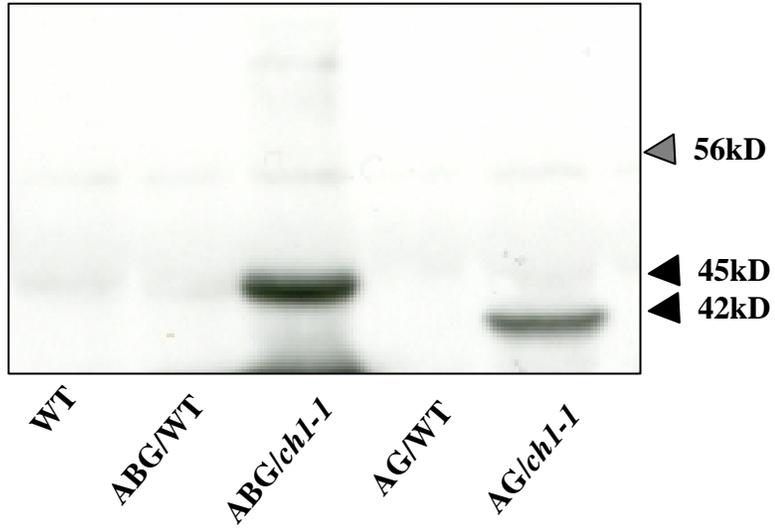
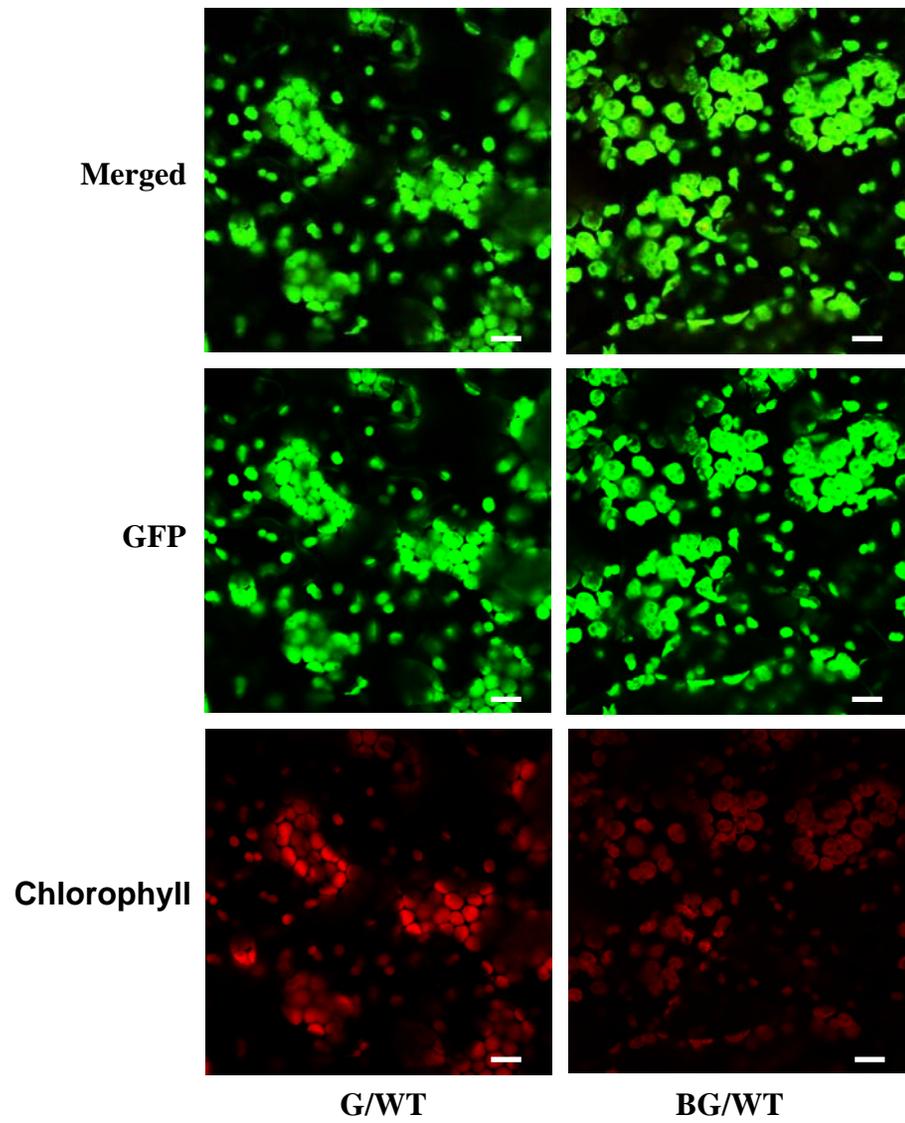


Figure 6

A



B

