

**Metabolic engineering of *Synechocystis*
sp. PCC6803 for plant type pigment
production, and identification of new
splicing factors in *Arabidopsis thaliana***



Dissertation der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

Vorgelegt von

Evgenia Vamvaka

München

18.04.2016

Metabolic engineering of *Synechocystis* sp. PCC6803 for
plant type pigment production, and identification of new
splicing factors in *Arabidopsis thaliana*

Zur Erlangung des Doktorgrades der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

Evgenia Vamvaka

Erstgutachter: Prof. Dr. Dario Leister

Zweitgutachter: Prof. Dr. Peter Geigenberger

Tag der Einreichung: 18.04.2016

Tag der mündlichen Prüfung: 23.06.2016

Summary:

Synechocystis sp. PCC6803 is a well studied cyanobacterium in regard to photosynthesis research. The fact that they can grow fast and are easily genetically modifiable make them ideal candidates for industrial applications. Lately, cyanobacteria are considered as promising cell factories for the generation of value-added products. Even though cyanobacteria are more efficient in photosynthesis than plants, their photosynthetic performance is limited. Therefore, different approaches to improve photosynthesis have been proposed like the introduction of light harvesting complexes (LHCs) from plants into *Synechocystis* or the extension of their absorption range. Three plant-specific xanthophylls are required for the correct assembly of LHCs in plants; lutein, violaxanthin and neoxanthin. In this study, genes for the production of lutein and violaxanthin were introduced into *Synechocystis*. For lutein synthesis, four *Arabidopsis* genes coding for enzymes of the lutein pathway were transformed in wildtype and into a *cruA* *Synechocystis* strain, which produces less β carotene. Lutein could only be detected in *cruA*, but not in the wild type, indicating that disruption of endogenous pigment synthesis pathways is a prerequisite for the production of lutein. For violaxanthin production, the *Arabidopsis* gene coding for the zeaxanthin epoxidase, which converts zeaxanthin to violaxanthin, was expressed in *Synechocystis*. Even though transgenic *Synechocystis* cells showed an altered carotenoid synthesis profile, violaxanthin could not be detected.

Many nuclear encoded factors are responsible for organelle gene expression in *Arabidopsis thaliana*. RNA splicing in plastids requires the action of proteinous factors and many of them have been recently identified and functionally characterized. In this study, a protein family, which consists of two proteins and which is present in the green lineage, was identified as a new component involved in organelle gene expression. The GreenCut proteins Gc16A and Gc16B are both targeted to plastids and are specifically involved in the biogenesis of the NADH dehydrogenase complex (NDH) and the Cyt b_6f complex. Lack of both proteins leads to a complete loss of the PSI-NDH- and to lowered accumulation of the Cyt b_6f -complex, which could be demonstrated by BN-PAGE, immunoblot and chlorophyll a fluorescence measurements. Transcript analyses on plastid-encoded subunits of the NDH and the Cyt b_6f complexes indicated a specific effect of Gc16 on *ndhA*, *petB* and *petD* transcript maturation. Furthermore, the

ability to bind RNA, which could be shown by electrophoretic mobility shift assay (EMSA) analyses, indicated that Gc16 is a novel factor involved in splicing of premature *ndhA*, *petB* and *petD* transcripts in *Arabidopsis*.

Zusammenfassung:

Synechocystis sp. PCC6803 ist in Bezug auf die Photosyntheforschung ein gut untersuchtes Cyanobakterium. Die Tatsache, dass Cyanobakterien schnell wachsen und leicht genetisch modifizierbar sind, macht sie zu idealen Kandidaten für eine industrielle Anwendung. Daher gelten Cyanobakterien als vielversprechende Zellfabriken zur Herstellung von Produkten mit hoher Wertschöpfung. Obwohl Cyanobakterien einen effizienteren Photosyntheseapparat gegenüber Pflanzen aufweisen, ist die Photokonversionseffizienz limitiert. Daher wurden unterschiedliche Strategien zur Verbesserung der Photosynthese vorgeschlagen, wie die Synthese von pflanzlichen Lichtsammelkomplexen (LHCs) oder die Erweiterung des Absorptionsbereichs der Photoynthesekomplexe. Für die korrekte Assemblierung von LHCs in Pflanzen sind drei pflanzenspezifische Xanthophylle erforderlich; Lutein, Violaxanthin und Neoxanthin. In dieser Studie wurden die Gene für die Herstellung von Lutein und Violaxanthin in *Synechocystis* eingebracht. Für die Luteinsynthese wurden vier *Arabidopsis*-Gene, die für Enzyme des Luteinsynthesewegs codieren, in den *Synechocystis*-Wildtyp und in die *cruA* Mutante, die weniger β-Carotin aber mehr Lycopen produziert, transformiert. Lutein konnte nicht im Wildtyp aber in *cruA* detektiert werden, was darauf hinweist, dass für die Luteinsynthese in *Synechocystis* eine Alterierung der endogenen Pigmentsynthesewege notwendig ist. Für die Violaxanthinproduktion wurde das *Arabidopsis* gen der Zeaxanthinepoxidase, die Zeaxanthin in Violaxanthin umwandelt, in *Synechocystis* heterolog exprimiert. Obwohl die transgenen *Synechocystis*-zellen ein verändertes Carotinoidsyntheseprofil aufwiesen, konnte Violaxanthin nicht über HPLC-Analysen detektiert werden.

Viele kernkodierte *Arabidopsis thaliana* Faktoren sind verantwortlich für die Genexpression in Organellen. RNA-Spleißen in Plastiden hängt dabei von Proteinen ab, von denen bereits einige identifiziert und funktionell charakterisiert

werden konnten. In dieser Studie wurde eine neue Proteinfamilie, die aus zwei Proteinen besteht und in der grünen Linie konserviert vorliegt, als neue Komponente in der plastidären Genexpression beschrieben. Die GreenCut Proteine Gc16A und Gc16B konnten im Chloroplasten nachgewiesen werden und sind spezifisch an der Biogenese des NDH- und des Cytb₆f Komplexes beteiligt. Die Abwesenheit von beiden Proteinen in einer generierten Doppelmutante führt zu einem vollständigen Verlust des PSI-NDH- und einer geringeren Akkumulation des Cytb₆f-Komplexes, was durch BN-PAGE-, Immunblot- und Chlorophyll a Fluoreszenzanalysen nachgewiesen werden konnte. Transkriptanalysen von plastidär kodierten Untereinheiten des NDH- und des Cytb₆f-Komplexes zeigten eine spezifische Wirkung von GC16 auf die Maturation von *ndhA*, *petB* und *petD* Transkripten. Darüber hinaus konnte über EMSA-Analysen nachgewiesen werden, dass Gc16B an RNA bindet. Daher ist GC16 ein neu identifizierter Faktor der für das korrekte Spleißen von *ndhA*, *petB* und *petD* notwendig ist.

Abbreviations:

amp	Ampicillin
<i>Arabidopsis</i>	<i>Arabidopsis thaliana</i>
ATP	Adenosine triphosphate
BN	Blue Native
bp	Base pair
BSA	Bovine serum albumin
CAI	Codon Adaptation Index
Car	Carotenoid
cDNA	Complementary deoxyribonucleic acid
CDS	Coding sequence
CET	Cyclic electron transport
Chl	Chlorophyll
Ci	Curie
CM	Cytoplasmic membrane
CO ₂	Carbon dioxide
Col-0	<i>Arabidopsis thaliana</i> Columbia-0
cTP	Chloroplast transit peptide
Cytb ₆ f	Cytochrome b ₆ f
Da	Dalton
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleotides
DTT	Dithiothreitol
<i>E.coli</i>	<i>Escherichia coli</i>
EDTA	Ethylene diamine tetraacetic acid
ETR	Electron transport rate
<i>g</i>	Gravity
g	Gram
GFP	Green Fluorescent Protein
h	Hour
HPLC	High Performance Liquid Chromatography
kan	Kanamycin
l	Litre
LB	Lysogeny broth

LET	Linear electron transport
LHC	Light harvesting complex
m	Metre
M	Molar
min	Min
ml	Millilitre
mM	Millimolar
mol	Molar
MOPS	3-(N-morpholino)propanesulfonic acid
NADPH	Nicotinamide adenine dinucleotide phosphate
NDH	NADPH dehydrogenase complex
nm	Nanometre
NPQ	Non-photochemical quenching
°C	Degree Celsius
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PAM	Pulse amplitude modulation
PCR	Polymerase chain reaction
PGR5	PROTON GRADIENT REGULATION 5
PGRL1	PROTON GRADIENT REGULATION 5-LIKE 1
PSI/II	Photosystem I/II
PVDF	Polyvinylidene difluoride
qRT-PCR	Quantitative Real Time PCR
rmp	Rounds per minute
RNA	Ribonucleic acid
SDS	Sodium dodecyl sulphate
sec	Second
Sm	Spectinomycin
suc	Sucrose
<i>Synechocystis</i>	<i>Synechocystis sp.</i> PCC6803
Taq	DNA Polymerase
T-DNA	Transfer DNA
TM	Thylakoid membrane
Tris	Tris (hydroxymethyl) aminomethane
UTR	Untranslated region
v	Volume

VAZ Xanthophyll cycle pigments
(Violaxanthin, antheraxanthin, zeaxanthin)
β-DM β-dodecylmaltoside
μ Micro

Table of contents:

1. Introduction	1
1.1. Expression of exogenous pigments in <i>Synechocystis</i> : Cyanobacteria as model organisms and cell factories	1
1.2. Improving photosynthesis in cyanobacteria	2
1.3. Biosynthetic pathway of carotenoids in cyanobacteria and plants	3
1.4. Carotenoid biosynthesis in <i>Synechocystis</i>	5
1.5. Carotenoid biosynthesis in <i>Arabidopsis thaliana</i>	6
1.6. Aim of <i>Synechocystis</i> project	9
1.7. Introduction of Gc16: Chloroplast Gene Expression	9
1.8. RNA splicing of introns in chloroplasts	10
1.9. Splicing factors in chloroplasts of land plants	11
1.10. NDH- and Cyt b_6f -complexes	12
1.11. Aim of Gc16 project	14
2. Material and Methods	15
2.1. Database analysis and software tools	15
2.2. Chemical materials	15
2.3. Enzymes, kits, primers and radiochemicals	15
2.4. Bacterial cell cultures and growth conditions	15
2.5. Construction of plasmids for <i>Synechocystis</i> transformation	18
2.5.1. Construction of pKS-LUT	18
2.5.2. Construction of pΔcruA:Spec and pΔcruA:5'Spec	19
2.5.3. Construction of pKS-ZEP	19
2.6. <i>Synechocystis</i> transformation	23
2.7. Nucleic acid isolation:	24
2.7.1. PCR (standard and high fidelity)	24
2.7.2. Genomic DNA isolation from <i>Synechocystis</i>	24

2.7.3. RNA isolation	25
2.8. cDNA synthesis	25
2.9. Northern blot analysis	25
2.10. Radioactive probe preparation	27
2.11. Spectroscopic analysis	27
2.12. Pigment extraction and HPLC analysis	27
2.13. Antibodies	28
2.14. Plant material and growth conditions	28
2.15. Complementation of mutant lines	29
2.16. Bioinformatic sources	29
2.17. Bacterial strains-vectors	30
2.18. DNA extraction	30
2.19. RNA extraction and Northern blot analysis	30
2.20. cDNA synthesis and RT-PCR analysis	31
2.21. Real Time PCR	31
2.22. Pigment analysis	31
2.23. Protoplast isolation and GFP detection	32
2.24. TEM analysis	32
2.25. Chlorophyll fluorescence measurements- NDH activity	32
2.26. Generation of Gc16 antibody	33
2.27. Total protein isolation	33
2.28. Thylakoid isolation	33
2.29. Immunoblot analysis	34
2.30. Blue Native gel analysis and second dimension gel	34
2.31. In vivo labeling assay with ^{35}S -Methionine	35
2.32. EMSA	35
2.33. Primer list	35

3. Results	38
3.1. Lutein Production	38
3.1.1. Generation of Lutein construct	38
3.1.2. Color complementation assay in <i>E.coli</i>	39
3.1.3. Generation of <i>Synechocystis</i> mutants producing Lutein	41
3.1.4. Pigment analysis of <i>SynLUT</i> mutants	42
3.1.5. Deletion of the putative β-cyclase (<i>sll0147</i>) from <i>Synechocystis</i>	44
3.1.6. Characterization of the mutants: Chromosomal segregation	45
3.1.7. Pigment analysis of Δ <i>cruA</i>	47
3.1.8. Generation of <i>Synechocystis</i> mutants combining lutein pathway introduction and <i>cruA</i> disruption.	48
3. 2. Violaxanthin Production	51
3.2.1. Introduction of Zeaxanthin Epoxidase from <i>Arabidopsis</i> into <i>Synechocystis</i>	51
3.2.2. Genetic characterization of <i>SynZEP</i> mutants	52
3.2.3. Pigment profile of the mutants	54
3. 3. Gc16 characterization	57
3.3.1. Greencut proteins	57
3.3.2. Mutant analysis of <i>gc16a</i> and <i>gc16b</i>	60
3.3.3. Complementation of Gc16 function	62
3.3.4. Subcellular and suborganellar localization of Gc16A and Gc16B	63
3.3.5. Pigment composition in <i>gc16</i> mutants	65
3.3.6. Gc16A and B are required in NDH and Cytb6f biogenesis	66
3.3.7. Thylakoid composition and translation rates in <i>gc16</i> mutants	67
3.3.8. Chloroplast gene expression in <i>gc16</i> mutants	69
3.3.9. NDH activity and photosynthetic parameters in <i>gc16</i> mutants	71
3.3.10. Gc16 is required for proper transcription in <i>ndhH-D</i> operon in the chloroplasts	74

3.3.11. Gc16A/B associates with group II introns in the chloroplasts	76
3.3.12. Gc16B is an RNA-binding protein	78
4. Discussion	80
4.1. Discussion <i>Synechocystis</i>	80
4.1.2. Synechocystis as a cell factory for lutein production	81
4.1.3. Strategies to increase lutein production in Synechocystis	82
4.1.4. Violaxanthin production in Synechocystis	83
4.2. Discussion Gc16	84
4.2.1. RNA splicing mechanism in plants	84
4.2.2. Defects in NDH and Cyt b_6f complexes in the absence of Gc16 proteins	85
4.2.3. Gc16 - a putative RNA splicing factor?	86
REFERENCES	88

1. Introduction

1.1. Expression of exogenous pigments in *Synechocystis*: Cyanobacteria as model organisms and cell factories

Cyanobacteria are the only organisms among prokaryotes that perform oxygenic photosynthesis. They can absorb sunlight and convert CO₂ into other products of interest, like natural compounds or chemicals. They also play an important role in ecology, since they can grow in a wide range of environmental conditions. Cyanobacteria (previously named as blue-green algae) are considered to be the evolutionary ancestors of chloroplasts, having a photosynthetic apparatus similar to that of higher plants. They belong to the Gram-negative bacteria and lack organelles but do have a cell wall (Rippka & Deruelles 1979).

Synechocystis sp. PCC6803 (hereafter *Synechocystis*) is one of the most extensively studied species among cyanobacteria. It is a unicellular, non-nitrogen fixing organism that belongs to the group of globular shaped Chroococcales. *Synechocystis* is a well established model organism, since its genome is fully sequenced (3,6 Mbp) (Kaneko & Tabata 1997). Their ability to naturally take up foreign DNA and integrate it into their genome via double homologous recombination is one of the main advantages working with this organism. Furthermore, their ability to grow from photoautotrophic to completely heterotrophic conditions in the presence of glucose in the dark makes them an ideal model organism to study photosynthesis (Williams 1988).

Recently, *Synechocystis* has been used in many studies as a platform for the production of a wide range of products, such as biofuels, chemicals, or other natural compounds (Lindberg et al. 2010a; Melis 2009; Bentley et al. 2014; Melis 2012). In contrast to heterotrophic bacteria, *Synechocystis* can be regarded as a photocatalyst, which captures light energy and converts it into bioproducts (Oliver & Atsumi 2014). Metabolic engineering has been applied to *Synechocystis* for many pathways using molecular tools developed for synthetic biology approaches.

1.2. Improving photosynthesis in cyanobacteria

Even though algae and cyanobacteria convert 2-3 fold more efficiently solar energy to biomass than plants, biomass yields are still low (Blankenship et al. 2011; Melis 2009). The efficiency of oxygenic photosynthesis has been proposed to be around 8-10% and drops to 2% under low light conditions (Melis 2012; Melis 2009). Nowadays, it is important to improve photosynthetic efficiency, so that solar irradiation can be better utilized to produce biofuels and value-added, chemical compounds. Photosynthetic microorganisms are ideal for this purpose, since they do not compete with traditional crops for land and can also be grown under controlled laboratory conditions (Rittmann 2008).

One drawback of photosynthetic organisms is that the photosynthesis rate is limited by photoinhibitory mechanisms, also defined as non-photochemical quenching mechanisms (NPQ). Excess, absorbed light is not used for photochemistry, but is wastefully dissipated as heat. Additionally, cyanobacteria cannot absorb all the sunlight, because of the limited absorption spectrum of their photosynthetic antennas. For light harvesting, they use phycobilisomes; large pigment-protein complexes that harbor phycobilins with a spectral range from 500 to 650 nm. Phycobilisomes in the cyanobacterium *Synechocystis* consist of a core with allophycocyanin (APC) pigments and six rods with phycocyanin (PC) (Arteni et al. 2009). They are attached to the cytoplasmic surface of the thylakoid membrane and primarily associate with photosystem II (MacColl 1998).

In order to overcome these limitations several ideas have been proposed, three of them are outlined in the following. The first one is to minimize the antenna complexes in photosynthetic organisms for better light penetration into deeper layers of photobioreactors. So far, this approach has been successful only for *Chlamydomonas*, where antenna truncation led to a threefold increase in efficiency (Beckmann et al. 2009). However, in *Synechocystis* the truncation of phycobilisomes (Beckmann et al. 2009) was followed by a decreased cell productivity due to the reduction of the PSI:PSII ratio (Collins et al. 2012; Page et al. 2012).

Another strategy for increasing photoconversion efficiency is to extend the light absorption range of one of the two photosystems up to ~1100 nm. This could be achieved by expressing bacteriochlorophyll *b* in cyanobacteria and algae (Blankenship et al. 2011). The third strategy is to introduce the light harvesting complexes from higher plants into cyanobacteria in order to extend the absorption range. This has been attempted in a study by He et al. (1999), where the *LHCB* gene from pea was heterologously expressed in a *Synechocystis* mutant producing also Chlorophyll *b* (Xu et al. 2001). Even though the gene was successfully expressed, the gene product was rapidly degraded and did not accumulate in thylakoid membranes, unless plant pigments and xanthophylls were added externally (Figure 1) (He et al. 1999). Hence, the authors concluded that successful LHC expression in *Synechocystis* depends on the presence of xanthophylls.

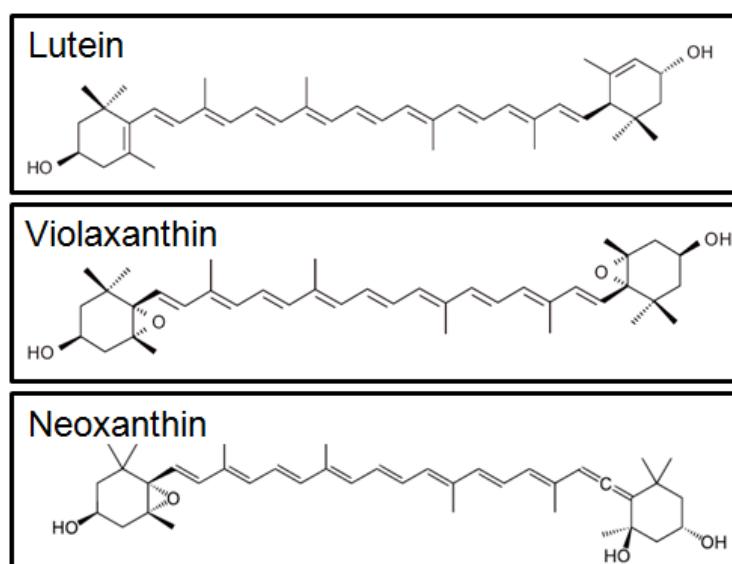


Figure 1. Xanthophylls produced in *Arabidopsis*, important for LHCs stabilization, that are not synthesized in *Synechocystis*.

1.3. Biosynthetic pathway of carotenoids in cyanobacteria and plants

Carotenoids are naturally occurring pigments and are present in all photosynthetic organisms and in some non-photosynthetic bacteria and fungi (Goodwin, 1980). They are isoprenoids and consist of polyene chains with up to 15 conjugated double bonds. Their main function is to act as light-harvesting pigments and to protect the organisms from photooxidation by scavenging singlet oxygen. Furthermore, they are essential components of

the photosynthetic complexes, stabilizing their structure (Jahns & Holzwarth 2011).

Carotenoids are divided into two groups; carotenes and xanthophylls, the latter to be the derivatives (epoxy, hydroxy, keto) of carotenes. Carotenes but not xanthophylls are present in all organisms. Cyanobacteria (including *Synechocystis*) lack the plant pigments lutein, violaxanthin and neoxanthin but produce the cyanobacterial-specific xanthophylls myxoxanthophyll and echinenone (Demmig-Adams & Adams 1992). In plants, carotenoid biosynthesis takes place in plastids and depends on low-abundant enzymes, which are nuclear-encoded.

In cyanobacteria and plants, the first steps of this pathway are similar and lead to the production of isopentenyl diphosphate (IPP), a C₅ compound. Pyruvate and glyceraldehyde 3-phosphate (G3P) react to IPP via the non-mevalonate pathway. Carotenoids are synthesized by several condensation steps that lead to the production of GGPP (geranylgeranyl diphosphate) (Cunningham & Gantt 1998). This is the precursor of carotenoids but also of many other compounds, like gibberellins, chlorophylls and tocopherols. The first specific step of the carotenoid pathway is the synthesis of the C₄₀ compound phytoene, which is catalyzed by phytoene synthase in a condensation reaction of two GGPP molecules. This is considered to be the rate-limiting step in the carotenoid synthesis pathway. Phytoene is a colorless carotene, its hydrocarbon chain harbors only three conjugated double bonds and absorbs only ultraviolet light (Armstrong 1997; Liang et al. 2006). After phytoene synthesis, the number of conjugated double bonds is increased by four desaturation steps, which ends up in the production of ζ -carotene. ζ -carotene is then converted into lycopene - a pink colored carotenoid - by crtQ in cyanobacteria and zds (zeta carotene desaturase) in plants. In principle, photosynthetic organisms share the same lycopene synthesis pathway, but subsequent cyclization steps differ between cyanobacteria and green plants which leads to species-specific carotenoid classes (Sandmann 1994; Ruiz-Sola and Rodriguez-Concepcion 2012).

1.4. Carotenoid biosynthesis in *Synechocystis*

Carotene production in *Synechocystis* starts with lycopene cyclization (Figure 2). The pathway is not completely understood and it is presumed that the lycopene β -cyclase in *Synechocystis* is encoded by *sll0147* (*cruA*) (Maresca et al., 2007). It produces sequentially a β -ionone ring in both ends of lycopene. In case one β -ring is formed, γ -carotene is produced, while the formation of two β -rings leads to the production of β -carotene, which is one of the major carotenoids in *Synechocystis*. These two carotenes are the substrate for the production of xanthophylls.

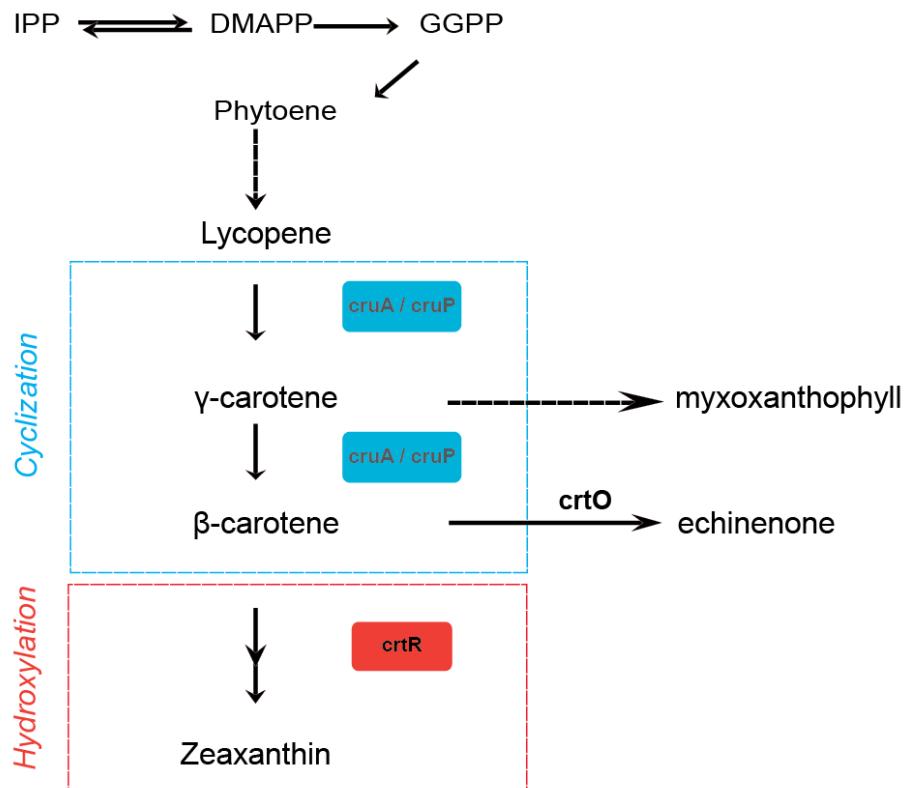


Figure 2. Biosynthetic pathway of carotenoids in *Synechocystis* in a simplified way. *cruA*: lycopene cyclase, *cruP*: lycopene cyclase paralao, *crtR*: β -carotene hydroxylase, *crtO*: β carotene ketolase.

The unique xanthophyll glycoside in *Synechocystis*, myoxanthophyll, is produced from γ -carotene and it is only present in cyanobacteria and some non-photosynthetic bacteria (Goodwin, 1980). The pathway is not fully identified yet, but for the formation of myoxanthophyll, γ -carotene hydroxylation of one β -ring and the addition of a unique glycoside linkage are

required. Myxoxanthophyll has been shown to be essential for the stabilization of thylakoid membranes and cell wall structure (Mohamed et al., 2005). β -carotene is also used as a precursor for the synthesis of the other two major xanthophylls echinenone and zeaxanthin. Echinenone is synthesized by a ketolase (encoded by *crtO*), adding a keto group on only one of the β -rings of β -carotene. Echinenone is bound to cytochrome b_6/f , acts as a structural component and is localized in the cytoplasmic membrane (CM) (Boronowsky et al. 2001). Zeaxanthin is formed from the hydroxylation of both β -rings at the C₃ position, via β -cryptoxanthin through the action of the enzyme β -carotene hydroxylase (*crtR*) (Takaichi et al. 2001; Takaichi and Mochimaru 2007). Zeaxanthin is localized in the thylakoid membrane (TM) of the cells and is an important component in a protective mechanism against photooxidation. Lack of zeaxanthin was shown to increase the susceptibility to reactive oxygen species (Schäfer et al. 2005).

1.5. Carotenoid biosynthesis in *Arabidopsis thaliana*

In higher plants, the carotenoid pathway downstream of lycopene differs from that of cyanobacteria. Lycopene cyclization is the branch point for several pathways, generating diverse carotenoids with two different cyclic groups; the β,β - and the β,ϵ -carotenoids. When acting alone, β -cyclase (*LCYb*) can either produce γ -carotene (β,ψ -carotenoid) or β -carotene (β,β -carotenoid). Whereas the production of δ -carotene and α -carotene depends on the action of the β - (*LCYb*) as well as the ϵ -cyclase (*LCYe*). The ϵ -cyclase can add an ϵ -ring in one end of lycopene, producing δ -carotene (ϵ,ψ -carotenoid).

The difference of β - and ϵ -rings is the position of the double bond in the cyclohexane ring, affecting the conformation of the ring. β -rings have the double bond in conjugation with the polyene chain, so that they have only one conformation. However, the ϵ -ring can rotate freely around its C_{6'}-C_{7'} bond. ϵ -rings are formed only in green plants, red algae and *Prochlorococcus*. Lycopene ϵ -cyclase (*LCYe*) is encoded by a single gene; the gene product is homologous to the β -cyclase and is a monoyclase. In contrast to *LCYb*, *LCYe* can only add one ϵ -ring on the chain of lycopene and then *LCYb* can take over and add the β -ring on the other end, converting it to α -carotene.

LCYe is the key enzyme for the production of β - and ϵ -carotenoids of higher plants (Cunningham et al., 1996).

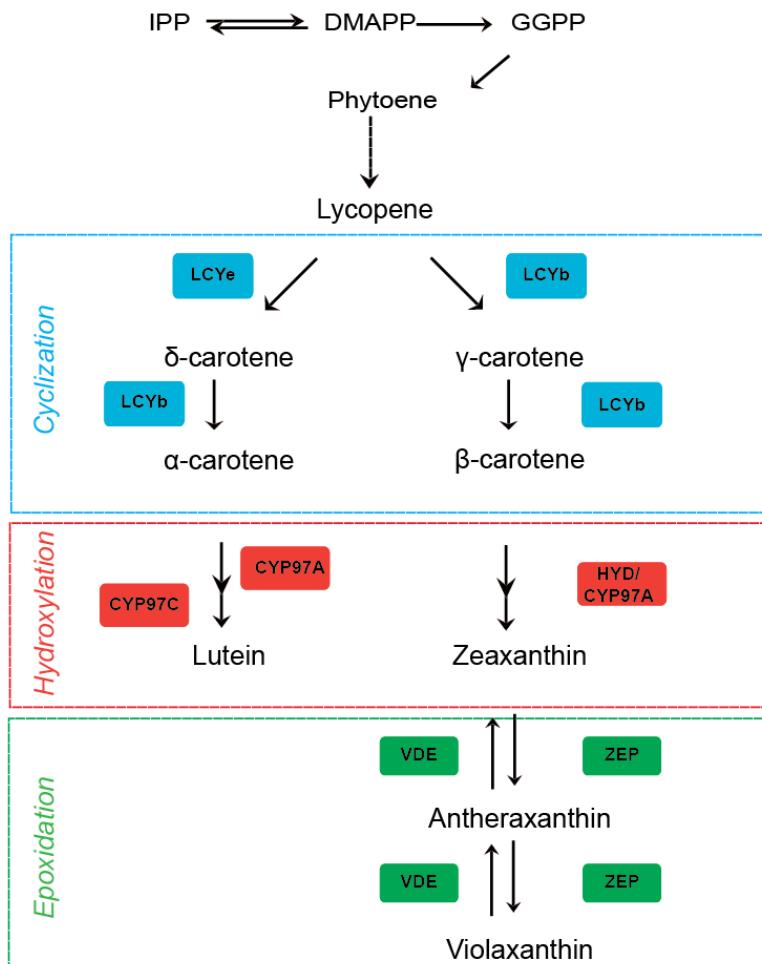


Figure 3. Carotenoid biosynthesis pathway in *Arabidopsis thaliana*.

LCYe: ϵ -cyclase, *LCYb*: β -cyclase, *CYP97A*: β -hydroxylase, *CYP97C*: ϵ -hydroxylase, *HYD*: β -hydroxylase, *ZEP*: zeaxanthin epoxidase, *VDE*: Violaxanthin de-epoxidase.

The next two steps of the pathway are hydroxylation reactions, using the carotenes as precursors for the formation of xanthophylls, lutein and zeaxanthin. Lutein is the most abundant carotene in photosynthetic plant tissues (>50% of total carotenoids) and it is conserved in all green plants due to the important role in LHC assembly. Two hydroxylases, a β -hydroxylase and an ϵ -hydroxylase, add a hydroxyl group at the C₃ position of each ring of α -carotene, catalyzing the formation of lutein. The production of zeaxanthin is catalyzed by a β -hydroxylase adding the hydroxyls in both ends of β -carotene (Figure 3) (Kim and DellaPenna 2006).

The enzymes responsible for this reaction are categorized in two distinct classes, the non-heme diiron enzymes (HYD) and the heme-containing cytochrome P450 (CYP97A3 and CYP97C1) (Tian and DellaPenna 2001; Sun et al. 1996). The first class of enzymes is involved in the hydroxylation of β -carotene only and they are dependent of ferredoxin, require iron and the transfer of a reducing equivalent from NADPH (Bouvier et al. 1998). The cytochrome P450 enzymes catalyze the hydroxylation of β - and ε -rings of α -carotene. These hemoproteins are momo-oxygenases and NADPH-dependent (Werck-Reichhart et al. 2002). All the enzymes used in this pathway are listed in Table 1.

The last part of this pathway involves the epoxidation of zeaxanthin to antheraxanthin and violaxanthin. This is part of the so called “xanthophyll cycle”, which includes also the reverse reaction and plays a central role in light adaptation of photosynthesis. This reaction is catalyzed by zeaxanthin epoxidase (*ZEP*), an enzyme that uses oxygen, ferredoxin, NADPH and FAD as cofactors (Bouvier et al. 1996). The optimal conditions for this enzyme is dark or low light and pH at 7,5 (Hager & Holocher 1994). The zeaxanthin epoxidase is localized at the stromal side of the thylakoids (Siefermann & Yamamoto 1975). The reverse reaction is catalyzed by vioalxanthin de-epoxidase (*VDE*), an enzyme responsible for the conversion of violaxanthin back to zeaxanthin, concomitant with the stepwise removal of epoxy-groups. Under high light conditions, the luminal proton concentration increases, creating a pH around 5, which is the optimum for their enzyme activity (Hager and Holocher 1994). Then, *VDE* binds to the thylakoid membranes and uses ascorbate for the reaction (Yamamoto et al. 1972). These two reactions occur in the opposite sites of thylakoids. Both of the enzymes belong to the plant lipocalin family, with a conserved structure, binding and transfer small hydrophobic molecules (Bugos et al. 1998).

Table 1. *Arabidopsis* genes involved in biosynthetic pathway of carotenoids.

These enzymes are involved in reactions downstream of lycopene synthesis and are not present in *Synechocystis*. The molecular mass of the proteins are indicated according to the Plant Proteome Database PPDB (<http://ppdb.tc.cornell.edu/>), the transmembrane domain are

predicted according to Aramemnon (<http://aramemnon.botanik.uni-koeln.de/>) and the chloroplast localization is depicted from Joyarg et al. (2009).

Accession	Gene	Function	Size (aa)		TM domain (Aramemnon)	Localization in Chloroplasts
			Full	cTP		
AT5G57030	LCY-e	Lycopene ε cyclase	524	45	v	nd
AT3G10230	LCY-b	Lycopene β cyclase	501	46	-	envelope
AT1G31800	CYP97A3	β-Hydroxylase	595	28	-	envelope
AT3G53130	CYP97C1	ε-Hydroxylase, LUT1	539	36	-	envelope
AT5G67030	ZEP	Zeaxanthin epoxidase, ABA1	667	59	-	envelope (75%), thylakoid (11%), stroma (14%)
AT1G08550	VDE	Violaxanthin de- epoxidase, NPQ1	462	82	v	thylakoid

1.6. Aim of *Synechocystis* project

The aim of this thesis was to introduce and express heterologously the genes that encode the enzymes responsible for the production of plant xanthophylls lutein and violaxanthin in *Synechocystis*. The resulting organism can be used as a platform, to introduce *lhcb* genes and to determine whether the synthesis of plant-type LHCs is more stable with the constant production of these pigments.

1.7. Introduction of Gc16: Chloroplast Gene Expression

Mitochondria and chloroplasts contain their own genome, which is a reminiscence of endosymbiotic events, where a cyanobacterial or an α-proteobacterial ancestor was engulfed by an eukaryotic host cell (Raven & Allen 2003; Andersson et al. 2003). During evolution, the organellar genomes shrank continuously, and most of the genetic information was transferred to the nucleus. Out of six possible ways to transfer DNA, from organelles to nucleus and vice versa, and between the plastids and mitochondria, five of them have been confirmed so far and only the transfer from nucleus to plastids has not been yet confirmed (Kleine et al. 2009). The most common event is the one from organelles to the nucleus.

The plastid genome of higher plants codes only for around 90 proteins, which are important components of the photosynthetic, the transcriptional and the

translational machinery (Timmis et al. 2004). Most of the proteins that have been transferred to the nucleus either are needed for chloroplast gene expression or have acquired new functions in different processes. Thus, mechanisms have to exist which guarantee a close communication between the compartments and which allow effective regulation of the gene expression of both genomes.

Many chloroplast genes are part of operons and are expressed as polycistronic transcripts. These transcripts can be processed through the action of nuclear-encoded proteins, acting on the transcriptional and the post transcriptional level resulting in the generation of multi-, di- or monocistronic transcripts (Barkan & Goldschmidt-Clermont 2000). Those factors are involved in RNA transcription, processing (5' and 3'), editing, splicing and translation. Transcription of chloroplast genes depends on nucleus encoded RNA Polymerase (NEP) and plastid encoded RNA Polymerase (PEP), (Hajdukiewicz et al. 1997). Then, after plastid differentiation, gene expression is primarily controlled by post transcriptional steps, like transcript processing or stabilization (Deng & Gruissem 1987).

1.8. RNA splicing of introns in chloroplasts

RNA splicing is part of the maturation of the transcripts, together with RNA processing and editing. During this event, the introns of the pre-mRNAs are removed and the exons are joined to form the mature mRNA. This process depends on nuclear-encoded proteins, which allow correct maturation of the transcripts. The introns in the chloroplasts are divided according to their conserved structure and splicing mechanisms into two groups (group I and group II). In *Arabidopsis thaliana*, there is only one intron belonging to group I (the *trnL* gene (UAA)) and 20 introns that belong to group II (De Longevialle et al. 2010). Group II is divided into four subclasses and only two subclasses (IIA and IIB) are present in land plants (Figure 4) (Michel & Ferat 1995). Group I is usually only present in eukaryotic organisms (plants, algae and fungi) while group IIC and IID are only present in bacteria (Haugen et al. 2005).

Group II introns are ribozymes and their catalytic mechanism is closely related to that of spliceosome in the nucleus. They consist of six domains and splicing

occurs in two stepwise trans-esterification reactions. It has been proposed that they act in folding the intron into the correct secondary structure to obtain a catalytic active ribozyme and also to bring the exons close to each other (Michel & Ferat 1995).

Group I	<i>trnL</i>
IIA	<i>atpF</i>
Group II	<i>rpl2</i>
IIB	<i>rps12-2</i>
	<i>trnK</i>
	<i>trnA</i>
	<i>trnV</i>
	<i>tml</i>
	<i>clpP-2</i>
	<i>ycf3-1</i>
	<i>ycf3-2</i>
	<i>rps12-1</i>
	<i>rpl16</i>
	<i>rpoC1</i>
	<i>trnG</i>
	<i>clpP-1</i>
	<i>ndhA</i>
	<i>ndhB</i>
	<i>petB</i>
	<i>petD</i>

Figure 4. Group I and group II introns in the chloroplast of *Arabidopsis thaliana* genes.

1.9. Splicing factors in chloroplasts of land plants

In bacteria, group II introns encode for a specific enzyme, a maturase that facilitates splicing of its own, coding sequence. However, the plastidial introns of vascular plants have lost their ability of self-splicing and only two maturases are present in the chloroplasts, MatK and Mat4, which are encoded by the plastid genome (Zoschke et al. 2010). The loss of these enzymes led to the recruitment of other proteins to act as splicing factors. However, these proteins are encoded in the nucleus, synthesized in the cytosol and transported into the chloroplasts (Vogel et al. 1999).

Chloroplast RNA splicing and ribosome maturation (CRM) domain proteins are RNA splicing factors, which are found in prokaryotes and plants and contain one or more RNA-binding domains (Till et al. 2001). The splicing factor CRS1 (Chloroplast RNA Splicing 1) is involved in *atpF* intron splicing and consists of three CRM domains (Jenkins & Barkan 2001; Till et al. 2001). CAF1 and CAF2 (CRS-2 associated factors 1 and 2) are members of this family and have two CRM domains, forming a complex with CRS2. The complexes CAF1-CRS2 and CAF2-CRS2 are involved in splicing of eleven group II introns in *Arabidopsis* (*petB*, *petD*, *ndhA*, *ndhB*, *rps12*, *ycf3*, *rpl16*, *rps16*, *trnG*, *clpP1* and *rpoC1*) (for overview Watkins et al. 2011) (Till et al. 2001; Asakura & Barkan 2006; Ostheimer et al. 2003). CRS2 belongs to another family that is involved in RNA splicing, the peptidyl-tRNA hydrolase (Till et al. 2001). Pentatricopeptide repeat proteins (PPR) are a group of

proteins with RNA binding ability. At least three of them (PPR4, PPR5 and OTP51) are present in the chloroplast and are involved in splicing (*rpS12-1*, *trnG* and *ycf3-2nd* intron, respectively) (Schmitz-Linneweber et al. 2006; De Longevialle et al. 2008). RNC1 is another splicing factor, which belongs to the ribonuclease II proteins, and forms a heterodimer with WTF1 (What's this factor 1), which contains a domain of unknown function (PORR-DUF860). It could be demonstrated that this heterodimer is involved in splicing of subgroup IIB introns found in *petB*, *ndhA* and *trnG* and all introns of subgroup IIA, except *clpP* (Watkins et al. 2007; Kroeger et al. 2009). Another factor, APO1, has been identified to bind RNA and mediates splicing of *petD*, *clpP-1* and *ycf3-2* introns (Watkins et al. 2011).

1.10. NDH- and Cytb_{6f}-complexes

The NADH dehydrogenase-like complex (NDH) in the chloroplasts mediates the cyclic electron transport (CET) around Photosystem I (PSI) and is also involved in chlororespiration. It reduces in a light-independent way the plastoquinone pool (Shikanai et al. 1998; Burrows et al. 1998). It has been shown that even the complete disruption of the NDH complex does not affect the overall electron transport (Kofer et al. 1998), but is essential in the absence of PGR5 (Munekage et al. 2004). The NDH complex is related to stress tolerance in plants, alleviating the stromal over reduction. It has been shown that tobacco mutants are sensitive to supra-saturating light and humidity stress (Horváth et al. 2000; Endo et al. 1999).

The NDH complex in chloroplasts consists of four subcomplexes, the membrane, the A, B and the lumen subcomplex. Eleven genes of the NDH complex (*ndhA-K*) are encoded by the plastid genome, while the other subunits are nuclear encoded (Peng & Shikanai 2011; Ifuku et al. 2011). The plastid genes are arranged in two operons, *ndhH-D* and *ndhC-J*, and two single genes, *ndhB* and *ndhF*. The *ndhH-D* operon includes the genes *ndhH*, *ndhA*, *ndhI*, *ndhG*, *ndhE*, *psaC* and *ndhD* in this order (Maier et al. 1995), giving rise to a complex pattern of transcripts that are produced after post transcriptional modifications, splicing (*ndhA* and *ndhD*) and editing (*ndhA*) (Figure 5) (del Campo 2009). This feature makes this operon the ideal

candidate to study post transcriptional control in gene expression in the chloroplasts.

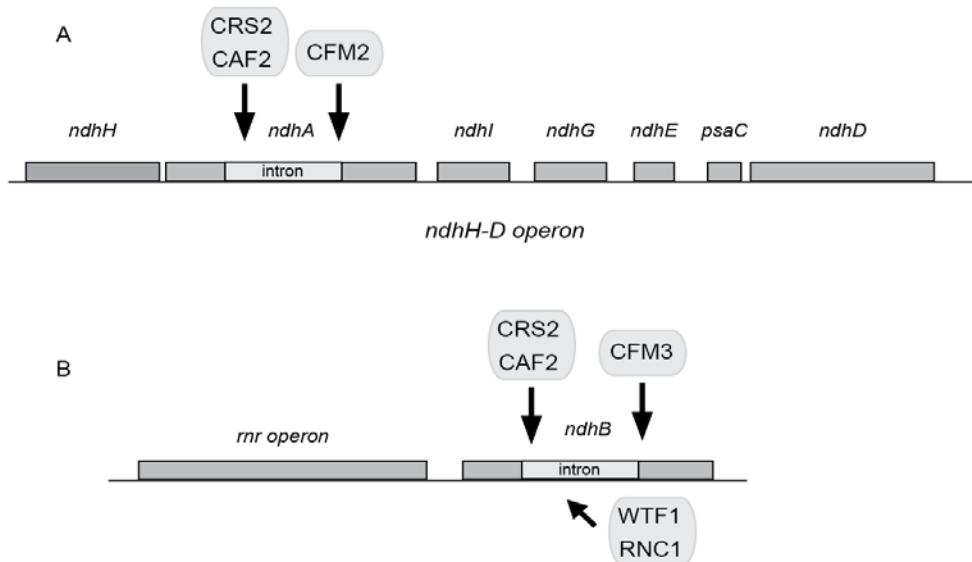


Figure 5. *ndh* genes and splicing factors in *Arabidopsis*.

ndhH-D operon (A) and *ndhB* gene (B) in *Arabidopsis* chloroplasts and the nucleus encoded splicing factors involved. All the splicing factors are described in 5.3.

The operon *ndhC-J* includes only three genes, the *ndhC*, *ndhK* and *ndhJ*. The last two genes, that are not part of an operon, are *ndhB* and *ndhF*, with the *ndhB* to have an intron in its sequence and also requiring post transcriptional modification (RNA splicing). Three splicing factors were identified for *ndhA* intron processing, belonging to the CRM and peptydyl tRNA hydrolase proteins, while splicing of the *ndhB* intron requires five factors; CRM, peptydyl tRNA hydrolases, PORR and ribonuclease II proteins (Figure 5.A, B).

Cytochrome *b₆f* (plastohydroquinone:plastocyanin oxidoreductase) is a complex that is embedded in the thylakoids of chloroplasts and has oxidoreductase activity. It is involved in linear electron transport (LET), transfers electrons between PSII and PSI and creates a proton gradient (ΔpH) across the thylakoid membrane (Munekage et al. 2004). Cytochrome *b₆f* is also involved in CET and in state transitions (Joliot & Joliot 2006). This complex consists of nine subunits, Cyt*b₆* (*petB*), Cyt*f* (*petA*), subunit IV (*petD*),

and the small subunits *petN*, *L* and *G* that are encoded in the plastid genome, and also *petC* (Rieske protein) and *petM* which are encoded in the nucleus. The *petB* and *petD* genes are part of *psbB* operon, transcribed with *psbB*, *psbT* and *psbH* genes. Both genes contain a group II intron in their sequences and require six splicing factors for correct processing (APO, CRM, PORR and ribonuclease proteins) (Figure 6).

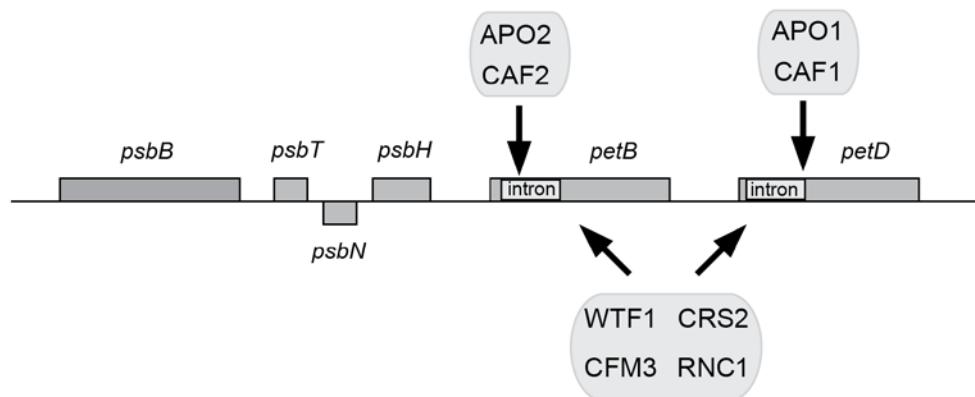


Figure 6. *psbB* operon in *Arabidopsis thaliana* and the splicing factors involved.

Adapted from (Stoppel & Meurer 2013).

1.11. Aim of Gc16 project

The aim of this project was to elucidate the function of two nuclear encoded proteins Gc16A (AT2G17240) and Gc16B (AT3G24506) that are part of the GreenCut project (Grossman et al. 2010; Karpowicz et al. 2011). The proteins were characterized by genetic and biochemical analyses. The effect of Gc16 disruption in photosynthetic complexes was analyzed by BN-PAGE analysis. The defect in NDH and *Cytb₆f* assembly was further studied on transcriptional and post-transcriptional level, by revealing the role of Gc16A/B.

2. Material and Methods

2.1. Database analysis and software tools

All gene sequences were obtained from the databases TAIR (www.arabidopsis.org), NCBI (www.ncbi.nlm.nih.gov) and Cyanobase (<http://genome.microbedb.jp/cyanobase/Synechocystis>). Transit peptide lengths were predicted by ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>) and TargetP (<http://www.cbs.dtu.dk/services/TargetP/>). The Vector NTI software (Invitrogen) was used for all sequence alignment and the generation of the constructs.

2.2. Chemical materials

All chemicals were purchased from Sigma-Aldrich (Steinheim, Germany), Roth (Karlsruhe, Germany), Duchefa (Haarlen, Netherlands), Applichem (Darmstadt, Germany), Serva (Heidelberg, Germany) and Invitrogen (Darmstadt, Germany).

2.3. Enzymes, kits, primers and radiochemicals

All the restriction enzymes used for cloning and Phusion High Fidelity DNA Polymerase were obtained from New England Biolabs (Ipswich, USA). Taq DNA Polymerase, DNA purification and Plasmid kits were purchased from Qiagen (www.qiagen.com). All primers were purchased from Metabion GmbH (Martinsried, Germany) and the radiochemicals from Hartmann Analytic (Braunschweig, Germany).

2.4. Bacterial cell cultures and growth conditions

All *E.coli* strains (DH5a, TOP10) used in this study were grown at 37°C in LB medium under continuous shaking at 225 rpm. The lycopene containing strain, harboring pAC-LYC plasmid (Cunningham et al. 1994) was also grown at 30°C, in LB supplemented with 34 µg/ml chloramphenicol. For the *E.coli* strain producing lutein, harboring both plasmids (pAC-LYC and pKS-LUT) the same growth conditions were used, but the LB medium was supplemented with 34 µg/ml chloramphenicol and 100 µg/ml kanamycin.

Synechocystis sp. PCC 6803 glucose tolerant wild type strain (GT, H. Pakrasi, Department of Biology, Washington University, St. Louis) and all the mutants

were grown at 30°C in BG11 medium containing 5 mM glucose (Rippka & Deruelles 1979), under continuous illumination at 30 µmol photons m⁻² s⁻¹, unless otherwise stated. In the BG11 of the plates for *Synechocystis* growth, 1% (w/v) agar was added and they were buffered with 10 mM TES-KOH (pH 8.0).

Table 2. *Synechocystis* strains used in this study

Strain	Characteristics	Marker	Source
<i>Synechocystis</i>			
WT	WT <i>Synechocystis</i> , glucose tolerant		H. Pakrasi
<i>SynLUT</i>	Introduction of lutein pathway genes, <i>nptl-sacB</i> and flanking regions for <i>sll0168</i> into <i>Synechocystis</i> WT	kan ^R , suc ^R	This study
<i>ΔcruA</i>	Spectinomycin cassette replacing <i>sll0147</i> in <i>Synechocystis</i> WT	Sm ^R	This study
<i>SynLUT/ΔcruA</i>	Spectinomycin cassette replacing <i>sll0147</i> , into <i>SynLUT</i> mutant strain	kan ^R , suc ^R , Sm ^R	This study
<i>SynZEP</i>	<i>ZEP_opt</i> , <i>nptl-sacB</i> into <i>sll0168</i> in <i>Synechocystis</i> WT	kan ^R , suc ^R	This study

Table 3. Plasmids used in this study

Plasmid	Characteristics	Marker	Source
pGEM Teasy	Backbone for many plasmids	amp ^R	Promega, Madison
pRL250	<i>nptl-sacB</i> , double selection cassette	Kan ^R , suc ^R	P. Wolk (Michigan University)
pKS	pGEM ^R -T Easy with <i>nptl-sacB</i> from pRL250	Kan ^R , suc ^R	This study
pFS-us	pGEM ^R -T Easy with upstream region of <i>sll0168</i>	amp ^R	This study
pFS-ds	pGEM ^R -T Easy with downstream	amp ^R	This study

	region of <i>sll0168</i>			
pPpsbAII	pGEM ^R -T Easy with <i>psbAII</i> promoter from <i>Synechocystis</i>	amp ^R	This study	
pPrbcL	pGEM ^R -T Easy with <i>rbcL</i> promoter from <i>Synechocystis</i>	amp ^R	This study	
pUC57-LCYe_opt	Codon optimized <i>LCYe</i>	amp ^R	GenScript	
pUC57-LCYb_opt	Codon optimized <i>LCYb</i>	amp ^R	GenScript	
pUC57-CYP97A3_opt	Codon optimized <i>CYP97A3</i>	amp ^R	GenScript	
pUC57-CYP97C1_opt	Codon optimized <i>CYP97C1</i>	amp ^R	GenScript	
pKS-LUT	piCH30791 modified with <i>LCYe</i> , <i>LCYb</i> , <i>CYP97A3</i> and <i>CYP97C1</i> optimized sequences, with <i>psbAII</i> and <i>rbcL</i> promoters, <i>nptI-sacB</i> selection marker and <i>sll0168</i> flanking regions	Sm ^R , kan ^R , Suc ^R	This study	
psll0147-up	pGEM ^R -T Easy with upstream Flanking Sequence of <i>sll0147</i>	amp ^R	This study	
psll0147-ds	pGEM ^R -T Easy with downstream Flanking sequence of <i>sll0147</i>	amp ^R	This study	
piCH30791mod	Destination vector for Golden Gate cloning. From piCH30791, with modified Bsal sites	Sm ^R	This study	
piCH30791	Entry vector for Golden Gate, used for Spectinomycin amplification	Sm ^R	E. Weber, Icon Genetics	
pSpec	pGEM ^R -T Easy +Spec	amp ^R , Sm ^R	This study	
pΔcruA	piCH30791 modified with Spectinomycin and flanking sequences of <i>sll0147</i>	Sm ^R	This study	

pUC57-ZEP_opt	Codon optimized Zeaxanthin Epoxidase	amp ^R	GenScript
pSR-ZEP_opt	pGEM ^R -T Easy with <i>ZEP</i> region for double recombination	amp ^R	This study
pZEP_opt	pICH30791 modified with <i>psbAII</i> promoter, <i>ZEP</i> optimized, <i>nptI-sacB</i> and flanking regions of <i>slr0168</i>	Sm ^R , kan ^R , suc ^R	This study

2.5. Construction of plasmids for *Synechocystis* transformation

For the assembly of multiple DNA fragments, the Golden Gate technique was chosen as cloning strategy (Engler et al. 2008; Engler et al. 2009). For all constructs a modified pICH30791 destination vector was used (pICH30791mod). This vector is derived from pUC19, where the ampicillin resistance-mediating gene has been replaced by a spectinomycin resistance-mediating gene. For this study, Bsal restriction sites were added flanking the lacZ gene (at the Pvull site, 2011 bp; at Zral site, 2708 bp) using the primers pAssembly Fw/Rv, shown in Table 4. In this vector, the positions of the Bsal sites were identical with those in the destination vector pICH69822 (Engler et al., 2008).

2.5.1. Construction of pKS-LUT

For the generation of the pKS-LUT construct, the whole lutein pathway genes from *Arabidopsis thaliana* were used and placed in the same construct. The genes responsible to convert lycopene to lutein were *LCYe* (GI: 19310742), *LCYb* (GI: 30681202), *CYP97A3* (GI: 145336316) and *CYP97C1* (GI: 42565880). The sequences for all four genes were codon optimized to improve the expression in *Synechocystis* (GenScript, USA) and cloned into EcoRV sites of pUC57. For each gene the cTP (length) was removed from the coding sequence, a tag (His for *LCYe*, c-myc for *LCYb*, Flag for *CYP97A3* and HA for *CYP97C1*) and a terminator (*Synechocystis psbAII* and *rbcL*) was placed at the C-terminus (<http://transterm.ccb.umd.edu/index.php>). Every gene was under the control of a strong *Synechocystis* promoter, for *LCYe* and *CYP97A* *PpsbAII* was used, while for the *LCYb* and *CYP97C* the *Prbcl* promoter (Mohamed et al. 1993; Kelly et al. 2009). *PpsbAII* was 94 bp long,

starting immediately upstream from ATG of *psbAII* gene and *PrbcL* was 100 bp long, starting from first ATG of the gene. The double selection cassette, *nptI-sacB* was amplified from pRL250, and placed downstream of the CYP97C gene. For homologous recombination in *Synechocystis*, the neutral integration site *sll0168* was used as flanking region (Kunert et al. 2000). For each flanking region 600 bp were amplified from *sll0168*, with the primers FS us Fw/Rv for the upstream region and FS ds Fw/Rv for the downstream region (Table 4). All the fragments were cloned into the pGEM-Teasy vector for sequencing and better assembly in the Golden Gate cloning. The whole construct was introduced into the pICH30791mod vector, according to the protocol described in Engler et al. (2009).

2.5.2. Construction of pΔcruA:Spec and pΔcruA:5'Spec

For the generation of pΔcruA:Spec construct, the spectinomycin resistance cassette was amplified from pICH30971, using the primers Spec Fw/Rv (Table 4) and was placed between flanking regions. For the upstream and downstream flanking regions 500 bp were amplified from *Synechocystis* genomic DNA allowing the exchange of the whole *sll0147* gene after homologous recombination. The whole construct was assembled in pICH30791mod, as mentioned earlier. In order to obtain a knock-down mutant of *cruA*, the construct pΔcruA:5'Spec was designed to replace a part of the promoter and the 5'-end of the *sll0147* gene with a spectinomycin-resistance mediating cassette. The primers used for the amplification of these parts are listed in Table 4 (Prom FS Fw/Rv and cruA FS Fw/Rv).

2.5.3. Construction of pKS-ZEP

To generate the pKS-ZEP vector, the coding sequence of zeaxanthin epoxidase gene from *Arabidopsis thaliana*, lacking the predicted cTP, was codon optimized for expression in *Synechocystis* (OptimumGene™ – Codon optimization algorithm, GeneScript, USA). The synthetic gene, with a 6xHis tag in the C-terminal followed by the terminator, of *rbcL* gene from *Synechocystis*, 44 bp long (<http://transterm.cbcn.umd.edu/index.php>) was purchased (*AtZEP_opt*) and cloned into EcoRV site of pUC57 vector (GenScript, USA). The gene was placed under the transcriptional control of

the strong and endogenous *Synechocystis psbAII* promoter (Mohamed et al., 1993), followed by a double selection cassette (*nptI-sacB*). Immediately downstream of the cassette, a fragment of the gene was placed, (the last 450 bp) that was amplified from pUC57+AtZEP_opt, for the second recombination (Viola et al., 2013) and the whole construct was placed between the *slr0168* flanking regions (600 bp each) for the stable integration into the *Synechocystis* genome.

Table 4. Primers used for in this study for the generation of the constructs and the genotyping of *Synechocystis* mutants.

Name	Sequence (5'→3')	Purpose
pAssembly Fw	TTT AGGTAGAGACC CTG GCA CGA CAG GTT TCC CG	Amplification of <i>lacZ</i> gene in pICH30791 with Bsal sites
pAssembly Rv	TTT AAGCAGAGACC GTC TAA GAA ACC ATT ATT ATC AT	
pKS-LUT		
FS us-Fw	TTT GGTCTC TAGGT TGCAGTTCAAACCTCGATAAC	Amplification of 600 bp of <i>slr0168</i> for upstream flanking region
FS us-Rv	TTT GGTCTC CAGTA AAATCGCTCCCTCCGTGCCT	
FS ds-Fw	TTT GGTCTC TCAAG CTCAAAGGGACGAAGCCGCAG	Amplification of 600 bp of <i>slr0168</i> for downstream flanking region
FS ds-Rv	TTT GGTCTC TAAGC ATCTGCCAAAGCTGCTTCTT	
<i>PpsbAII-1</i> Fw	TTT GGTCTC CTACT TTACAAAGCTTACAAAAC	Amplify 100 bp of <i>Synechocystis psbAII</i> promoter, between FS us and <i>LCYe</i>
<i>PpsbAII-1</i> Rv	TTT GGTCTC GCATT TGGTTATAATTCCATTATGTAT	
<i>PpsbAII-2</i> Fw	TTT GGTCTC ATGAA TTACAAAGCTTACAAAAC	Amplify 100 bp of <i>Synechocystis psbAII</i> promoter,

for CYP97A

<i>PrbcL-1</i> Fw	TTT GGTCTC CGGTA AAATAATAACTGTCTCTGGG	Amplify 100 bp of <i>Synechocystis rbcL</i> promoter, in front of <i>LCYb</i>
<i>PrbcL-1</i> Rv	TTT GGTCTC GCATT CTAGGTCAAGTCCTCCATAA	Amplify 100 bp of <i>Synechocystis rbcL</i> promoter, in front of <i>LCYb</i>
<i>PrbcL-2</i> Fw	TTT GGTCTC TATCA AAATAATAACTGTCTCTGGG	Amplify 100 bp of <i>Synechocystis rbcL</i> promoter, in front of <i>CYP97C</i>
Sel. Cassette Fw	TTT GGTCTC ACGTT GGAATTGATTGATCCGTCGAC	Amplify <i>nptI-sacB</i> from pRL250
Sel. Cassette Rv	TTT GGTCTC CCATA CTT TAG GCC CGT AGT CTG CA	

pΔcruA:Spec

<i>sII0147</i> us Fw	TTT GGTCTC TAGGT TTAAAGAACATGAATTGGGGTG	Amplify 500 bp upstream of <i>cruA</i> gene in <i>Synechocystis</i>
<i>sII0147</i> us Rv	TTT GGTCTC CAGTA CAAAATTGCCACTGAAATCA	
<i>sII0147</i> ds Fw	TTT GGTCTC TCAAG TAAATTAGGGCGACATTAAG	Amplify 500 bp upstream of <i>cruA</i> gene in <i>Synechocystis</i>
<i>sII0147</i> ds Rv	TTT GGTCTC TAAGC TGGGTCGCCCTTCTGCACCT	
Spec Fw	TTT GGTCTC CTACT CCCTGATAAAATGCTTCAATAA	Amplify Spec from pIC30791
Spec Rv	TTT GGTCTC TCTTG TTATTGCCCCACTACCTTGG	

pΔcruA:5'Spec

Prom FS Fw	TTT GGTCTC TAGGT GGCCTATGATTCCCTAACCA	Amplify promoter part of <i>sII0147</i> , for upstream flanking sequence
Prom FS Rv	TTT GGTCTC CAGTA	

	TGGGGTGTATAAACCCGC	
cruA FS Fw	TTT GGTCTC TCAAG TTCCTCCGGCCAAAGCAGTG	Amplify part of <i>sll0147</i> , for downstream flanking sequence
cruA FS Rv	TTT GGTCTC TAAGC TCATTGGGCCGAAGGGCAAT	
pKS-ZEP	-	-
PpsbAII 3 Fw	TTT GGTCTC CTACT GCCCTCTGTTACCCATGGAAA	Amplify 180 bp of <i>psbAII</i> promoter from <i>Synechocystis</i>
PpsbAII 3 Rv	TTT GGTCTC GCATT TGGTTATAATTCCATTATGT	
SR_ZEP Fw	TTT GGTCTC TTATG TTAACCAAAGACGAAGATCAAC	Amplify 450 bp from <i>ZEP_opt</i> for second recombination and probe
SR_ZEP Rv	TTT GGTCTC TCTTG AAAAGCAACAAAAACCCGC	for Northern blot
Sel.cassette 2 Fw	TTT GGTCTC AGGTA GGAATTGATTGATCCGTCGAC	Amplify <i>nptI-sacB</i> with sel.cassette Rv

General:

cruA probe Fw	GGCCGAATGAATCGGGATGGAATA	Detection of <i>cruA-sll0147</i> ,
cruA probe Rv	TCCCCCTGCTTGCTGGAGTTA	for Northern blot analysis
cruA gene Fw	ACTGCTCTACTGTGAAGTCCCCACC	Genotyping of Δ <i>cruA</i> for complete segregation
cruA gene Rv	GCAGGTCATACTGCAAAGCTGTGTG	
LCYe Rv	TACCAAAAGCAACAAAAACCCGCG	Genotyping of SynLUT, with FS us-Fw
LCYb Fw	ATGGTTAGCAGCGTTGTGAGTG	Genotyping of SynLUT
LCYb Rv	CACGGTCTTGCACCAGGTTG	

CYP97C Fw	GGGTGAGTCCCGATTGGTTGA	Genotyping of <i>SynLUT</i>
CYP97C Rv	AACGTCAGTCAAGCAAAAGCAC	
<i>sacB</i> Fw	AGCATATCATGGCGTGTAATATGGG	Genotyping of <i>SynLUT</i> , with FS ds-Rv
<i>sll0146</i> Fw	GCTAATGTTATTAGCTTAAACC	Genotyping of Δ <i>cruA</i> for correct integration, with Spec Rv

2.6. *Synechocystis* transformation

The transformation of *Synechocystis* wild type or mutant strains was performed with the natural transformation method (homologous double recombination) (Williams 1988). The plasmids used in this case were purified with the Qiagen plasmid Midiprep kit. For each transformation, 10 ml of cells at OD₇₃₀ of 0.4 were pelleted at 6,000xg for 15 min and then resuspended with BG11 at 1/20 of the initial volume. In this suspension, 2-3 µg circular DNA was added and incubated for 6 h at 25°C in the light, with the last 3 h shaking. For the recovery of the cells, BG11 medium was added to the transformation assay and incubated overnight at 28°C in the dark. The next day, cells were harvested by centrifugation at 4,500xg for 10 min and after resuspending them in a small volume of BG11, they were plated on BG11 agar plates supplied by the appropriate antibiotic. For selection of mutants carrying the lutein cassette (*SynLUT*) and the *ZEP* gene (*SynZEP*) 10 µg/ml kanamycin was added to the BG11 agar plates, whereas for the selection of Δ *cruA* mutants 5 µg/ml spectinomycin was added. In the case of *SynLUT*/ Δ *cruA* the selection was performed at 100 µg/ml kanamycin and 5 µg/ml spectinomycin, since both selection cassettes were present and the *SynLUT* mutant used as background strain was fully segregated at 100 µg/ml kanamycin. For the complete segregation of the mutants concentrations of antibiotics were steadily increased, for kanamycin up to 100 µg/ml and for spectinomycin up to 200 µg/ml.

2.7. Nucleic acid isolation:

2.7.1. PCR (standard and high fidelity)

PCR analysis was performed for genotyping of *Synechocystis* mutant strains and *E.coli*. The PCRs were performed in a total volume of 20 µl, containing 2 µl 10x PCR-buffer, 100 mM dNTPs, 200 µM primers (listed at Table 3) and 0.5 units of Taq DNA Polymerase (Qiagen). The PCR products were loaded on a 1% agarose gel.

DNA fragments were amplified from *Synechocystis* cDNA (for cruA probe) and *Synechocystis* genomic DNA (for P_{psbAI} , P_{rbcL} , FS $slr0168$ upstream flanking sequence/downstream flanking sequence) with the Phusion High-Fidelity DNA Polymerase (NEB). All the reactions were performed in a total volume of 20 µl, where 1x Phusion HF reaction, 200 µM dNTPs, 0.5 µM of each primer (listed at Table 4) and 0.4 units Phusion DNA Polymerase were added to the reaction. The amplified fragments were loaded on an agarose gel and purified with the QIAquick Gel Extraction kit, according to the manufacturer.

2.7.2. Genomic DNA isolation from *Synechocystis*

Genomic DNA isolation for small scale preparations was performed according to the xanthogenate-SDS method from (Tillett & Neilan 2000). Briefly, 1 ml of bacterial medium of exponentially growing cell cultures were pelleted and resuspended in 50 µl of TER buffer (10 mM Tris/HCl pH 7.4, 1 mM EDTA pH 8.0 and 100 µg/ml RNAse A). 750 µl freshly made XS buffer were added (1% calciummethylxanthogenate, 100 mM Tris/HCl pH 7.4, 20 mM EDTA pH 8.0, 1% SDS, 800 mM ammonium acetate) and mixed by inversion. The samples were incubated at 70°C for 2 h in order to dissolve membranes. The samples were vortexed for 10 seconds and incubated on ice for 30 min. Samples were centrifuged for 10 min at 13,000xg to remove all cell debris and the supernatant was transferred to a new tube containing 750 µl isopropanol for precipitation. The DNA was collected by centrifugation at 12,000xg for 10 min and washed with 70% ethanol. Finally, the DNA was air dried and resuspended in 100 µl TE buffer (10mM Tris/HCl pH 7.4, 1 mM EDTA pH 8).

2.7.3. RNA isolation

Total RNA was isolated from *Synechocystis* samples using the TRIzol method. Cells from 50 ml liquid cultures ($OD_{730}=0.7$) were pelleted at 6,000xg for 15 min, resuspended in 1 ml TRIzol (phenol 38%, guanidine thiocyanate 0.8 M, ammonium thiocyanate 0.4 M, sodium acetate 0.1 M pH=5.0, glycerol 5%) and vortexed thoroughly. The samples were immediately frozen in liquid nitrogen and incubated at 65°C for 15 min. After this procedure was repeated at least two times, cell debris was removed by centrifugation at 12,000xg for 15 min at 4°C. The supernatant was transferred into a new tube, mixed with 0.2 volumes of chloroform and incubated at room temperature for 15 min. After phase separation by centrifugation at 12,000xg for 15 min at 4°C, the aqueous phase was transferred into a new tube. RNA was precipitated with 0.25 volume of isopropanol and 0.25 volume of a high salt solution (1.2 M NaCl and 0.4 M sodium citrate) and washed with 70% ethanol. The dried RNA pellet was resuspended in 100 μ l H₂O and concentration was measured by Nanodrop (Nanodrop 200, Peqlab).

2.8. cDNA synthesis

Synthesis of cDNA from *Synechocystis* was performed with the iScript cDNA synthesis kit (Bio-Rad, www.bio-rad.com). Prior to the reaction, DNase I (New England Biolabs) treatment of the RNA was performed, in order to remove genomic DNA contamination. 1 μ g of RNA was incubated with 0.5 μ l of 10x PCR buffer (Qiagen) and 0.5 units of DNase I at room temperature for 30 min in total reaction volume of 10 μ l. For enzyme deactivation 2.5 mM EDTA, was added and incubated at 65°C for 15 min. The whole reaction was further used for the reverse transcription reaction, in a total volume of 20 μ l. The reaction contained 1x iScript reaction mix, 1 μ l iScript Reverse Transcriptase and DEPC H₂O. The cDNA synthesis was performed according to the manufacturer's instructions.

2.9. Northern blot analysis

Northern blot analysis was performed according to Sambrook and Russel (2001). 10 μ g total RNA was fractionated in a denaturing 1.2% (w/v) agarose gel, which contains 1x MOPS buffer (200 mM MOPS, 50 mM sodium acetate and 10 mM EDTA, pH 7) and 1% formaldehyde. The RNA samples were

mixed with 5x RNA loading dye and incubated at 95°C for 5 min. After denaturation, the samples were loaded on an agarose gel and run for 2-3 h, at ~40 V, in a 1x MOPS running buffer.

Then, the gel was subjected to the following manipulations for the RNA capillary transfer to the membrane. The gel was equilibrated for 10-15 min in 10x SSC (1.5M Na-chloride and 150 mM Na-citrate, pH 7). The capillary transfer was performed by placing the gel upside down on the transfer bridge. The positively charged nylon membrane (Hybond N+; GE Healthcare, Freiburg, Germany) was shortly equilibrated in 2x SSC and layered on top of the gel, as well as three Whatman papers (3MM). On top of this stack, paper towels and an extra weight were placed in order to start the capillary transfer. The transfer solution was 10x SSC and the transfer was carried out overnight (16-20 h).

Next, the membrane was washed with 2x SSC and dried for 30 min. RNAs were cross-linked on the membrane by UV radiation (Stratalinker® UV Crosslinker 1800, Stratagene, USA) at 1200 μ J cm⁻².

The membrane was equilibrated in 2x SSC solution and placed into a glass cylinder containing 20 ml prehybridization buffer (7% SDS, 0.25 M Na₂HPO₄ pH 7) and 160 μ l denatured herring sperm. The membrane was incubated at 65°C for at least 4 h prior to hybridization. Finally, the radioactive labeled probe was added and the hybridization was performed for 16 h at 65°C. Hybridization buffer was discarded and 10 ml of pre-warmed washing buffer (0.1% SDS, 0.2 M NaCl, 20 mM NaH₂PO₄, 5 mM EDTA; pH7.4) was added to the membrane and incubated for 30 min at 65°C. The second washing step was performed with the same buffer but incubated for 15 min at 65°C. The final washing step was carried out with 1x RT buffer (6 mM NaH₂PO₄, 1 mM EDTA, 0.2% SDS; pH 7.0) for at least one hour on a shaker at room temperature. The membrane was then exposed to a radioactive sensitive screen (Storage Phosphor Screen, Fuji) overnight. The signals were detected with the Phosphorimager (Typhoon, GE Healthcare).

2.10. Radioactive probe preparation

Approximately 100 ng of DNA (purified PCR product) was diluted in 12 µl ddH₂O and used as template for radioactive labeling. The template was denatured at 100°C for 5 min and then placed on ice for 2 min. 4 µl of 1x OLB buffer (50 mM Tris pH6.8, 10 mM MgOAc, 50 mM DTT, 0.5 mg/ml BSA, 33 µM of dATP, dTTP and dGTP), 1 µl Klenow DNA polymerase (New England Biolabs) and 3 µl radioactive ³²P-dCTP were added to the probe. The reaction was incubated at 37°C for 60 min and the probe was purified with Illustra MicrospinTM G-25 Columns (Freiburg, Germany) according to the manufacturer's instructions. The probe was subsequently denatured at 100°C for 5 min in the presence of 40 µl Herring sperm DNA (10 ng/ µl) and added to the hybridization solution. Primers used for the amplification of the probes are listed in Table 4.

2.11. Spectroscopic analysis

Absorption spectra of *Synechocystis* cells cultures were recorded using a spectrophotometer (Shimadzu). The optical density of the suspensions was measured at 730 nm and adjusted for all the samples to 0.5.

2.12. Pigment extraction and HPLC analysis

Pigment analysis was performed by High Performance Liquid Chromatography (HPLC). *Synechocystis* cell samples were grown until they reached the exponential growth phase (OD₇₃₀ approximately 0.5). For all extractions, 1 ml of liquid culture was harvested by centrifugation and pigments were extracted using 100% acetone. The samples were incubated in the dark on ice for 15 min. To remove the cell debris the samples were centrifuged and the supernatant was collected. This procedure was repeated for the complete extraction of carotenoids and the extracts were combined. One final centrifugation was performed to remove any contaminants and the samples were directly subjected to HPLC analysis. Pigments were analyzed by reverse phase HPLC as described previously by Färber et al. (1997).

Pigment extraction from *E.coli* was carried out using pure acetone and samples were incubated at 55°C for 15 min in the dark with intermediate vortexing. The extraction was repeated until cell pellets were colorless. The

cell debris was removed by centrifugation and the supernatants were subjected directly to HPLC for analysing and quantifying pigment composition and amounts, respectively. A lutein standard solution (Lutein analytical standard, Sigma Aldrich) with a known lutein concentration was used to verify the presence of lutein, while pKS-LUT was expressed in *E.coli* cells producing lycopene.

Aliquots of the extractions were separated on C₁₈ column (Grom Sil 120 ODS-5 ST, 3 µm particle size, 150 mm length, 2-mm inner diameter). Elution was performed for 3 min with 60:40 acetone/ water mixture (pH 3.5), followed by a linear gradient for 15 min at a flow rate of 100 µl/minute. Afterwards, 100% acetone was subjected through the column for 15 min and in the end 60% acetone for 10 min, so the column can be rinsed. Absorption spectra for individual peaks were recorded continuously and obtained with a photodiode array detector in the range of 350-750 nm. Carotenoids were identified by the absorption spectra and retention times.

2.13. Antibodies

Commercially available antibodies were used against PSI (PsaA and PsaC), PSII (D2 and CP43), RbcL and Cytb_{6f} (Cytb₆, Cytf) were obtained by Agrisera (Vänas, Sweden). Antibody against AtpF was kindly provided by Jörg Meurer (LMU, Munich, Germany) and NDH antibodies from Peter Westhoff (Düsseldorf University). Lhcb2 antibody was provided by S. Jansson, Umeå University, Sweden.

2.14. Plant material and growth conditions

Three different *Arabidopsis thaliana* T-DNA insertions mutants were used for this study. The *gc16a* line carrying an integration in gene AT2G17240 was obtained from the SALK collection (SALK_133989) (Alonso et al. 2003), whereas for *gc16b* (AT3G24506) the line *gc16b-1* were isolated from the GABI-KAT collection (GABI 295A02) (Li et al. 2003). A second insertion line (*gc16b-2*) for gene AT3G24506 was obtained from the SAIL collection (SAIL_71_A01) (Sessions et al. 2002). All three lines were derived from the Columbia-0 ecotype background.

Arabidopsis seeds from wild type and mutant lines were incubated at 4°C for two days, for stratification and then transferred on soil. The plants were grown under controlled greenhouse conditions or in growth chambers (PDF: 70-90 µmol photons m⁻²s⁻¹, 16h light/ 8 h dark). Fertilization of the plants was added according to the manufacturer's instruction (Osmocote Plus). For growth kinetics, leaf areas were recorded by photography at distinct time points after germination (7 to 27 days) and the surface was measured with ImageJ software (Abramoff et al., 2004). Murashige and Skoog (MS, Duchefa) plates supplemented with 1% sucrose were used for growth of the double mutants (*gc16a/b-2*) and to select transformed plants on antibiotics.

2.15. Complementation of mutant lines

Gc16A and *Gc16B* coding sequences without the stop codon were amplified with Phusion DNA Polymerase (NEB,) and cloned into an entry vector using the Gateway system (Invitrogen®). Then the genes were transferred into the binary vector pB7FWG2.0 (Karimi et al. 2002) where they were placed under the control of the cauliflower mosaic virus 35S promoter and fused with their 3' end to the sequence coding for the green fluorescent protein, eGFP of *Aequorea victoria*. The final constructs, 35S:*Gc16A:GFP* and 35S:*Gc16B:GFP*, were first introduced into *Agrobacterium tumefaciens* strain GV3101 and then into *gc16a/b-1* and *gc16a/b-2* plants by the floral dip method (Clough, S.J. and Bent 1998). Transformed plants were selected by several BASTA treatments (glufosinate ammonium) and resistant plants were checked by PCR for insertion of *Gc16A:GFP* and *Gc16B:GFP*. The protein synthesis in the transformed plants was examined by immunoblot analysis and by NDH activity measurements.

2.16. Bioinformatic sources

Gene and protein sequences were obtained from NCBI and TAIR (www.ncbi.nlm.nih.gov, www.arabidopsis.org). Sequence analysis was performed with Vector NTI software (Invitrogen®) and the chloroplast transit peptide lengths were predicted using ChloroP 1.1 (<http://www.cbs.dtu.dk/services/ChloroP/>). The sequences were aligned using MAFFT (<http://mafft.cbrc.jp/alignment/software/>) and formatted using the

Boxshade server (<http://mobyle.pasteur.fr/cgi-bin/portal.py#forms::boxshade>). Protein molecular weight was calculated using PPDB database (<http://ppdb.tc.cornell.edu/>) and the co regulation of expressed genes was performed with ATTED-II (<http://atted.jp/>).

The sequence identifiers for the alignment of Gc16A and Gc16B with homolog proteins in eukaryotic organisms are: *Vitis vinifera* (GI: 225445256), *Oryza sativa* (GI: 115447287), *Populus trichocarpa* (GI: 224143529), *Camelina sativa* (GI: 727588954), *Zea mays* (GI: 413937772), *Chlamydomonas reinhardtii* (GI: 159484224), *Ostreococcus lucimarinus* (GI: 145340865), *Jatropha curcas* (GI: 802627410), *Beta vulgaris* (GI: 731359496), *Malus domestica* (GI: 658004660) and *Glycine max* (GI: 356521959).

2.17. Bacterial strains-vectors

The bacterial strains used in this study were: *E.coli* TOP10 (Life technologies) and BL21- CodonPlus (DE3)-Ripl (Stratagene). The strain of *Agrobacterium tumefaciens* was GV3101 (pMP90RK).

2.18. DNA extraction

Arabidopsis genomic DNA was isolated from three to four-week-old leaves. After freezing in liquid nitrogen plant material was disrupted with metal beads and extraction buffer (200 mM Tris HCl pH 7.5, 250 mM NaCl, 25 mM EDTA and 0.5 % SDS) was added. The DNA was precipitated by addition of 0.8 volumes of isopropanol and by centrifugation at 16000xg for 20 min. The pellet was washed with 70% ethanol and resuspended in 100 µl ddH₂O. 1 µl of genomic DNA was used for further PCR analyses, in a total volume of 10 µl reaction. The primers used for genotyping of the mutants are shown in Table 5.

2.19. RNA extraction and Northern blot analysis

For expression analysis RNA was isolated with TRIzol reagent according to manufacturer's instructions. Northern blot analysis was performed as described in Chapter 2 (Sambrook, J. and Russel 2001).

For the detection of *tRNAs* the end labeling procedure was used. 10 µM DNA used for template and 2 µl T4 PNK, 5 µl 10x PNK buffer and radioactive (γ -³²P)-ATP (3 µl) were added in the reaction, in a final volume of 50 µl. The whole reaction was incubated for 30 min at 37°C. Before hybridization, the probe was denatured with 80 µl herring sperm DNA and added to the solution.

2.20. cDNA synthesis and RT-PCR analysis

Synthesis of cDNA from *Arabidopsis* was performed with the iScript cDNA synthesis kit (Bio-Rad, www.bio-rad.com), as previously described in Chapter 2. The primers used for the amplification of radioactive probes for Northern blot analysis are listed in Table 5.

2.21. Real Time PCR

RNA used for Real time PCR analysis was isolated using the RNeasy Mini Kit (Qiagen). The reactions were performed using iQ™ SYBR Green Supermix (Bio-Rad), cDNA as template and gene specific primers, as designed by the web tool of Roche (Universal ProbeLibrary Assay Design Center) (Table 5). Actin was used as reference gene and a negative control was also included in the reactions. The PCR program was a two-step protocol, using the following thermocycling profile: 2 min at 50°C and 95°C for 3 min for activation of DNA Polymerase and initial denaturation, followed by 48 cycles of denaturation (95°C for 15 seconds) and annealing (60°C for 30 seconds) and a last step of denaturation at 95°C for 10 seconds. The amplification was monitored by SYBR-Green fluorescence signals in a IQ™5 Multicolor Real time PCR Detection system (Bio-Rad). Gene expression amounts were calculated using the standard curve method of IQ™5 Optical System software.

2.22. Pigment analysis

Pigments were analyzed by reverse-phase HPLC as described previously (Färber et al., 1997). Pigments were extracted from 4 week-old wild type, *gc16a*, *gc16b-1*, *gc16b-2*, *gc16a/b-1* and *gc16a/b-2* leaves. The leaves were frozen with liquid nitrogen and grinded with beads in the presence of 100% aceton. The supernatants were collected after centrifugation (at 16000xg for 20 min at 4°C) and used directly for HPLC or stored at -80°C for 2 days (in collaboration with Peter Jahns, Düsseldorf, Germany).

2.23. Protoplast isolation and GFP detection

Protoplasts were isolated from *Arabidopsis* leaves using the TAPE-Arabidopsis Sandwich method (Wu et al. 2009). Leaves from 3-4 weeks old plants, carrying the 35S:Gc16A:eGFP or 35S:Gc16B:eGFP construct were collected. After the stabilization of the upper surface on a tape, the lower epidermal phase was removed carefully with a transparent tape. The peeled leaves were transferred to a petri dish containing 10 ml enzyme solution (1% cellulose R10, 0.25% macerozyme, 0.4 M mannitol, 10 mM CaCl₂, 20 mM KCl 0.1% BSA and 20 mM MES, pH 5.7). The leaves were shaken at 40 rpm for 60 min under light, until the protoplasts were released to the solution. The protoplasts were pelleted (3 min at 100xg) and washed two times with W5 solution (154 mM NaCl, 125 mM CaCl₂, 5 mM KCl, 5 mM glucose and 2 mM MES, pH 5.7). After an incubation of 30 min on ice the protoplasts were resuspended in modified MMg solution (0.4 M mannitol, 15 mM MgCl₂ and 4 mM MES, pH 5.7). The GFP signal and chlorophyll fluorescence from the protoplasts was detected using an Axio Imager fluorescence microscope with an integrated ApoTome (Zeiss, Jena, Germany).

2.24. TEM analysis

TEM analysis was carried out in 20 day old plants, in collaboration with Prof. G. Wanner (Department of Biology, LMU, Munich).

2.25. Chlorophyll fluorescence measurements- NDH activity

Chlorophyll fluorescence a measurements were performed using the DUAL PAM 10 (Walz GmbH, Effeltrich, Germany). Plants were dark adapted for 20 min before measurements. The measurements were performed for P700 and fluorescence simultaneously. Single leaves of plants grown under long day conditions were exposed shortly to measuring light for minimal chlorophyll fluorescence yield determination (F_0) and then to a saturating light pulse (10000 $\mu\text{E m}^{-2} \text{s}^{-1}$, 800 ms) for maximal chlorophyll fluorescence yield determination (F_m). Actinic red light was applied for 10 min and steady-state fluorescence yields (F_S) were measured. A saturating light pulse (10,000 $\mu\text{E m}^{-2} \text{s}^{-1}$, 800 ms) at the end of the light phase was applied to determine the maximal fluorescence yield in the light ($F_{m'}$). After the dark relaxation phase (10 min), F_m'' and F_0'' were measured by applying a saturating light pulse

($10,000 \mu\text{E m}^{-2} \text{s}^{-1}$, 800 ms). Parameters were calculated according to (Rühle et al. 2014). For P_{700} measurements, P_0 was measured after the dark adaptation and then P_m values for oxidized P700 from far red light were monitored every minute. The photochemical quantum yield of PSI (YI) is calculated as followed: $Y(I) = 1 - Y(ND) - Y(NA)$, $Y(ND)$ is the oxidized overall P700 and $Y(NA)$ is the P700 fraction that cannot be oxidized.

The transient increase in chlorophyll a fluorescence was also measured. The transition from dark to light was used for the determination of NDH activity in intact leaves, as previously described in (Shikanai et al. 1998). In brief, the cyclic electron flow around PSI was monitored after applying a saturating light to determine the maximal fluorescence (F_m), which was followed by a 5 min actinic light illumination ($80 \mu\text{mol photons m}^{-2} \text{s}^{-1}$) and a dark incubation for 5 min. The transient rise in fluorescence after switching off the light represents the activity of NDH complex.

2.26. Generation of Gc16 antibody

An antibody against both Gc16A and GC16B was produced in rabbits. The peptide synthesis, immunization of the animals and monospecific purification was performed by Biogenes (Berlin, Germany). The peptide was 14 amino acids long and it was designed to recognize both of the proteins (73-86 aa for Gc16a and 81-94 aa for Gc16B). The amino acid sequence of the peptide was PLDFPIEWERPKPG in a one letter code.

2.27. Total protein isolation

Leaves were disrupted in liquid nitrogen and the proteins were isolated in 2x Laemmli buffer (200 mM Tris-HCl pH6.8, 4% SDS, 20% glycerol and 5% β -mercaptoethanol). The samples were incubated for 15 min at 65°C and then centrifuged at 16000xg for 10 min. The supernatant was collected and then boiled at 95°C for 5 min for denaturation of the proteins. Total proteins were loaded on Tris-glycine (Schägger & von Jagow 1987) or Tris Tricine SDS-PAGE (Schägger 2006).

2.28. Thylakoid isolation

Leaves from 4-5 weeks old plants were harvested and homogenized with T1 buffer (0.4M sorbitol, 0.1M Tricine pH7.8 and 1 mM PMSF). The homogenate

was filtered through a Miracloth (Calbiochem®) and then centrifuged at 1500xg for 4 min at 4°C. Subsequently the chloroplasts were disrupted by resuspending the pellet in T2 buffer (20mM HEPES/KOH pH7.5, 10mM EDTA) and then centrifuge at 9500xg for 10 min at 4°C. Chlorophyll concentration of the samples was measured according to Porra et al. 1989.

2.29. Immunoblot analysis

The proteins that were separated in SDS-PAGE gels were subsequently transferred on polyvinylidene difluoride (PVDF) membranes (Immobilon-P, Millipore, Germany) by Western blotting according to the manufacturer's instructions using a semi dry blot system (Bio-Rad). The filters were blocked with 5% milk in TBST buffer (10 mM Tris pH8, 150 mM NaCl, 0.1% Tween-20) for at least one hour at room temperature and then incubated with specific antibodies for the proteins of interest. Bound secondary antibodies coupled to the HRP were visualized using the ECL kit (Amersham Bioscience®) and signals were detected using an ECL reader (Fusion FX7, Peqlab®, Germany).

2.30. Blue Native gel analysis and second dimension gel

Leaves from 4-5 weeks old plants grown under long day conditions were harvested and thylakoids were isolated according to 6.17. The samples were then adjusted according to chlorophyll concentration or fresh weight and subsequently washed four times with washing buffer (25mM BisTris/HCl pH7.0, 20% glycerol). Thylakoid membranes were solubilized with 25mM BisTris/HCl pH7.0, 20% glycerol, 1% β-dodecyl maltoside (β-DM) and incubated on ice for 10 min. Solubilized were separated from unsolubilized proteins by centrifugation at 16000xg for 20 min at 4°C. The supernatant was supplemented with 1/10 volume of BN loading buffer (100mM BisTris/HCl pH 7.0, 750mM ε-aminocaproic acid, 5% (w/v) Coomassie G-250) and loaded on BN-PAGE gels (4-12% acrylamide gradient). BN-PAGE was prepared as described in (Schägger et al. 1994).

To separate the complexes into their subunits, each stripe from the first dimension BN-PAGE was treated with 0.5M Na₂CO₃, 2% (w/v) SDS and 0.7% (v/v) β-mercaptoethanol, for 30 min at room temperature and then loaded on a Tricine-SDS-PAGE gel, according to (Schägger 2006). The gels were

subsequently used for immunoblot analysis with specific antibodies (NdhH and Cytb₆).

2.31. In vivo labeling assay with ³⁵S-Methionine

The labeling of the chloroplastic proteins was performed according to Pesaresi et al. (2011). Briefly, leaf disks of 4 mm were incubated with 20 µg/ml cycloheximide, 1 mM K₂HPO₄/ KH₂PO₄ pH 6.3 and 0.1 % (w/v) Tween-20. (³⁵S) Methionine was added to the buffer, in a final concentration of 0.1 mCi/ml, and incorporated into the samples by vacuum-infiltrated. Afterwards, leaves were exposed to low light (20 µmol photon m⁻² s⁻¹) and at each selected time point (10, 20 and 40 min) four leave disks were collected. Total protein extraction was performed as previously described and the proteins were fractionated in a 12% Tris-glycine SDS-PAGE.

2.32. EMSA

Protein purification was performed according to Stoppel et al. (2012) and the electrophoretic mobility shift assay according to Manavski et al. (2012).

2.33. Primer list

Table 5. Sequences for the primers used for this study.

Name:	Primer sequence (5'→3'):
<i>Genotyping:</i>	
<i>gc16a FOR:</i>	GAAGGTTCGAATTTCGAAAGG
<i>gc16a REV:</i>	ACTATCAAAACGCAAACGCAG
<i>LBb1.3:</i>	ATTTTGGCGATTTGGAAC
<i>gc16b-1 FOR:</i>	GCGAGTCTAGTCATGTCATTGG
<i>gc16b-1 REV:</i>	TAGCAATGTGAAATTCGACG
<i>LBGK1:</i>	CCCATTGGACGTGAATGTAGACAC
<i>gc16b-2 FOR:</i>	TAGGAACCATCGAAGAACACG
<i>gc16b-2 REV:</i>	AGCTTCCACTAAAGCGCTTTC
<i>LB3:</i>	TAGCATCTGAATTCATAACCAATCTCGATACAC
<i>Northern blot:</i>	
<i>ndhA FOR:</i>	ATGATAATTATGCAACAGCAGTCCAAACT
<i>ndhA REV:</i>	CAGCTCGCAGACCACCTAAAAAGAATAT
<i>ndhB FOR:</i>	AAAGCCTTCATTGCTTCTTCGATG

<i>ndhB</i> <i>REV:</i>	GCCCCACCCATGAGTAAATATTCATAGTA
<i>ndhC</i> <i>FOR:</i>	CAAGTGCTATTCCCTGTTTGGCATTTC
<i>ndhC</i> <i>REV:</i>	CCTTTGCCATGCATAAACTAAACCA
<i>ndhD</i> <i>FOR:</i>	GTGTATCTTGTCTTACCACG
<i>ndhD</i> <i>REV:</i>	AGACGTTCTTCCACCCA
<i>ndhE</i> <i>FOR:</i>	ATACTCGAACATGTACTTGTGAGTGCC
<i>ndhE</i> <i>REV:</i>	GGTCGATTGGTTATGCGAATTGAT
<i>ndhF</i> <i>FOR:</i>	AGGT CCTATTGCTAAATCCGC
<i>ndhF</i> <i>REV:</i>	CGGTTAATCCGCTGTTGAA
<i>ndhH</i> <i>FOR:</i>	TAAAACCCGGTGGTCGTATTTCC
<i>ndhH</i> <i>REV:</i>	GAGGATGTTGTTGACTGTGAACCCA
<i>ndhI</i> <i>FOR:</i>	TTGAGGTGAATTCAAATTGTTCGA
<i>ndhI</i> <i>REV:</i>	CCGGGTTCATGAATTATGGTCAAC
<i>ndhG</i> <i>FOR:</i>	ATTGACGAGCCACAGAAATTGCAC
<i>ndhG</i> <i>REV:</i>	CTTCCAACCCAATATTCAGCC
<i>ndhJ</i> <i>FOR:</i>	TCCTTACGTAAAGGCCACCCCTATCC
<i>ndhJ</i> <i>REV:</i>	GGTCATAGATCGTTGGCTTCG
<i>ndhA</i> <i>intr</i> <i>FOR:</i>	GAACCGTACATGAGGTCTGGCC
<i>ndhA</i> <i>intr</i> <i>REV:</i>	CAGAGTATGCTCCTATCCACCGACA
<i>petB</i> <i>FOR:</i>	ATGAGTAAAGTTATGATTGG
<i>petB</i> <i>REV:</i>	TTATAAGGGACCAGAAATGA
<i>rrn16</i> <i>FOR:</i>	AGTCATCATGCCCTTATGC
<i>rrn16</i> <i>REV:</i>	CAGTCACTAGCCCTGCCCTC
<i>rrn5</i> <i>FOR:</i>	TATTCTGGTGTCCCTAGGCGTAG
<i>rrn5</i> <i>REV:</i>	ATCCTGGCGTCGAGCTATTTCC
<i>rrn23</i> <i>FOR:</i>	GTTCGAGTACCAAGGCCTAC
<i>rrn23</i> <i>REV:</i>	CGGAGACCTGTGTTTGGT
<i>rrn4.5</i> <i>FOR:</i>	GAAGGTACGGCGAGACGAGCC
<i>rrn4.5</i> <i>REV:</i>	GTTCAAGTCTACCGGTCTGTTAGG
<i>trnA</i> <i>REV:</i>	CTACCAACTGAGCTATATCCCC
<i>trnI</i> <i>REV:</i>	CTACCACTGAGCTAATAGCCC
<i>trnR</i> <i>REV:</i>	CTAATCCTCTGAGCTACAAGCCC
<i>rpS3</i> <i>FOR:</i>	AAGGAACTCTGCCTTCTGATCCA

rpS3 *REV:* GAAGATAAACCCGAAGAGTCGAAG
rpS4 *FOR:* TGTGTGAAGAGTCAGATGGTTGGC
rpS4 *REV:* TCTTAGAAACCAATCACGCTCCGT
rpS7 *FOR:* AAGCTCTATTGCCTCTGCCATT
rpS7 *REV:* GCAATACGTGGAGTAACTCCGAT
rpS8 *FOR:* ATGGGGAAAGACACCATTGC
rpS8 *REV:* TCCGCCGATTCTTTAGTC
rpS18 *FOR:* ATCCAAGCGATCTTTCGTAGG
rpS18 *REV:* GGAGTCGACTCACTCTTCAA
psbA *FOR:* AAGCCGCCGGATCTTCAAC
psbA *REV:* CTAAGCCTTGTCCTCAAAGC
rbcL *FOR:* CGTTGGAGAGACC GTTCTT
rbcL *REV:* CAAAGCCC AAAGTTGACTCC

Real Time PCR:

Gc16a *qRT FOR:* GAATCAGCTGCAAGTTGTGTCT
Gc16a *qRT REV:* GGTTTAGGTCTTCCACTCAA
Gc16b *qRT FOR:* GCAATCCTCTCGATTCCCTA
Gc16b *qRT REV:* GCAACGGTGACTTCATAGGAC
Actin_8 *FOR:* GCAGCATGAAGATTAAGGTCGTG
Actin_8 *REV:* TGTGGACAATGCCTGGACCTGCT

3. Results

3.1. Lutein Production

3.1.1. Generation of Lutein construct

As the first step for plant-type pigment production in *Synechocystis*, the pigment lutein was chosen. *Synechocystis* accumulates one carotene (β -carotene) and three xanthophylls (zeaxanthin- lutein's isomer, echinenone and myxoxanthophyll). In *Synechocystis* only β -ring carotenoids are formed while in *Arabidopsis* both β - and ϵ -rings. Since *Synechocystis* lacks the ϵ -ring-forming enzymes, ϵ -cyclase (AT5G57030-LCYe) and ϵ -hydroxylase of *Arabidopsis* were introduced into *Synechocystis* (AT3G53130-CYP97C1). Even though β -carotenes are present in *Synechocystis*, *Arabidopsis* enzymes for β -carotene production were also considered for two reasons (i) the β -cyclase in *Synechocystis* has not been identified, yet (even though it is presumed to be *cruA* (Maresca et al. 2007)) and (ii) the plant-type ϵ -hydroxylase (CYP97C1) has been proven to interact with plant-type β -hydroxylase (CYP97A3) in order to form lutein from α -carotene (Quinlan et al. 2012). The strategy for the introduction was a 'one step approach', where all four *Arabidopsis* genes, cloned in one cassette, were transformed into the *Synechocystis* genome at once.

In order to overcome a potential codon usage bias, *Arabidopsis* genes were synthesized (OptimumGeneTM algorithm, GenScript, U.S.A.) for efficient expression in *Synechocystis*. The codon usage was optimized by eliminating RNA secondary structures, by removing cryptic splicing sites and by adjusting the GC content (Kane 1995; Lindberg et al. 2010b). To this end, the codon adaptation index (CAI) of each gene was significantly increased, for *LCYe* from 0.62 to 0.85, for *LCYb* from 0.63 to 0.86, for *CYP97A3* from 0.58 to 0.87 and for *CYP97C1* from 0.59 to 0.87, respectively.

For efficient transcription, strong and endogenous promoters were used and placed in front of each of the codon-optimized *Arabidopsis* gene. *Synechocystis* P_{psbAII} , the promoter of *psbAII* gene (Mohamed, 1992), and P_{rbcL} , *rbcL* gene promoter have been used alternatively for each gene and all important elements in those promoters for efficient transcription have been

kept; the motifs for σ^{70} promoters (-35 and -10), ribosome binding sites (Shine Dalgarno) and transcription activation sites. Moreover, the cis-element in the P_{psbAll} promoter responsible for light regulation has been preserved. The *Arabidopsis* genes were designed to be integrated into a neutral site in the *Synechocystis* genome (*sly0168*) via homologous recombination (Figure 7) (Kunert et al. 2000). The final plasmid carrying all these parts is referred to in the following as pKS-LUT (Figure 7).

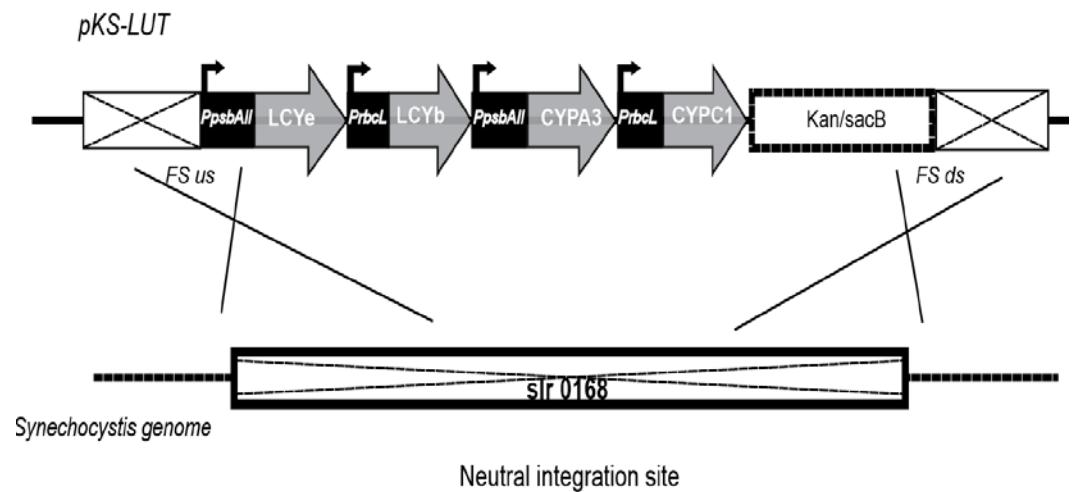


Figure 7. Schematic representation of pKS-LUT, which carries four genes to complete the lutein synthesis pathway in *Synechocystis*.

Genes, which are necessary for the conversion of lycopene to lutein (*LCYe* (ϵ -cyclase), *LCYb* (β -cyclase), *CYP97A* (β -hydroxylase) and *CYP97C* (ϵ -hydroxylase)) are depicted as gray arrows. The two promoters P_{psbAll} and P_{rbcL} are highlighted by black boxes. A double selection cassette was included in the construct for selection (*Kan/sacB*) and the genes were designed to be integrated into the neutral site in the *Synechocystis* genome (*sly0168*) by double homologous recombination (sites mediating homologous recombination are indicated by *FS us* and *FS ds*).

3.1.2. Color complementation assay in *E.coli*

In order to verify whether the genes coding for components of the lutein pathway can be successfully expressed in a bacterial host, a color complementation test was performed in *E. coli*. A lycopene-producing *E. coli* strain was used to test for heterologous expression of the genes involved in lutein biosynthesis. To this end, the plasmid pKS-LUT was transformed into *E.coli* cells containing the plasmid pAC-LYC, which - when transformed alone - accumulates lycopene, the substrate for the first enzymatic step in the plant-specific lutein synthesis pathway. In case all genes are expressed correctly

and all gene products are functional, *E.coli* (+pAC-LYC/+pKS-LUT) should convert lycopene into lutein. In fact, the color change (pink to yellow/orange) of the colonies in this assay was a first indication for efficient conversion of lycopene into carotenoids (Figure 8.A).

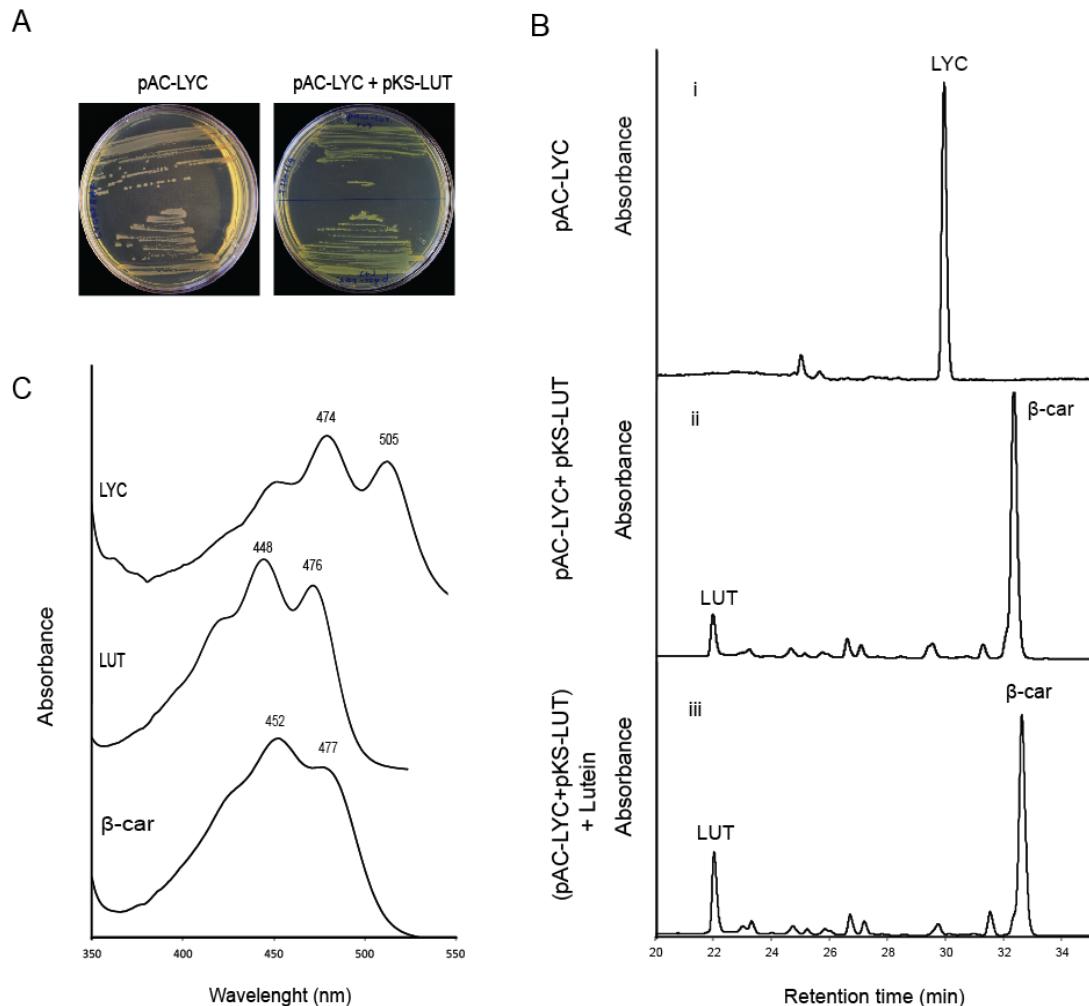


Figure 8. Heterologous expression of lutein pathway genes in *E.coli*.

A. Color complementation assay using a lycopene producing *E. coli* background strain (pAC-LYC, left panel). *E. coli* (pAC-LYC) cells after transformation of pKS-LUT (right panel) grown on LB plates. B. HPLC analyses of carotenoids extracted from *E.coli* cells containing pAC-LYC (profile i), pAC-LYC+pKS-LUT (profile ii) and the pAC-LYC+pKS-LUT, added exogenously lutein standard (profile iii). C. The absorption spectra of the major peaks from 5.B are shown. LYC: Lycopene, LUT: Lutein, β-car: β carotene.

To further determine which carotenoids were produced, HPLC analysis was performed. The elution profile of pAC-LYC+pKS-LUT yielded two main carotenoids, β-carotene and lutein, as indicated by the HPLC retention time

(Figure 8.B) and absorption spectra (Figure 8.C) of these carotenoids. Although β -carotene amounts were higher than those for lutein, the results indicated that all plant genes are properly expressed and that the enzymes necessary for lutein production are functional in a bacterial host.

3.1.3. Generation of *Synechocystis* mutants producing Lutein

The plasmid pKS-LUT, harboring the genes for lutein pathway, was used to transform the glucose tolerant wild type *Synechocystis* strain and the mutants were called *SynLUT*. The first attempt to produce lutein in *Synechocystis* was carried out with the wild-type strain. The transformants were selected on agar medium supplemented with kanamycin and grown under normal light conditions, but supplemented with glucose. The obtained mutants were re-streaked on increasing antibiotic concentration to achieve complete segregation.

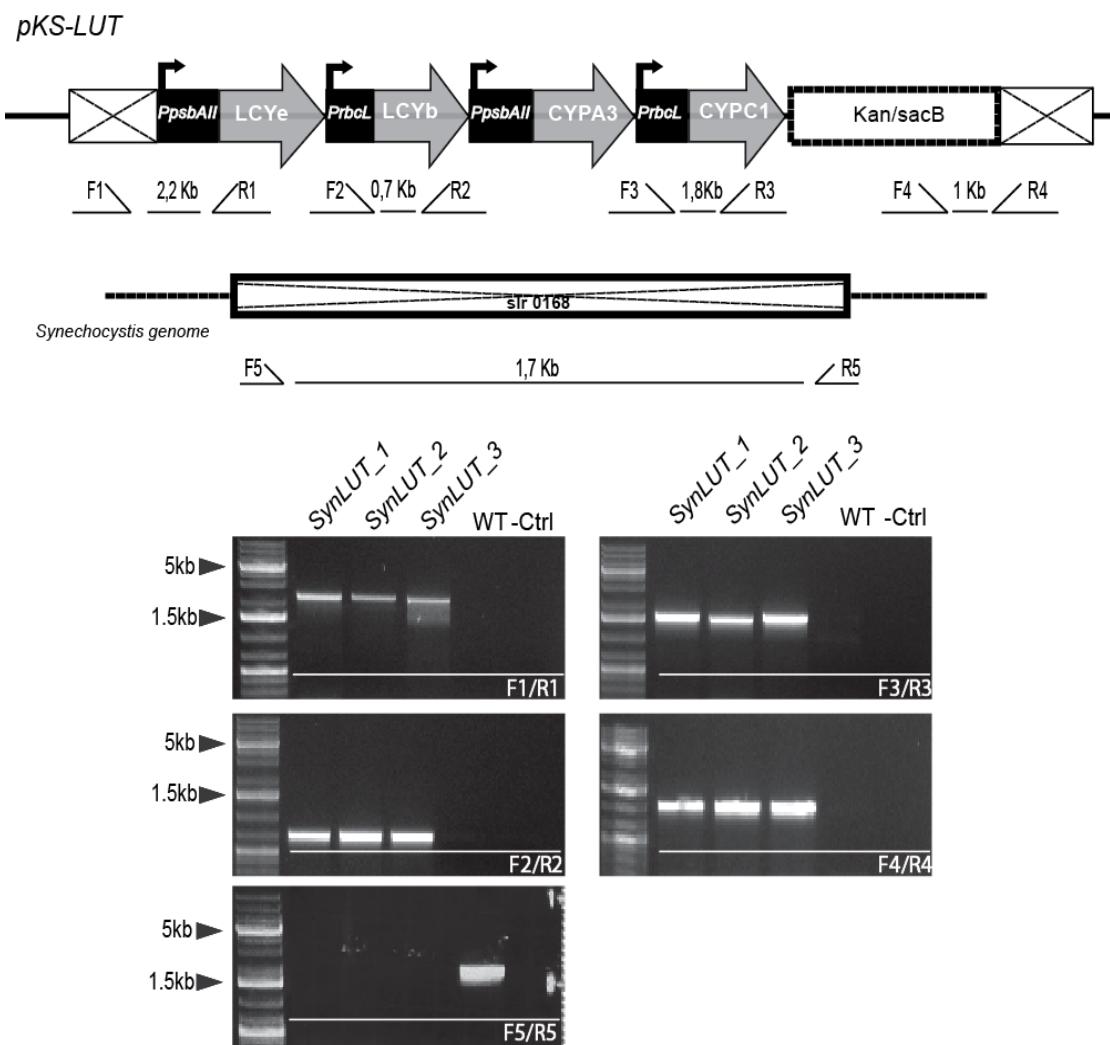


Figure 9. Analysis of three different colonies of *SynLUT* mutant strains.

A. Five different primer pairs were used for the analysis of the mutants and the expected sizes for each product are indicated. B. PCR analysis to test for the insertion into *Synechocystis* genome and for complete segregation of *SynLUT* mutants.

To verify the correct genome integration and the complete segregation of the mutations, genomic DNA was extracted from the *SynLUT* mutant and wild-type strain and subjected to PCR analysis (Figure 9). The primer pairs F1/R1 and F4/R4 were used to show that the construct was integrated into the correct genome position and also to verify that the first gene and the selection cassette were present. Amplicons of 2,2 Kb and 1 Kb were generated only in the mutants, while in the wild type no product formation could be observed. The primer pairs F2/R2 and F3/R3 were used to prove that the whole cassette was integrated. Amplification of the appropriate bands was achieved only in the mutants, confirming that the whole cassette was indeed integrated into the *Synechocystis* genome (Figure 9). Furthermore, complete segregation could be confirmed by PCR with primer pair F5/R5, which indicated that all the wild type copies of the *slr0168* gene were lost in the mutant strains.

3.1.4. Pigment analysis of *SynLUT* mutants

In order to determine whether the lutein pathway genes from *Arabidopsis* could be expressed in *Synechocystis* and whether they are able to produce lutein as the final product, HPLC analysis was performed with *SynLUT* mutants. The elution profile of the wild type strain was used as control. All cultures were grown under normal light conditions until an OD₇₃₀ of 0.5 and pigments were extracted with acetone and directly subjected to HPLC analysis. As shown in the HPLC chromatogram of figure 10.A no lutein could be detected in the mutants. This result suggests that the heterologous expression of lutein pathway genes from *Arabidopsis* was not sufficient to produce lutein in *Synechocystis*.

Lutein is the stereoisomer of zeaxanthin, so they are chemically similar (only an ε-ring in one end is different) with almost the same retention time. Taking into consideration that zeaxanthin is abundantly produced in *Synechocystis* and is one of the four major carotenoids, we can assume that it is rather difficult to detect lutein in case only traces of it are produced.

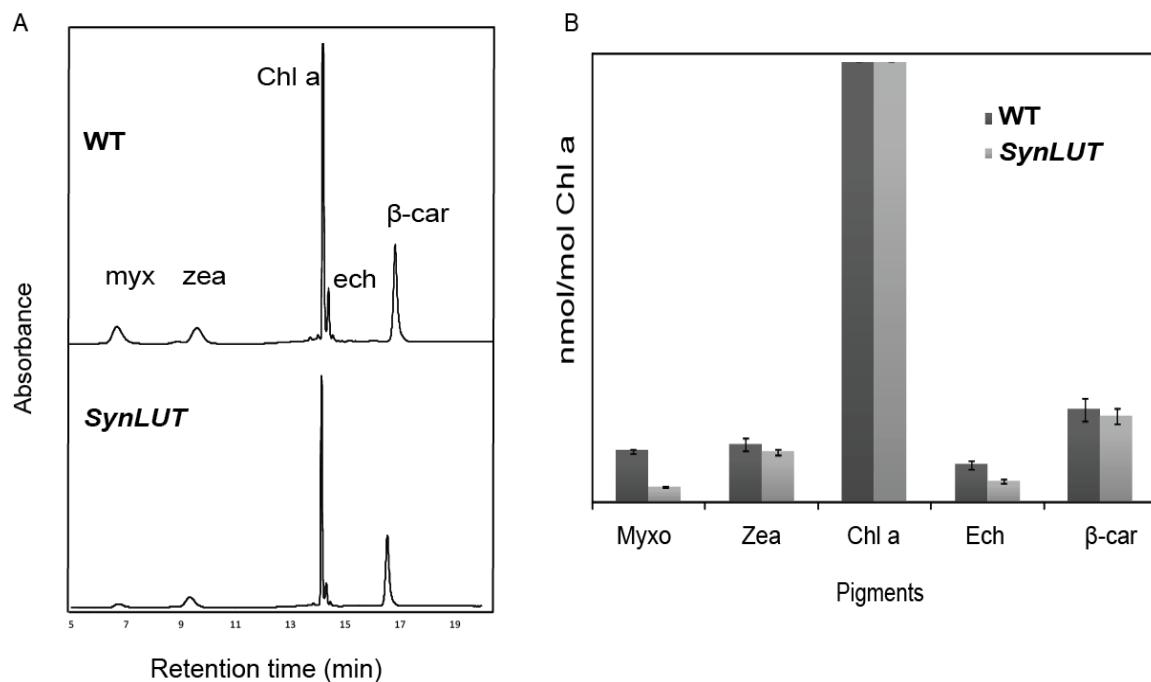


Figure 10. Pigment profile for *SynLUT* mutants.

A. HPLC analysis of pigments extracted from wild type and *SynLUT* mutants. All the major carotenoids are indicated; myx: myoxanthophyll, zea: zeaxanthin, Chl a: Chlorophyll a, ech: echinenone and β -car: β -carotene. B. Quantification of all the major carotenoids in *SynLUT* mutants compared to the wild type strain pigments, adjusted to Chlorophyll a level.

Another point to consider was whether the endogenously produced pigments were affected in *SynLUT* mutants. The production of extrinsic pigments may affect the balanced production of endogenous ones, particularly with regard to alteration of the substrate pool of lycopene. Since lycopene serves as a precursor of all carotenoids in *Synechocystis*, it could be depleted by the newly introduced pathway and affect the endogenous carotenoids. Quantification of all the main pigments from wild type and *SynLUT* mutant showed that indeed pigment amounts were altered (Figure 10.B). In *SynLUT* mutants the production of almost all pigments was decreased, with the most significant effect on myoxanthophyll and echinenone. Zeaxanthin and β -carotene were slightly reduced in the mutants. It was concluded that the synthesis pathways for zeaxanthin, β -carotene and lutein “compete” for lycopene as substrate and since no lutein was detected in *SynLUT*, lycopene seems to be favorably consumed in other synthesis pathways. For that reason, the disruption of competing pathways by mutagenesis of endogenous

enzymes was envisaged as an optimization strategy to produce lutein in *Synechocystis*.

3.1.5. Deletion of the putative β -cyclase (*sll0147*) from *Synechocystis*

To optimize lutein production in *Synechocystis*, the pathway producing the endogenous carotenoids β -carotene and zeaxanthin was disrupted in *Synechocystis*. In *Synechocystis*, the β -cyclase is presumed to be *sll0147* (Maresca et al., 2007), which was identified to be part of a new family of cyclases in photosynthetic bacteria like *Synechococcus* and *Chlorobium tepidum*. *cruA* from *Synechocystis* shares 65% identity with *cruA* from *Synechococcus*. An additional gene product *cruP*, which is a paralog of *cruA* could also function as a putative monoyclase in *Synechocystis*, but has not been proven to have this function, yet. The aim of this approach was to create *Synechocystis* mutants with no *cruA* activity, in order to reduce the amounts of carotenoids in this organism, especially β -carotene and zeaxanthin that are possibly affecting lutein production or detection. Furthermore, the knockout was supposed to be followed by an increase of lycopene production, which would redirect the reactions to lutein synthesis.

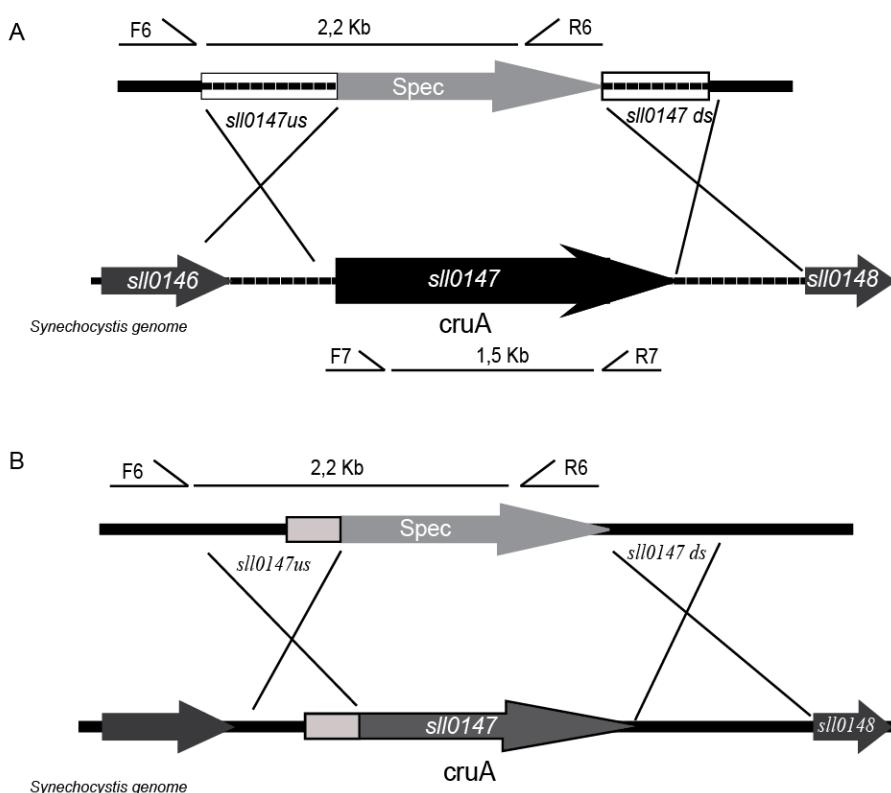


Figure 11. Scheme for Δ cruA mutant strains.

A. Δ cruA (Knock out) cassette. The whole *cruA* gene (*sll0147*) is completely replaced by a spectinomycin cassette. B. Scheme for *Synechocystis* Δ cruA (Knock down) mutants, where only a part of the promoter and the 5' end of the gene is replaced by spectinomycin. The sites of the primers for chromosomal integration and segregation are indicated.

Since β -cyclase is the gene responsible for lycopene conversion to other carotenoids, a complete knock out mutant might be not viable. For that reason two different constructs were designed to disrupt *cruA* in *Synechocystis*. In the first case, the entire *cruA* gene was deleted and replaced by a spectinomycin cassette, Δ cruA:Spec, generating the mutants Δ cruA_1/2, as shown in Figure 11.A, while only the last part of the promoter with the first part (5' end) of the gene were exchanged with spectinomycin (Δ cruA:5'Spec) in the second construct. In this case the important elements of the promoter (-35/-10 sequence) were replaced, but the predicted functional domain of the protein were kept intact and the mutant was called Δ cruA_3 (knock-down) (Figure 11.B).

3.1.6. Characterization of the mutants: Chromosomal segregation

Synechocystis wild type strain was transformed with the constructs Δ cruA:Spec and Δ cruA:5'Spec and the transformants were grown on plates supplemented with the respective antibiotic. It was only possible to obtain these mutants when grown under low light (5 μmol photons m^{-2} s^{-1}) and glucose was added to the medium. The correct integration of the construct was verified via PCR using the primer pair F6/R6 for the mutants, where a band of approximately 2,2 Kb was amplified. To check whether all the *cruA* copies in the genome were replaced, the primer pair F7/R7 was used, which amplify the endogenous *cruA* gene of *Synechocystis*. The Δ cruA mutants were not able to segregate completely, no matter how high the antibiotic concentration was (Figure 12.A).

Apparently, *cruA* cannot be completely absent from this organism, probably because of its important role, since the gene product is involved in the synthesis of all main carotenoids in *Synechocystis*. Although the mutants were not completely segregated, the Δ cruA mutants showed a severe growth defect, even under low light conditions (Figure 12.B).

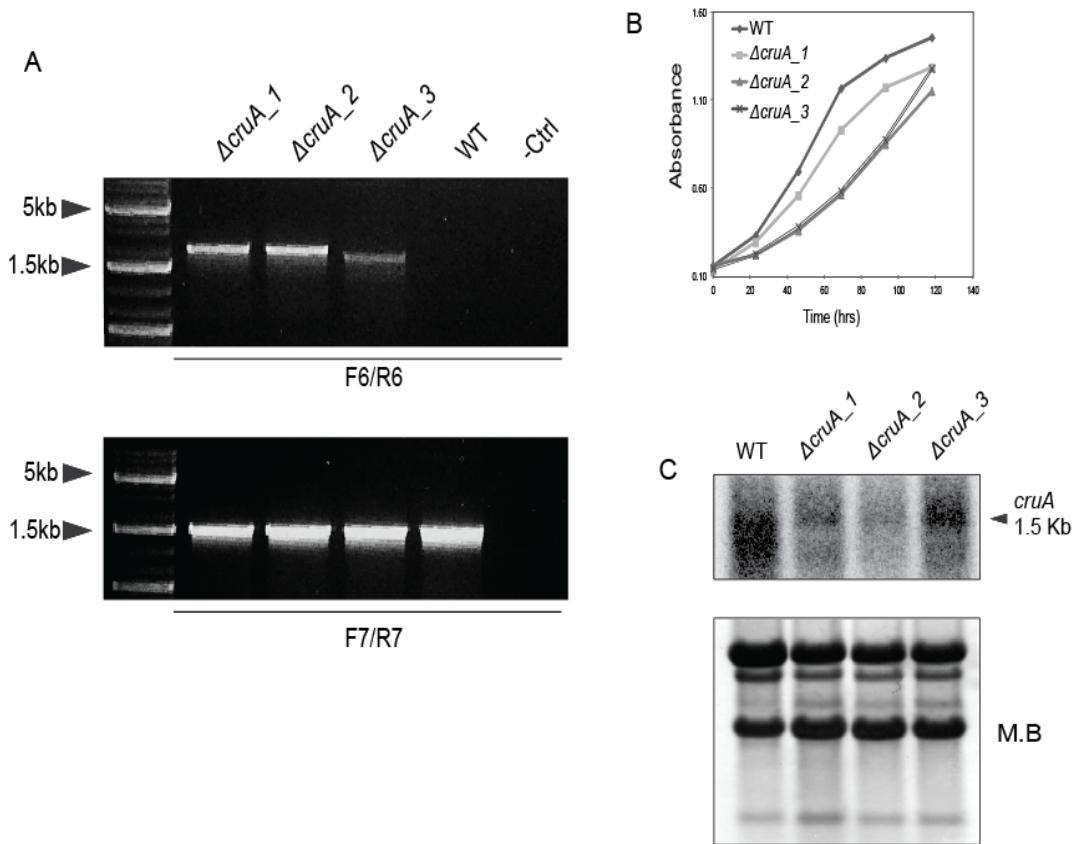


Figure 12. Characterization of ΔcruA mutants of *Synechocystis*.

A. Chromosomal integration of spectinomycin cassette in *sll0147* and incomplete segregation of the generated mutants. B. Growth rate analysis of wild type and ΔcruA mutants under photomixotrophic conditions. C. Northern blot analysis for the expression of *cruA* genes in wild type and ΔcruA mutants. $\Delta\text{cruA}_1/2$ were transformed with $\Delta\text{cruA}:\text{Spec}$ construct, while ΔcruA_3 with $\Delta\text{cruA}:5'\text{Spec}$. 10 μg of RNA were loaded in each lane and the membrane was stained with methylene blue (M.B.) and is shown as loading control. The probe was amplified with primers from Table 4 from Material and Methods.

In order to determine how the transcription of *cruA* is affected in ΔcruA mutants, Northern blot analysis was performed (Figure 12.C). Even though the mutants were not completely segregated, transcript amounts of *cruA* in the mutants were significantly reduced, with the more severe effect in the $\Delta\text{cruA}_1/2$ mutants (~80%). In ΔcruA_3 expression levels were reduced approximately to 50% of wild type levels, but were higher compared to levels of the other two mutants (Figure 12.C). Since a gradual reduction of *cruA* expression in the mutants could be observed, their pigment composition was examined in more detail.

3.1.7. Pigment analysis of *ΔcruA*

The pigments from both wild type and *Synechocystis ΔcruA* were analyzed by HPLC, according to the method described in chapter 2 (Färber et al. 1997). Interestingly, the elution profiles of the mutants showed an extra peak, which was identified as lycopene according to the absorption spectrum and the retention time (Figure 13).

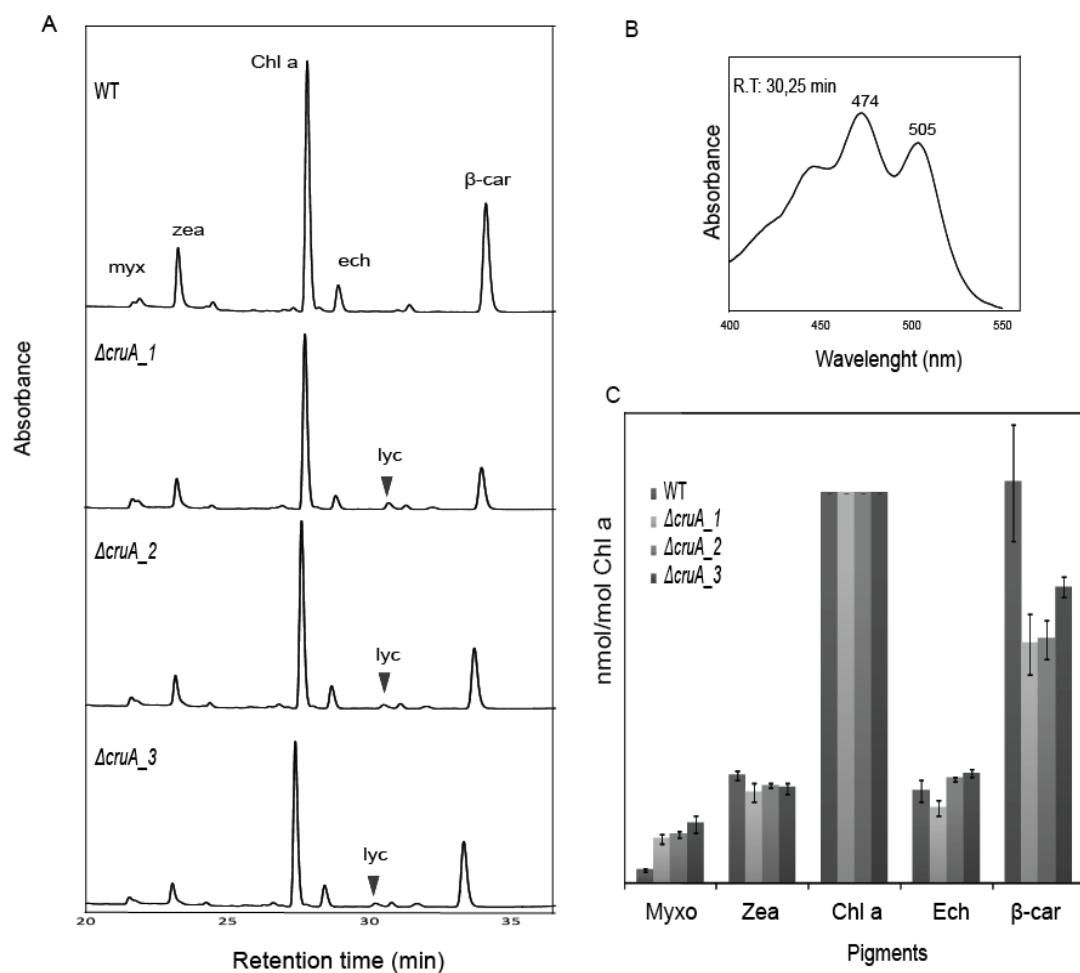


Figure 13. Pigment elution profile of *Synechocystis ΔcruA* mutants.

A. HPLC analysis of wild type (top profile) and $\Delta c\text{ruA}$ mutants grown under low light conditions. Pigments were extracted from the same amount of cells and immediately subjected for analysis. B. Absorption spectrum of the extra peak in $\Delta c\text{ruA}$ mutants, with retention time 30,25 min corresponding to lycopene. C. Quantification of all the major pigments in $\Delta c\text{ruA}$ and wild type strains. Myx: myoxanthophyll, zea: zeaxanthin, Chla: Chlorophyll a, ech: echinenone, β -car: β -carotene, lyc: lycopene.

This was the first indication that the pathway for the production of carotenoids is disrupted, since the precursor of this pathway accumulates in the mutants.

All the major pigments from Δ cruA mutants and wild type samples were quantified and it was evident that β -carotene and zeaxanthin levels were reduced in the mutants, suggesting that cruA is indeed a β -cyclase involved in the cyclization of lycopene for the production of β -carotene and zeaxanthin. In Δ cruA_1/2 mutants, the amounts of β -carotene were reduced to around 40% compared to the wild type strain. Δ cruA_3 was also affected, but not as severe as the other mutants; almost 25% less β -carotene was produced. Taking those results together, disrupting the carotenoid pathway in *Synechocystis* via cruA mutation leads to lycopene production.

3.1.8. Generation of *Synechocystis* mutants combining lutein pathway introduction and cruA disruption.

The next step was the combination of lutein introduction with the disruption of cruA. To this end, two different approaches were chosen. In the first approach, the SynLUT mutant strain was used as background to introduce the Δ cruA:Spec and Δ cruA:5'Spec cassettes, while in the second case, the *Synechocystis* mutant Δ cruA was used as background strain to introduce the pKS-LUT cassette. Transformants were only obtained for the first combination (SynLUT with Δ cruA:Spec), but not for the second approach. Transformants were plated on agar medium containing both kanamycin and spectinomycin for selection of positive clones. The medium was also supplemented with glucose and all mutants were grown under normal light conditions.

The integration of the cassette into the genome of *Synechocystis* was verified by PCR analysis, using the primer pair F8/R8, which specifically amplified the spectinomycin gene from the cassette. An amplicon of the expected size was (1,5 Kb) only detectable in the mutants, which indicates that the cassette was present in SynLUT mutants and integrated in the correct genome position. The complete segregation of the mutants was verified by the primer pair F9/R9 (Figure 14). The fact that the mutant lost all cruA copies is interesting, because when cruA was disrupted (Δ cruA), the mutants were not able to segregate (Figure 12.A). This observation can be an indication that the presence of lutein pathway genes from *Arabidopsis* leads to the production of other carotenoids or enhances the production of existing ones.

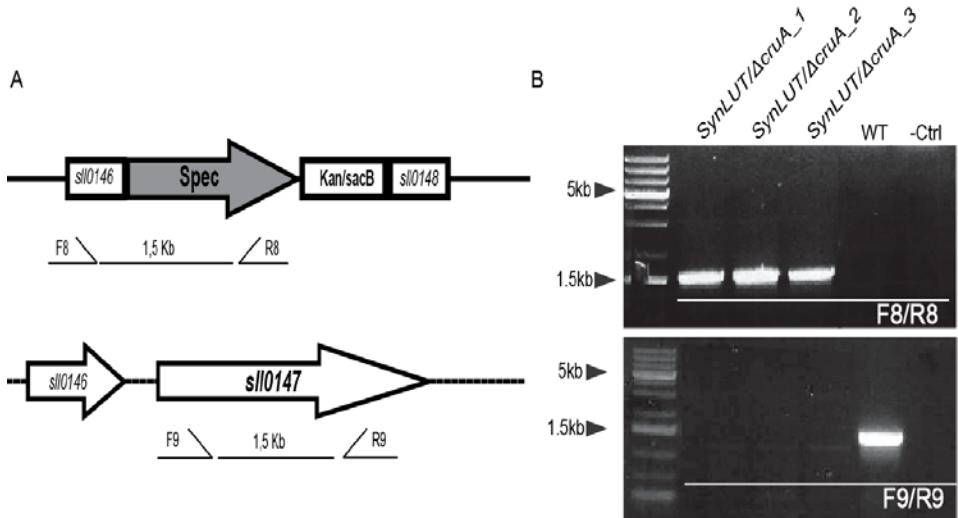


Figure 14. Analysis of *SynLUT/ΔcruA* mutant strains.

A. Schematic presentation of the mutants with the annealing sites of the primers used for genotyping. B. Correct segregation of the generated mutants with two primer pairs.

To identify lutein, HPLC analysis was performed for the wild type, *SynLUT* mutant and *SynLUT/ΔcruA* strains (Figure 15). A new peak could be detected in *SynLUT/ΔcruA* extracts that was neither present in wild type nor in *SynLUT* extracts. The retention time of this peak was 8 min, just one minute prior to zeaxanthin. To verify that this peak corresponds to lutein, a lutein standard was added to the same sample (4 pmol). Indeed, the peak could be identified as lutein, since the area of the peak of interest was increased by the external addition of lutein (Figure 15.A, B). Thus, the introduction of the *ΔcruA*:Spec cassette into *SynLUT* mutants leads to the production of lutein, even though the amounts were low.

Additionally, the endogenous pigments of *Synechocystis* were quantified, to control if there was any alteration or even restoration of the amounts of them. As shown in figure 15, the amounts of myoxanthophyll were increased in *SynLUT/ΔcruA* mutants, compared to the myoxanthophyll amounts in *SynLUT* (from 30% to 70%), but not reaching the wild type concentration. Echinenone was still less expressed in *SynLUT/ΔcruA*, like in *SynLUT*, with only a small increase (7%) after the expression of lutein in *Synechocystis*. Interestingly, zeaxanthin that previously was reduced in *SynLUT* mutants by 13%, accumulate to the levels of wild type in *SynLUT/ΔcruA*. β -carotene is also produced differently in these mutants, in *SynLUT* and *SynLUT/ΔcruA*

levels were reduced by 8% and 22%, respectively, compared to wild type. The drop of β -carotene amounts was around 40% when *cruA* is disrupted, but β -carotene synthesis is increased, when lutein genes are introduced, by around 20% compared to Δ *cruA* mutants. Taking all the results together, lutein could be produced in *Synechocystis*, although in low amounts only, in *SynLUT*/ Δ *cruA* mutants, where the β -cyclase from *Synechocystis* was disrupted and the lutein pathway genes from *Arabidopsis* were introduced.

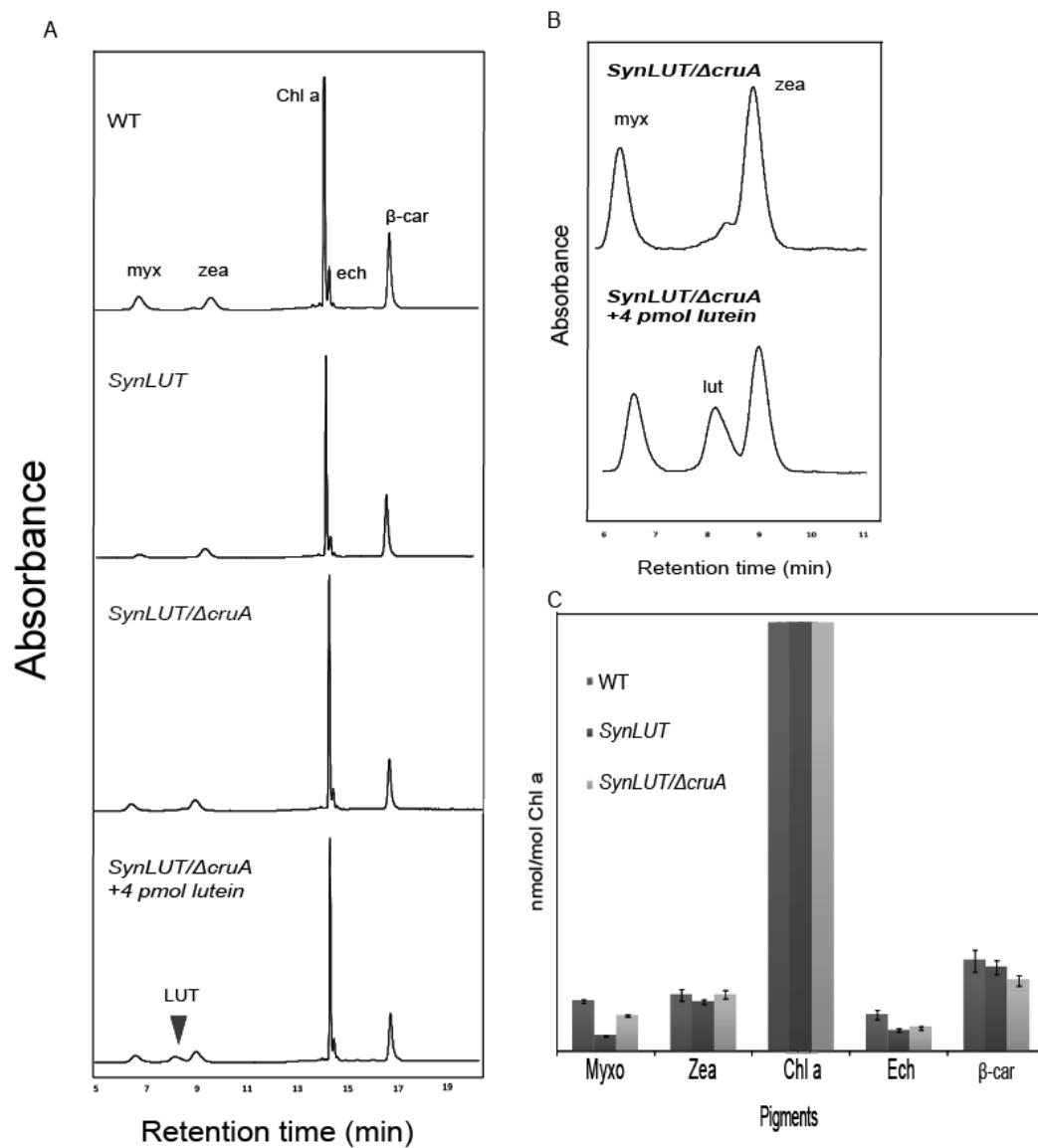


Figure 15. Pigment analysis of *Synechocystis* mutants.

A. Pigment elution profiles of *Synechocystis* mutant (*SynLUT*, *SynLUT/ΔcruA*) and wild type. The last profile shows the sample with the external addition of lutein (4 pmol). B. Part of the chromatogram, 6 to 11 min, from myoxanthophyll until zeaxanthin peaks shown, for sample *SynLUT/ΔcruA* and *SynLUT/ΔcruA* plus 4 pmol of lutein. C: Quantification of the four major

carotenoids in *Synechocystis* mutants and wild type strains referring to the HPLC analysis from A. The normalization was done with respect to chlorophyll a amounts.

3. 2. Violaxanthin Production

3.2.1. Introduction of Zeaxanthin Epoxidase from *Arabidopsis* into *Synechocystis*

Since cyanobacteria do not produce the plant-type pigment violaxanthin, *Synechocystis* mutants were generated carrying the Zeaxanthin epoxidase (*ZEP*) gene, which converts zeaxanthin to violaxanthin through antheraxanthin. *ZEP* consists of a short-chain dehydrogenase/reductase (SDR) domain, which includes an NAD and FAD binding site and a Fork_head Associated domain (FHA) in the C-terminal part of the protein and contains a binding site for phospho-amino acids. The protein is 667 amino acids long with a MW of 73 kDa.

For heterologous expression of *ZEP* in *Synechocystis*, a plasmid was generated carrying *Arabidopsis ZEP* gene. The nucleotide sequence of *ZEP* (AT5G67030) was codon optimized (*AtZEP_opt*) for efficient translation in the host organism (OptimumGeneTM algorithm, GenScript, U.S.A.) (Kane 1995; Lindberg et al. 2010b) and yielded a higher Codon Adaptation Index (CAI) value (0.63 to 0.87). The predicted transit peptide coding sequence (59 aa according to ChloroP) was not included. In addition, a His tag was fused to the N-terminus.

To generate the *AtZEP* expressor strain, *AtZEP_opt* was placed under control of the native and strong promoter *psbAII* and the *rcbL* terminator. Moreover, a double selection cassette (*nptI-sacB*) was included in the construct followed by a 450 bp DNA sequence of the *AtZEP* 3' end (Cai & Wolk 1990). This arrangement allows removing the selection cassette from the genome by a negative selection strategy and facilitates further transformation steps (Viola et al. 2014). The whole construct was designed to be integrated into the neutral site of *Synechocystis* genome, *slr0168* (Figure 16). The final vector was named pKS-ZEP_opt. The glucose-tolerant wild type strain of *Synechocystis* was used for the transformation and the mutant was named *SynZEP*.

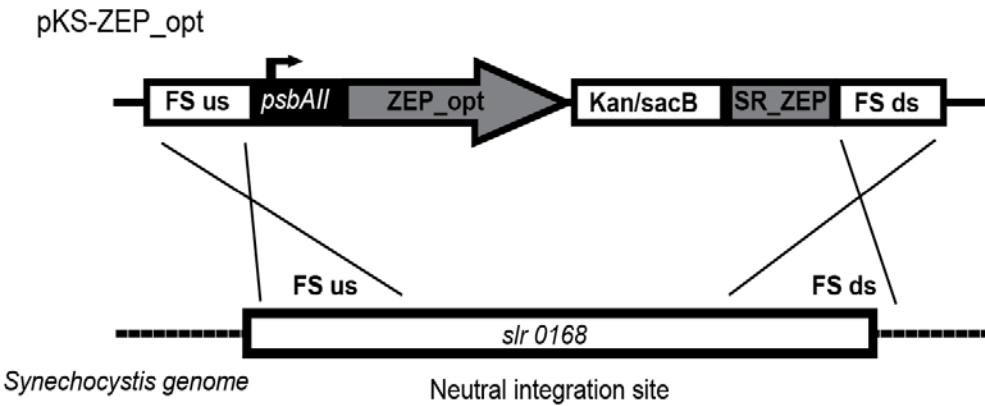
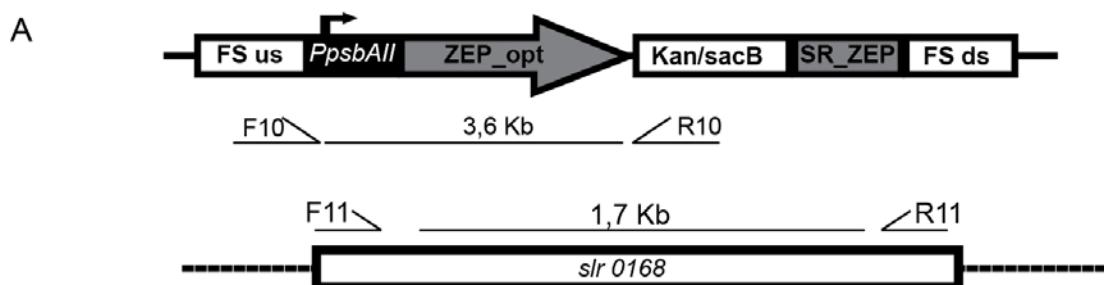


Figure 16. Schematic representation of the integration cassette into the genome of *Synechocystis*.

FS: Flanking sequence for homologous recombination, P_{psbAll} : psball promoter, *ZEP_opt*: optimized sequence of *Arabidopsis ZEP*, *kan/sacB*: double selection cassette for kanamycin and *sacB*, *SR_ZEP*: selection region for double recombination (450 bp from 3' end of *ZEP_opt*).

3.2.2. Genetic characterization of *SynZEP* mutants

The integration of the synthetic gene into the neutral genome site was confirmed by PCR with primers designed to amplify the transition of the integration site, with primers F10/R10 (Figure 17.A,B). The complete segregation of the mutants was confirmed by PCR analysis (primer pair F11/R11). An amplified band was only present in wild type strain and not in the mutants, which indicates that there was a complete segregation for the mutants *SynZEP*. The growth rate under low light conditions of the completely segregated *SynZEP* did not differ significantly compared to the wild type (Figure 17.C), indicating that the gene product is not toxic.



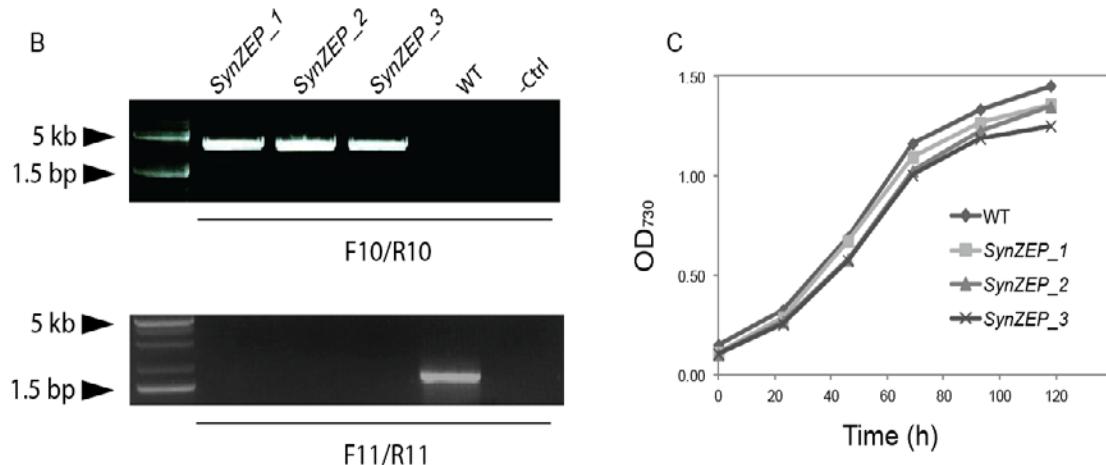


Figure 17. Characterization of *SynZEP* mutants.

A. Schematic presentation of the genetic modification on *Synechocystis* sp. PCC6803 genome. B. Integration of the *ZEP* gene in *Synechocystis*, with primers on the genome and the introducing gene (F1/R1). C. Complete segregation of the *SynZEP* (1-3) mutants, using primers flanking the integration platform (F2/R2) on the genome. D. Growth curve of *SynZEP* mutants.

To investigate whether transcripts of the synthetic *ZEP* gene accumulated in *Synechocystis* mutants, Northern blot analysis was performed. Three independent lines of *SynZEP* mutants were tested using an *AtZEP*-specific probe (Figure 18). A transcript at the expected size was detected only in the mutant samples, indicating that the synthetic *ZEP* gene was actually transcribed. An additional band with a size of about 500 bp could be detected in the mutants, which could be a degradation product of the main transcript. This result confirms that the plant-type *AtZEP* gene could be expressed successfully in *Synechocystis*. Protein levels of *AtZEP* were analysed by immunoblot assays using an antibody against the His tag. However, it was not possible to detect any protein (data not shown).

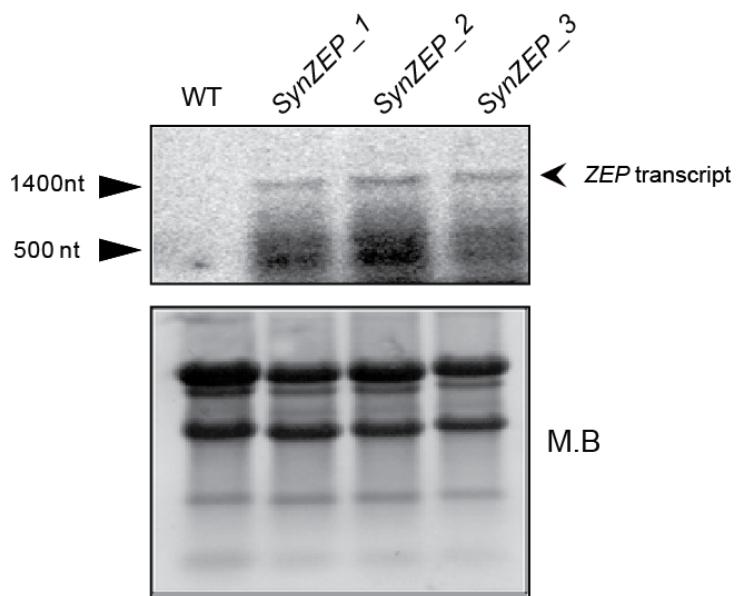


Figure 18. Transcriptional analysis of the *SynZEP* mutants.

RNA was isolated from wild type and 3 independent lines of the mutants and fractionated on denaturing RNA gels. A gene-specific primer labeled with ^{32}P was used for the detection of the transcript. Staining of the membrane with methylene blue (M.B) was used as loading control.

3.2.3. Pigment profile of the mutants

In order to determine the pigment composition in *SynZEP*, High Performance Liquid Chromatography (HPLC) was performed using acetone extracts of pigments from equal amounts of wt and *SynZEP* cells (Figure 19). The first profile shows the typical carotenoid content of wild-type strain, with Chlorophyll *a* and the four major carotenoids to be present, myoxanthophyll, zeaxanthin, echinenone and β -carotene. In case of a functional *AtZEP* which is able to use zeaxanthin as substrate and to produce violaxanthin, one would expect a peak shortly before the myoxanthophyll peak in the elution profile (Qingfang He et al. 1999). However, no peak corresponding to violaxanthin was detectable. Interestingly two new peaks which might be carotenoids not present in the wild type could be identified in the *SynZEP* chromatogramm (Figure 19). These results demonstrate that expression of *AtZEP* in *Synechocystis* resulted in newly synthesized carotenoids but does not produce violaxanthin.

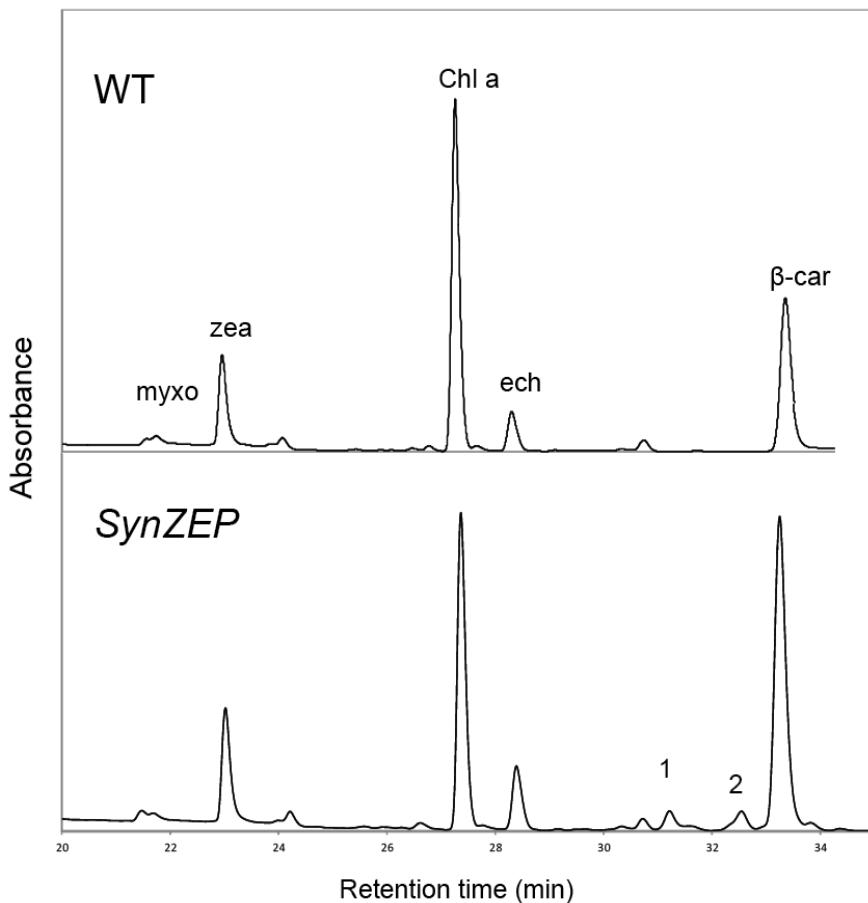


Figure 19. HPLC elution profile of pigments from *Synechocystis* wild type strain (upper panel) and the *SynZEP* mutants (lower panel).

Cultures were grown in mixotrophic conditions under continuous low light and pigments were extracted from equal amount of cells. Peaks corresponding to the abundant pigments are indicated; myxo, myoxanthophyll, zea, zeaxanthin, Chl, chlorophyll a, ech, echinenone, β -car, β -carotene. Pigment assignment is based on retention time and absorption spectra.

Newly synthesized carotenoids in *SynZEP* had a retention time of 30.5 and 32.5 min, respectively, and eluted in the chromatogram just before β -carotene. This indicates that these carotenoids might be carotenes and not xanthophylls. Although the desired final product could not be detected in the mutants, the results indicate that *AtZEP* is functional. In order to clarify the identity of the pigments corresponding to peak 1 and 2, their absorption spectra were recorded (Figure 20). The pigment corresponding to peak 1 showed maxima at 441 and 471 nm, while pigment corresponding to peak 2 had maxima at 402 and 427 nm. Therefore, the products of *AtZEP* in *Synechocystis* are carotenoids and not xanthophylls as expected by the action of an epoxidase.

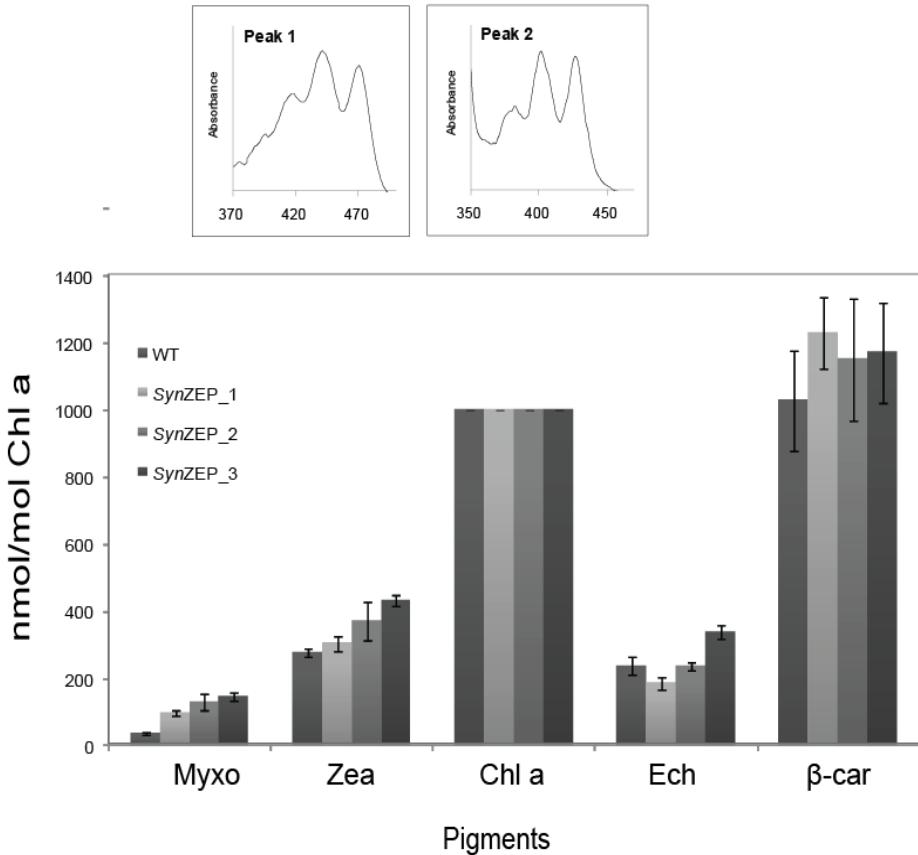


Figure 20. Pigment analysis of *SynZEP* mutants.

Absorption spectra of peak 1 and peak 2 from Figure 19 and pigment quantification in *SynZEP* mutants compared to *Synechocystis* wild-type strain. Three independent lines (three replicates of each) were analyzed with HPLC and the quantification was adjusted to chlorophyll amounts (in cooperation with Peter Jahns). Averages and SDs are provided.

To analyze whether the abundance of endogenous pigments was affected by the presence of *AtZEP*, pigments were quantified and referred to the chlorophyll a amount (Figure 20). Zeaxanthin and myxoxanthophyll were more abundant in the three mutant lines tested, while echinenone and β -carotene levels were not significantly affected.

3.3. Gc16 characterization

3.3.1. Greencut proteins

Phylogenomic comparative analyses have revealed a set of proteins, which are only found in eukaryotic organisms from the green lineage and might be involved in photosynthesis (Merchant et al. 2007). One of those GreenCut (GC) proteins with unknown function, Gc16A/CGL20A (AT2G17240), is conserved in algae and plants, but not in cyanobacteria. A second gene product (AT3G24506) found in the *Arabidopsis* nuclear genome, named Gc16B in the following, shares high similarity (92%) and identity (67%) in a sequence alignment with Gc16A (Figure 21).

Gc16A consists of 140 amino acids and contains a predicted (ChloroP) transit peptide sequence (cTP) of 59 amino acids (ChloroP, Emanuelsson et al., 1999), while Gc16B is 149 amino acids long and contains a predicted cTP of 67 amino acids. The calculated molecular weight of mature Gc16A and Gc16B are 16.19 kDa and 17.06 kDa, respectively. No transmembrane or functional domain is predicted for Gc16A and Gc16B.

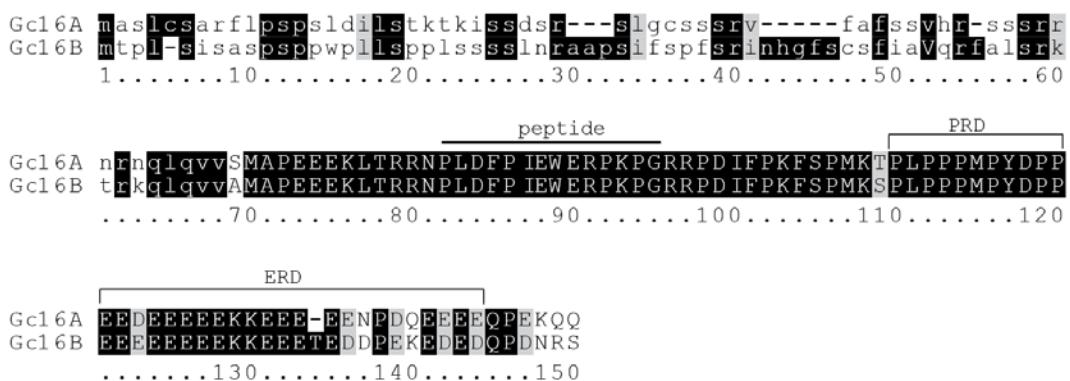


Figure 21. Sequence alignment of the two Gc16 proteins from *Arabidopsis thaliana*.

Sequences of Gc16A and Gc16B were aligned with MAFFT and formatted with Boxshade for similarities. Conserved amino acids to more than 50% were shaded with black/grey boxes. Predicted chloroplast targeting sequences (ChloroP) are shown in lowercase letters. PRD: Proline rich domain, ERD: Glutamic acid rich domain. The sequence used for antibody production is shown.

Expression analysis data (ATTED-II) showed that GC16A/B are co-regulated with genes, whose products are involved in translation, like the ribosomal

proteins L35, L28, S27, but also with genes, whose products have RNA-binding capacities (L34, S20 and S17). Furthermore, both genes are highly co-regulated with a gene, coding for SWIB/MDM2 (Bennett-Lovsey et al. 2002), which is involved in plastid transcription (Table 6). Thus, Gc16A and B are chloroplast proteins conserved in the green lineage and - according to co-regulation studies - might play a role in plastid transcription or translation.

Table 6. Coexpression of Gc16 proteins with other proteins in *Arabidopsis*.

The locus of each gene, the average correlation to query loci (Ave. cor to query), the subcellular localization (C: Chloroplast, M: Mitochondria, Y: cytosolic, O: others) and the function of each gene (Function) are indicated. All information were obtained from ATTED-II (atted.jp).

Gc16A			
Locus	Ave. cor to query	Target	Function
At2g14880	0.63	C,C	SWIB/MDM2 domain superfamily protein
At3g24506	0.5	C,C	Gc16B
At1g11430	0.57	C,C	plastid developmental protein DAG, putative
At1g29070	0.56	C,C	Ribosomal protein L34
At3g15190	0.56	C,C	chloroplast 30S ribosomal protein S20, putative
At2g24090	0.56	C,M	Ribosomal protein L35
At2g33450	0.56	C,C	Ribosomal L28 family
At1g66820	0.53	O,N	glycine rich protein
At1g79850	0.56	C,C	ribosomal protein S17
At5g40950	0.54	C,C	ribosomal protein large subunit 27
At3g56910	0.56	C,C	plastid-specific 50S ribosomal protein 5

Gc16B			
Locus	Ave. cor to query	Target	Function
At2g17240	0.5	C,N	Gc16A
At2g14880	0.43	C,C	SWIB/MDM2 domain superfamily protein
At2g38140	0.44	C,C	plastid-specific ribosomal protein 4
At4g32915	0.42	C,C	regulation of translational fidelity
At2g21580	0.43	O,N	Ribosomal protein S25 family protein
At2g47840	0.39	C,C	Uncharacterised conserved protein ycf60
At5g20130	0.41	C,C	unknown function
At5g52960	0.43	C,C	unknown function
At5g17710	0.41	C,C	Co-chaperone GrpE family protein
At5g51960	0.41	O,Y	unknown function
At2g33430	0.41	M,C	differentiation and greening-like 1
At2g20450	0.41	O,Y	Ribosomal protein L14

To find orthologous proteins in other species, the whole protein sequence of Gc16 was blasted against a non-redundant protein sequence database (NCBI

server). The sequence alignment confirmed that the protein is conserved in many different species ranging from green plants (*Camelina sativa*, *Populus trichocarpa*, *Vitis vinifera*, *Oryza sativa* and *Zea mays*) to green algae (*Chlamydomonas reinhardtii*, *Ostreococcus lucimarinus*). The highest conservation between these two proteins is from 60-110 amino acid, which contains a proline (PRD) - and glutamic-acid-rich (ERD) domain (Figure 22, 22.A).

A

<i>Arabidopsis-Gc16A</i>	M----ASLC[SARFLPSPSPLSDI[L-----STKTKISSSDSR--SLGCSSSR-----
<i>Arabidopsis-Gc16B</i>	M----TPLSI[SASPSPSPWPLI-----S[PPLSSSSSLNRAAPSISFPFSRINHGFSCS-----
<i>C.sativa</i>	MA----STLC[SVRFLPSPSPLDVL-----G-KTKISSSDSR--SLVSSSSR-----
<i>O.sativa</i>	M----ASWC[LSA-----SASAAAPVAAAAAAAGPL-----GASVA-----
<i>Z.mays</i>	M----AFPWCLSA-----PAA-----VP[PVAPAPGTF-----GVSAS-----
<i>P.trichocarpa</i>	MASIYLSIC[TSK-----PIA-----APPKSAGRTTASASSSSCFK-----SQFCG-----
<i>V.vinifera</i>	M----ASFSS-----SIL-----MPPASFTSKTLELGYASSSSL-----SCRCSLILT
<i>C.reinhardtii</i>	-----MRA-----QFA-----RPLGLAHRSTVVVKPPSG-----
<i>O.lucimarinus</i>	-----MRAA-----ATTMI-----GRSTTTGARAR-----ARAAPAR-----ASTRK-----
	1.....10.....20.....30.....40.....50.....60
<i>Arabidopsis-Gc16A</i>	--VFAFSSVHRS[RNNRNQLO[VVSMAPEEEKLTRRN-----LDFPIEWERPKPGRRPDIFPK
<i>Arabidopsis-Gc16B</i>	--FIAVQR[FAL---SRKTRKOL[VVSMAPEEEKLTRRN-----LDFPIEWERPKPGRRPDIFPK
<i>C.sativa</i>	--VYGFSSVH-----SSRNNRNQLO[VVSMAPEEEKLTRRN-----LDFPIEWERPKPGRRPDIFPK
<i>O.sativa</i>	--SVSLARA[VPSRRRRRNDAI[VVCAPDEEKITRSP-----LDFPIEWERPKPGRRPDIFPK
<i>Z.mays</i>	--VTALAR[VPLVLAGR[RNNSTI[VVCAPDEEKITRSP-----LDFPIEWERPKPGRRPDIFPK
<i>P.trichocarpa</i>	--WFLAGRILVCSGS[RKQKQKMQ[VVSMAPEEEKLTRRN-----LDFPIEWERPKPGRRPDIFPK
<i>V.vinifera</i>	PLFIPQRKQMFMHKNRSKLQ[VQV[VSPDTEE-----QRSP-----LD[AP[OEWEA[MP[S]R[RPDIFPE
<i>C.reinhardtii</i>	--FCRPLALPVARRRPAFVVC[VSPDTEE-----QRSP-----LD[AP[OEWEA[MP[S]R[RPDIFPE
<i>O.lucimarinus</i>	--YPVVS[RF-----SA[RDARTGAGAFRNP-----AP[EQT[EEE[EGSVLDFP[EYQRAVE[S]R[RPDIFPD
70.....80.....90.....100.....110.....120
<i>Arabidopsis-Gc16A</i>	F-SPMKTP[LP[PMPYDPPGEDE-EEEEKK-----EEEEE[NPDQEE-----E-----EQ
<i>Arabidopsis-Gc16B</i>	F-SPMK[SPL[PMPYDPPGEDE-EEEEKK-----EEETEDDP[EKE[D-----E-----DQ
<i>C.sativa</i>	F-SPMKTP[LP[PMPYDPPGEDE-EEEEKK-----EEEEE[NPDQEE-----E-----EQ
<i>O.sativa</i>	F-SPMKTP[LP[HPLPADDPLDD-----EEEEE[EEEPQ-----PQEFPQEDDEDKEE[PEE-----DD
<i>Z.mays</i>	F-SPMKTP[LP[HPLPADDPLDD-----EEEEE[EEEQP-----PQEFPQEDDEDKEE[PEE-----DD
<i>P.trichocarpa</i>	F-SPMKTP[LP[PPLPYDPPGEDE-EEEEKK-----KKEEEETDPE[KEE-----BP-----DK
<i>V.vinifera</i>	F-SPMKTP[LP[PPLPYDPPGEDE-EEEEKK-----EEEEE[EEDEKEE-----S-----DN
<i>C.reinhardtii</i>	FEK[PQRVFLPKLPKGDP[PEMDE-----ELEESAKRTTPGDPEQKGP-----PFFEE[EA[PA[EGAPE
<i>O.lucimarinus</i>	L-KEPKTPMP[RPMP[GDP[PEMDE-----EEEEE[KERKE-----NP[G[PG[EKPE[DEE-----EN
130.....140.....150.....160.....170.....180
<i>Arabidopsis-Gc16A</i>	PEKQQ-----
<i>Arabidopsis-Gc16B</i>	PDNR-----
<i>C.sativa</i>	PD[KQQ-----
<i>O.sativa</i>	PD[KPT-----
<i>Z.mays</i>	PD[KPT-----
<i>P.trichocarpa</i>	PEKQ-----
<i>V.vinifera</i>	PEKQ-----
<i>C.reinhardtii</i>	PERKTD- IPTVPE-----
<i>O.lucimarinus</i>	EKKKKKKIKKKDDEEEE-----
190.....

B

Gc16A	MA----S[LC[SARFLPSPSPLSDI[L-----STKTKIS-----SDSRS[LGCS[SSR[VAFSSV
<i>C.sativa</i>	MAS----TLC[S-VRFLPSPSPLDV[G-KTKIS-----SDSRS[LVSSSSR[VYGFSSV
<i>J.curcas</i>	MASICL[SCSTS[SKPLV[PPKSAAGVASSSSSSPS[CHGCRC[SLVFSTPIFGSGRMISSTS
<i>B.vulgaris</i>	MA-----MAT[SC[STLSCV[SPS[CSWSNWRK[HSSLSLV[SSSPAS[AL
<i>M.domestica</i>	MA-----MALVSNT[H-TSP-----SFSFSSPKPY[SSSNYAPL
<i>G.max</i>	1.....10.....20.....30.....40.....50.....60
Gc16A	HRSS[SRRRNQLO[VVSMAPEEEKLTRRN-----LDFPIEWERPKPGRRPDIFPK-----FSPMKTP[LP
<i>C.sativa</i>	HR-[SRRRNQLO[VVSMAPEEEKLTRRN-----LDFPIEWERPKPGRRPDIFPK-----FSPMKTP[LP
<i>J.curcas</i>	S-[NTSRSSRQK[QVVCMA[DEEKLT[RNN-----LDFPIEWERPKPGRRPDIFPK-----FSPMKTP[LP
<i>B.vulgaris</i>	S-[TSNARARRS[QIV[VCMA[DEEKLT[RNS-----PLDFPIEWERPKPGRRPDIFPK-----FSPMKTP[LP
<i>M.domestica</i>	S-[FESRQHRRS[QVVCMA[DEEKLT[RNN-----LDFPIEWERPKPGRRPDIFPK-----FSPMKTP[LP
<i>G.max</i>	T-[LLPR-[FRRHLV[RMAP[DEEKLT[RNS-----PLDFPIEWERPKPGRRPDIFPK-----FSPMKTP[LP
70.....80.....90.....100.....110.....120
Gc16A	PMPYDPPPEDEEEEE-----KKEEEEN[PQEEEEE[Q[PEKQ-----Q
<i>C.sativa</i>	PMPYDPPPEDEEEEE-----KKEEEEN[PQEEEEE[Q[PEKQ-----Q
<i>J.curcas</i>	P[PYDPPPEDEEEEE-----KKEEEEN[PQEEEEE[Q[PEKQ-----E[KQ
<i>B.vulgaris</i>	PMPADPPPEDEEEEV-----KKEEEEN[PQEEEEE[Q[PEKQ-----E[KQ
<i>M.domestica</i>	PMPYDPPPEDEDEEE-----KKEEEEEE[Q[PEKQ-----X[SP[EKMTQIEQY
<i>G.max</i>	P[PEADPPPEDEDEEE-----KKEEEQE-[PQEEEEE[Q[PEKQ-----E[KQ
130.....140.....150.....160.....

Figure 22. Protein alignment of Gc16A and Gc16B.

A. Protein sequences of Gc16 proteins from *Arabidopsis* were compared to homologs from other plants and algal species (MAFFT software). Similarities/identities in at least 50% of the

aligned sequences were shaded in grey and in black colors, respectively. B. Protein sequence alignment of Gc16A with plant proteins predicted to be splicing factors and which contain an arginine/serine rich 19-like protein domain.

Gc16 proteins share also high similarity with the protein from *Camelina sativa*, which is predicted to be a splicing factor and harbors an arginine/serine rich 19-like protein domain. There are also other predictions for proteins with high similarity with Gc16, like transcriptional regulator myc-1-like, histone chaperone ASF1, tau-tubulin kinase 1-like and tRNA2-thiocytidine biosynthesis TtcA. Although, a further alignment with other Gc16 proteins from plant species containing a predicted argine/serine rich 19-like protein domain revealed that indeed the conservation is high, with the highest conservation (91% identity) found in an alignment of Gc16A with Gc16 of *Camelina sativa* (Figure 22.B). The domain prediction is defined on a consensus sequence found in metazoan proteins (arginine/serine rich-SR), which are essential for splicing or other functions related to RNA metabolism (Haynes & Iakoucheva 2006).

3.3.2. Mutant analysis of *gc16a* and *gc16b*

To identify the function of Gc16A/B in plants, one T-DNA insertion mutant for Gc16A (*gc16a*- SALK_133989) and two different insertion mutants for Gc16B (*gc16b-1*- GABI 295A02 and *gc16b-2*- SAIL_71_A01) were obtained from publicly available collections. Both genes consist of two exons and one intron (Figure 23.A). The T-DNA in *gc16a* is inserted in the first intron, while for *gc16b-1* the insertion is located in the 3'UTR and in the intron for *gc16b-2*. The growth rate of the mutants was monitored for 27 days showing that the single mutants (*gc16a*, *gc16b-1*and *gc16b-2*) were not significantly affected. For that reason, two different double mutants were created, with one of them to be already available from the DUPLO collection (Bolle et al. 2013) (*gc16a/gc16b-1*), while the second was created in this thesis (*gc16a/gc16b-2*). The double mutants showed slower growth rates compared to the wild type under long day conditions. Particularly, the mutant *gc16a/b-2* was more severely affected than *gc16a/b-1* with respect to growth rates (Figure 23.B, C).

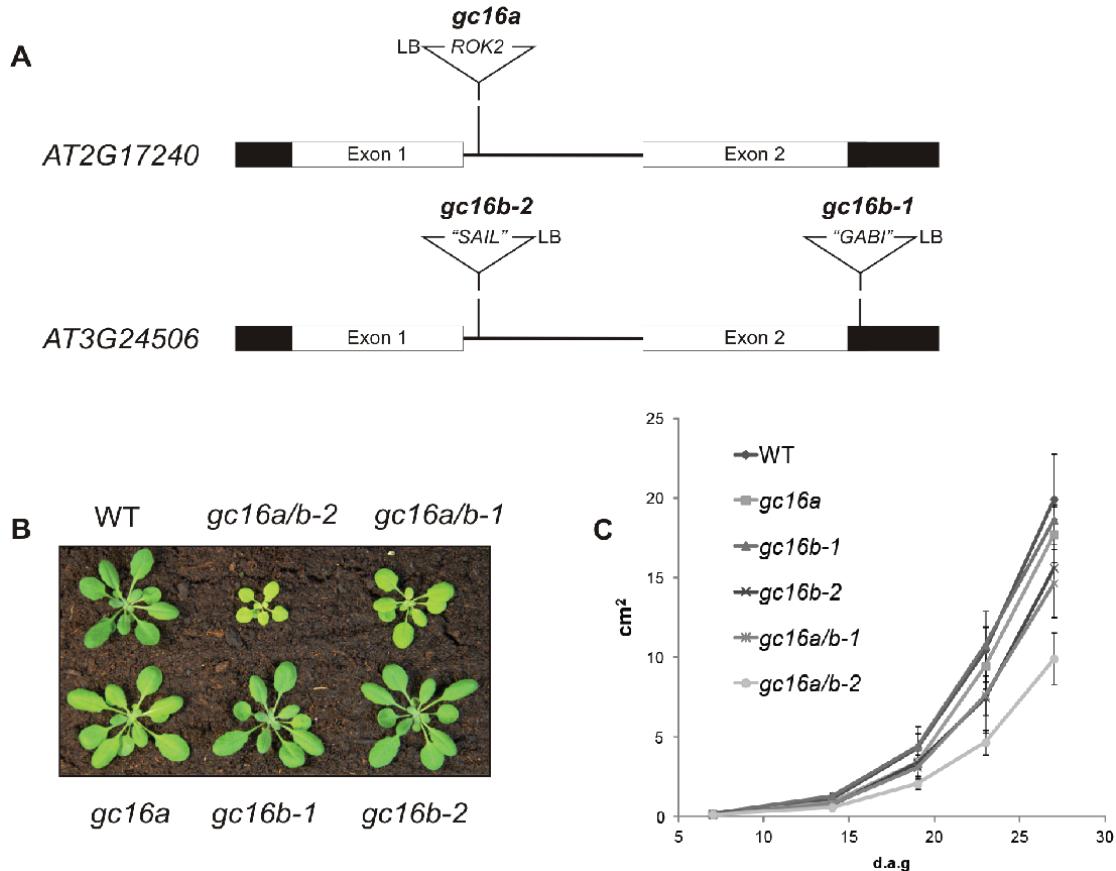


Figure 23. T-DNA insertion lines and growth phenotype of *gc16a* and *gc16b* mutants.

A. Insertion map of the T-DNA in Gc16A and Gc16B. 5' and 3' UTRs are shown in black boxes, exons in white boxes and introns in lines. B. Growth phenotype of four-week-old wild type, *gc16a/b-1*, *gc16a/b-2* and the respective single mutants, grown under long day conditions (16h light/8h dark) in the climate chamber. C. Growth rate analysis of wild type and *gc16* mutants calculated by leaf area (ImageJ software). Values are the average of 12 plants and standard deviations are indicated by error bars.

Real time PCR was performed to check *Gc16* gene expression in all mutant lines. *Gc16A* transcripts were completely absent from *gc16a* and both double mutants, whereas *GC16B* was still expressed in *gc16b-1* but not in *gc16b-2*. *Gc16B* transcripts were still detectable in *gc16a/b-1*, but not in *gc16a/b-2*, which is in line with the more severe growth phenotype observed for *gc16a/b-2* (Figure 24). Taking those results together, *gc16a/b-1* is a knock-down mutant line still expressing *Gc16B*, while *gc16a/b-2* is a knock-out mutant line for both *GC16* genes.

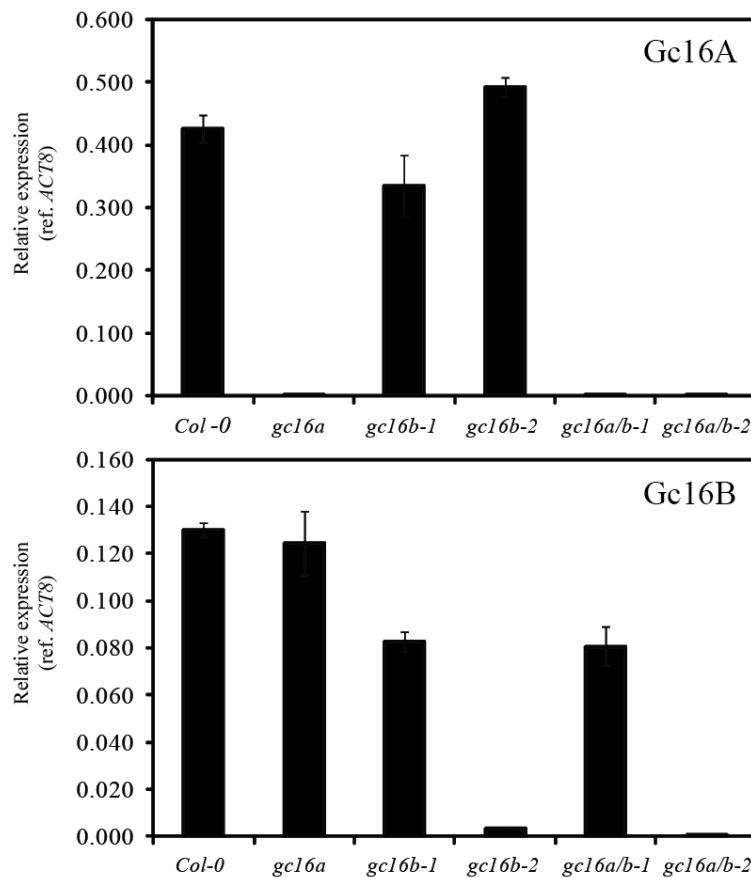


Figure 24. Transcription analysis for *gc16* mutant lines.

Gc16A and *Gc16B* transcript levels were detected by real time PCR analysis. Primers for the amplification have been designed by the web tool from Roche (Universal ProbeLibrary Assay Design Center).

3.3.3. Complementation of *Gc16* function

To determine whether the loss of *Gc16A/B* is responsible for the mutant phenotype, complementation analysis was performed by transforming the mutant line *gcl6a/b-2* with two overexpression constructs. To this end, *Gc16A* or *Gc16B* were placed under the control of the 35S promoter and fused to the eGFP-coding sequence (35S:*Gc16A:GFP* or 35S:*Gc16B:GFP*). After selection of BASTA-resistant plants, the integration of the constructs in the genome was tested via PCR (for primer information, see table 5). The successful complementation of 35S:*Gc16A:GFP* or 35S:*Gc16B:GFP* lines was confirmed for several independent lines, which rescued the wild type phenotype (Figure 25). Immunoblot blot analysis with a *Gc16*-specific antibody raised against a synthesized peptide sequence (peptide sequence; refer to Figure 21) was also performed to test fusion protein levels in the complemented lines. *Gc16A/B* is

only detectable in the overexpressor lines (~40 kDa) (data not shown), but undetectable in wild type plants (expected molecular mass 10 kDa), which is most probably due to low, native protein levels. Taking those results together we can assume that the disruption of Gc16 proteins causes the growth retardation and the visible pigment phenotype of the knockout lines.

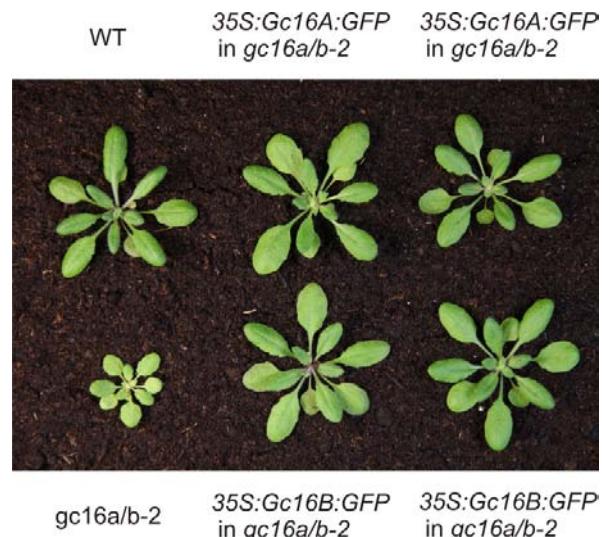


Figure 25. Complementation analyses of *gc16a/b-2* line with 35S:Gc16A:GFP or 35S:Gc16B:GFP construct.

Growth rate of four week-old wild type, mutant (*gc16a/b-2*) and the complemented lines (35S:Gc16A:GFP and 35S:Gc16B:GFP) growing under long day conditions in climate chamber. Several independent lines rescued the phenotype of wild type plants.

3.3.4. Subcellular and suborganellar localization of Gc16A and Gc16B

According to several prediction algorithms (ChloroP, TargetP, PredSL, PProowler) the subcellular localization of Gc16A and Gc16B is the chloroplast. To examine protein localization, protoplasts from stably transformed plants with 35S:Gc16A:GFP and 35S:Gc16B:GFP were isolated from 3-week-old plants grown under long day conditions. Gc16A/B:GFP fusion proteins were localized only to the chloroplast, since eGFP signals coincided with chlorophyll autofluorescence signals from the chloroplast (Figure 26).

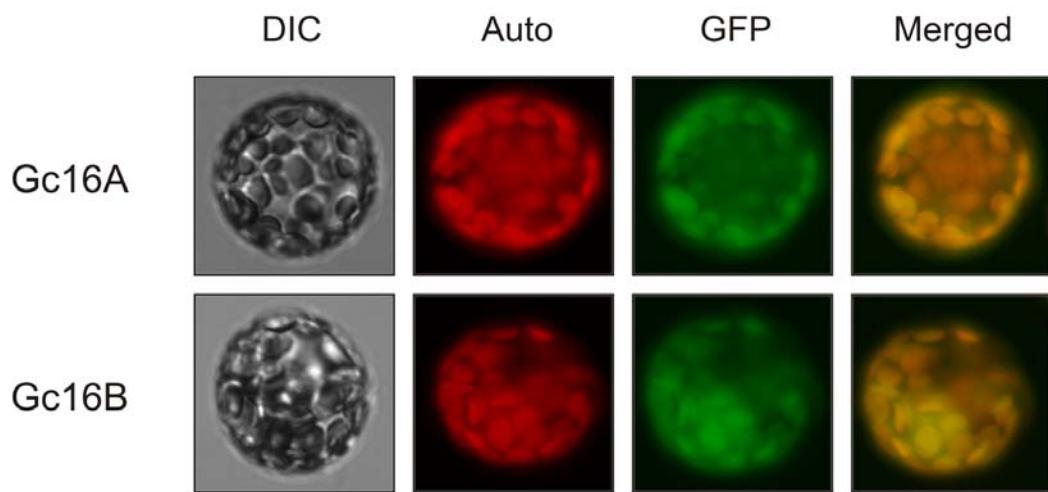


Figure 26. Subcellular localization of Gc16A and Gc16B.

Fluorescence microscopy of protoplasts reveal the localization of Gc16 in the chloroplasts. DIC: image of a protoplast in bright field, Auto: autofluorescence signal, GFP: GFP signal, Merged: Merged image of autofluorescence of protoplasts with GFP signal.

The chloroplast ultrastructure was examined by transmission electron microscopy (TEM). The chloroplasts of all the mutants were smaller than the wild type, gradually reduced from the single to double mutant. The thylakoids in the mutants were not affected significantly and any differences could be due to the different developmental stage between wild type and the mutants. In total, Gc16 proteins do not affect chloroplast architecture (Figure 27).

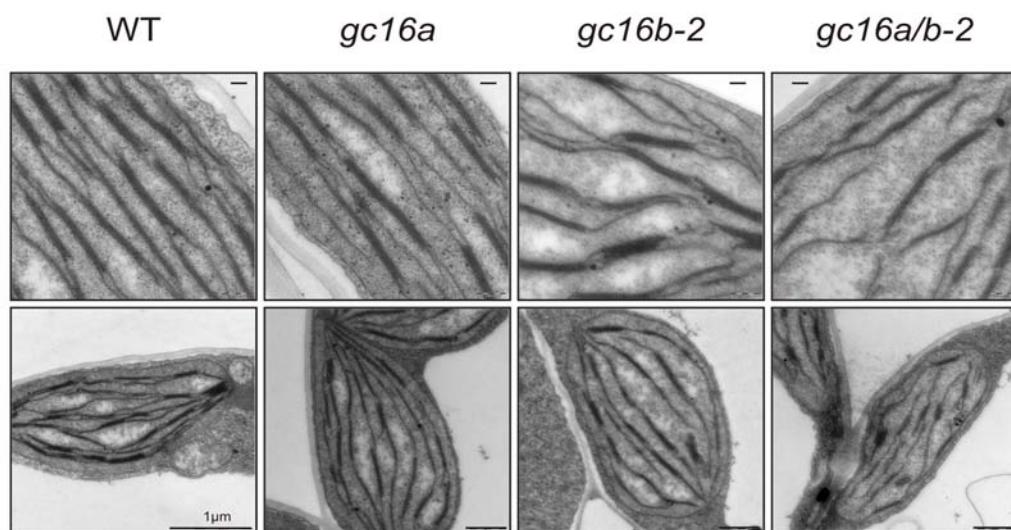


Figure 27. Chloroplast ultrastructure of wild type and gc16a/b-2 mutants.

Transmission electron microscopy (TEM) pictures of 20 days old leaves of gc16 mutants and the wild type.

3.3.5. Pigment composition in *gc16* mutants

Pigments were extracted from the mutant lines grown under long day conditions (4 week old plants), at 80-100 $\mu\text{mol m}^{-2}\text{s}^{-1}$ light intensity and subjected to reverse-HPLC, according to Färber et al. (1997). Total chlorophyll concentration was affected only in the *gc16a* single mutant (Chla+b: 789 \pm 215 pmol/mg) but not significantly changed in the *gc16b* in comparison to the wild type plants (Chla+b: 911 \pm 215 pmol/mg). Furthermore a gradual decrease in chlorophyll content was observed for the double mutants (*gc16a/b-1* Chla+b: 604 \pm 11 pmol/mg and for *gc16a/b-2* Chla+b: 410 \pm 30 pmol/mg).

Table 7. Leaf pigment content of *gc16a*, *gc16b-1*, *gc16b-2*, *gc16a/b-1* and *gc16a/b-2* mutants. The amount of pigments is shown in pmol/mg fresh weight obtained from four physiological replicates. Nx: Neoxanthis, Vx: Violaxanthin, Ax: Antheraxanthin, Zx: Zeaxanthin, VAZ: Violaxanthin+Antheraxanthin+Zeaxanthin, Chl a: Chlorophyll a, Chlb: Chlorophyll b, Chla+b: total Chlorophyll, Chla/b: Chlorophyll a to b ratio.

Pigment	pmol / mg fresh weight					
	Col-0	<i>gc16a</i>	<i>gc16b-1</i>	<i>gc16b-2</i>	<i>gc16a/b-1</i>	<i>gc16a/b-2</i>
Nx	28 \pm 6	23 \pm 4	27 \pm 2	29 \pm 5	17 \pm 3	11 \pm 1
Vx	25 \pm 5	24 \pm 5	25 \pm 2	26 \pm 4	27 \pm 5	21 \pm 1
Ax	1 \pm 0	1 \pm 0	1 \pm 0	1 \pm 0	1 \pm 0	1 \pm 0
Zx	0	0	0	0	0	0
VAZ	26 \pm 6	25 \pm 6	25 \pm 2	27 \pm 5	28 \pm 5	22 \pm 2
Lutein	91 \pm 21	83 \pm 18	89 \pm 5	96 \pm 19	73 \pm 14	50 \pm 4
β -car	71 \pm 15	63 \pm 9	66 \pm 4	74 \pm 11	47 \pm 9	31 \pm 2
Chl a	693 \pm 162	604 \pm 11	653 \pm 42	706 \pm 124	467 \pm 85	318 \pm 24
Chl b	218 \pm 53	185 \pm 34	211 \pm 17	226 \pm 38	136 \pm 26	92 \pm 7
Chl a+b	911 \pm 215	789 \pm 146	865 \pm 58	932 \pm 162	604 \pm 11	410 \pm 30
Chl a/b ratio	3,19 \pm 0,07	3,26 \pm 0,03	3,09 \pm 0,06	3,13 \pm 0,02	3,43 \pm 0,02	3,44 \pm 0,04

The concentration of xanthophylls, violaxanthin and antheraxanthin, was not affected in any mutant line, but neoxanthin was reduced in the double mutant (*gc16a/b-2*) with respect to wild type (28 \pm 6 to 11 \pm 1 pmol/mg). In contrast, lutein and β -carotene levels declined gradually with lowered *Gc16AB*

expression levels with the more significant effect on the knock out mutant *gc16a/b-2* (Lut: 50 ± 4 pmol/mg, β -car: 31 ± 2 pmol/mg) compared to the wild type (Lut: 91 ± 21 pmol/mg, β -car: 71 ± 15 pmol/mg) (Table 7). Interestingly, the Chlorophyll a/b ratio was increased for the double mutants, from 3.19 ± 0.07 for wild type to 3.43 ± 0.02 for *gc16a/b-1* and 3.44 ± 0.04 for *gc16a/b-2*.

3.3.6. Gc16A and B are required in NDH and *Cytb6f* biogenesis

To study the influence of Gc16 disruption on thylakoid membrane complex composition, BN/SDS-PAGE analysis was performed. Thylakoids were solubilized and complexes were fractionated from single and double knockout mutants in Blue Native PAGEs (Figure 28.A). According to BN analysis, a reduction in all photosynthetic complexes could be observed in the double mutant. The most prominent difference was the significant reduction of PSI-NDH from the double mutant, which indicates a defect either in PSI or in the NDH complex (Figure 28.A).

A strong reduction of PSI-NDH was identified in *gc16a/b-2* in the second dimension of BN-PAGE also. Furthermore, three different spots were absent in the double mutant, between 20-35 kDa area (Figure 28.C,D). The same observations were also shown for the knock down mutant (*gc16a/b-1*) with its respective single ones (data not shown). Proteins migrating in this area correspond to proteins from the complex of *Cytb6f/PSII* monomer and for the verification of this, mass spectrometry was performed. Indeed, it was confirmed that these three spots correspond to *Cytb6f* proteins. Taking together from the BN analysis there is a defect in PSI-NDH and cytochrome *b6f* complexes.

Immunoblot analysis was performed in the first and second dimension with *NdhH*- and *Cytb6*-specific antibodies. *NdhH* is plastid-encoded and part of the NDH subcomplex A, while the *petB* gene encodes for *Cytb6* protein. *NdhH* completely lacks in *gc16a/b-2* samples in the first dimension detection, while there is still some residual protein detected in the second dimension gel. *Cytb6* could still be detected in both assays, however was severely reduced in the double mutant while the accumulation was not severely affected in the single mutants (Figure 28.B,C).

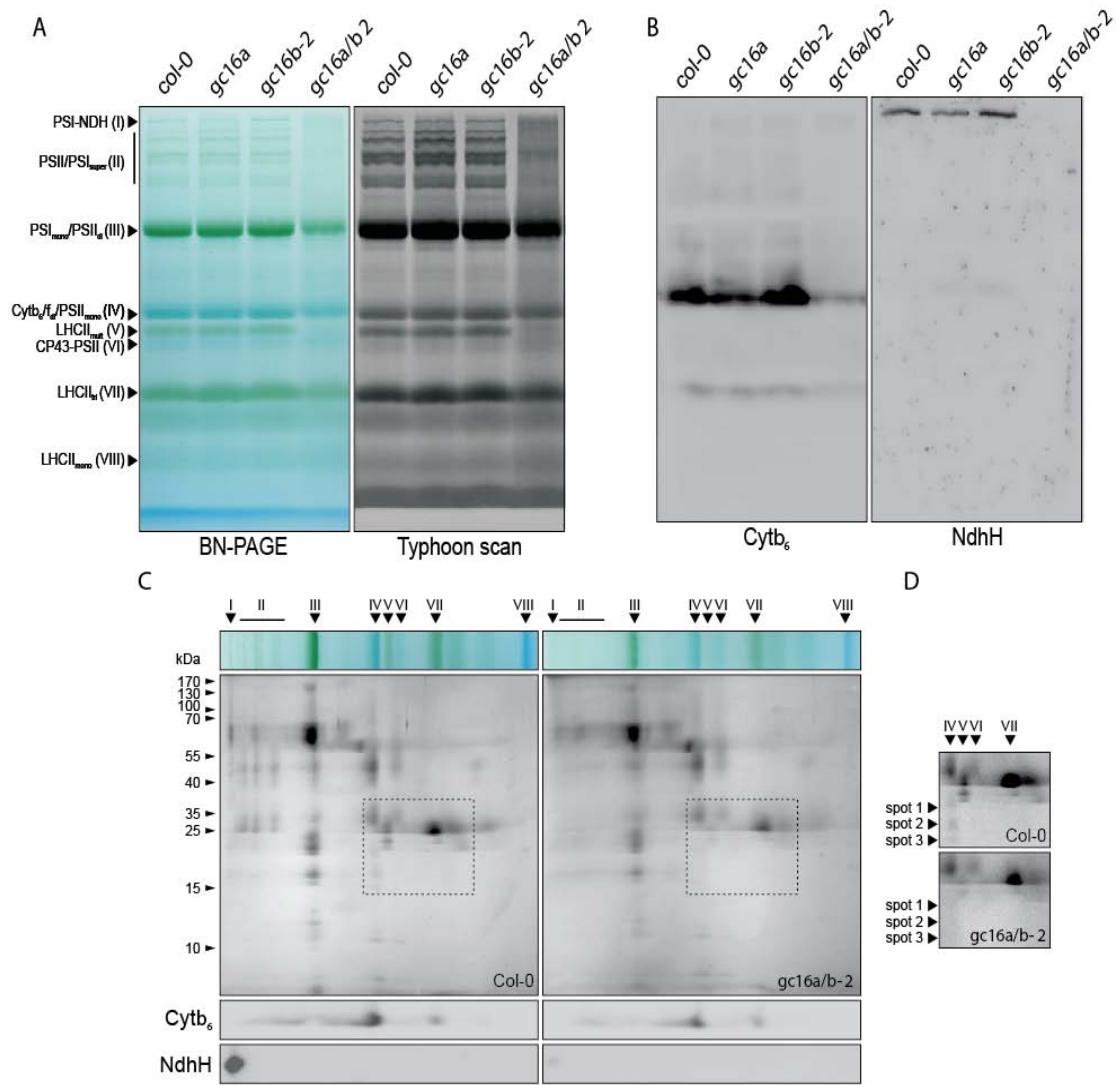


Figure 28. Blue Native Analyses of wild type, single mutants (*gc16a*, *gc16b-2*) and the double mutant (*gc16a/b-2*).

A. Isolated thylakoid membranes (corresponding to 20 µg Chl) were solubilized with β-DM (1%) and fractionated on BN gels (5-12% gradient). A Typhoon scan was performed to visualize low-abundant complexes. B. Immunodetection of Cytb_f and NdhH in the first dimension. C. Coomassie staining of size-fractionated complexes in the second dimension. Proteins were blotted on PVDF membranes and Cytb_f and NdhH were immunodetected. D. Magnified area of the 2D gel staining shown in C. Protein spots, which are less abundant in the mutant (in cooperation with Thilo Rühle).

3.3.7. Thylakoid composition and translation rates in *gc16* mutants

The abundance of thylakoid complexes was quantified by immunoblot analysis. Since the double mutant is characterized by a lowered Chl

concentration (Table 7), RbcL and Comassie (C.B.B) staining of the membranes were used as loading controls. *gc16a/b-2* showed a reduction of RbcL amounts to 50% wild type levels, which we interpreted as a general alteration in chloroplast protein content in *gc16a/b-2*.

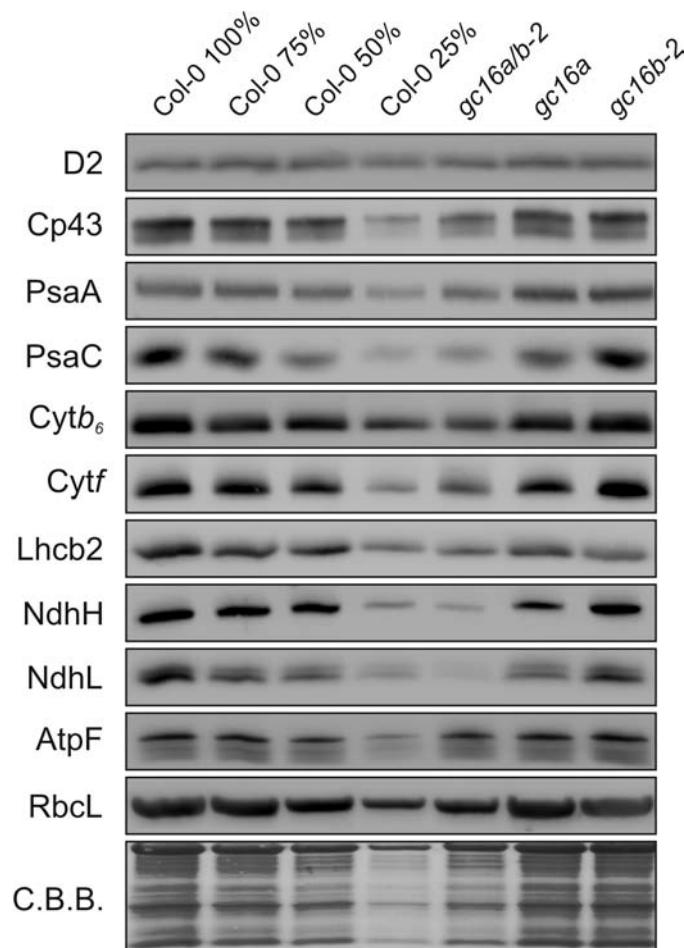


Figure 29. Abundance of photosynthetic proteins in *gc16a/b-2*, *gc16a*, *gc16b-2* mutants and wild type.

Serial dilutions of wild type and *gc16* samples were size-fractionated on SDS-PAGEs and blotted onto PVDF membranes. Specific antibodies were used for marker subunit detection of different thylakoid complexes (PSII:D2, CP43, PSI: PsaA, PsaC, Cytb₆/f: cyt_b₆, cyt_f, LHCs: Lhcb2, NDH: NdhH, NdhL, NdhT, ATP synthase: AtpF and RbcL). (in cooperation with Thilo Röhle and Thomas Huber).

All marker subunits tested were reduced to at least 50% wild type levels (PSII, PSI and cpATPase subunits), whereas Lhcb2 (30%), Cytb₆ (30-40%), NdhH (10%) and NdhL (10%) in *gc16a/b-2* were more severely affected than RbcL

contents. Interestingly, the reduction of Cytb_{6f}- and NDH-complex levels could already be observed in the single mutant *gc16a* (Figure 29).

To check if the reduced levels of NDH- and Cytb_{6f}-complexes in *gc16* mutants are caused by a defect in protein synthesis *in vivo* labeling experiments were performed with *gc16a/b-1* (knock down mutant), the respective single mutants and the wild type. Leaf disks were incubated with (³⁵S) Methionine under light and cytosolic translation was blocked with cycloheximide. Newly synthesized proteins were extracted after 10, 20 and 40 min. The synthesis rate was monitored according to the synthesis of RbcL and D1 proteins. The mutants of *gc16* showed a reduced RbcL and D1 synthesis rate, with a more pronounced effect on *gc16a* and *gc16a/b-1*. The amount of RbcL labeled for this experiment was also approximately 50% of the wild type levels.

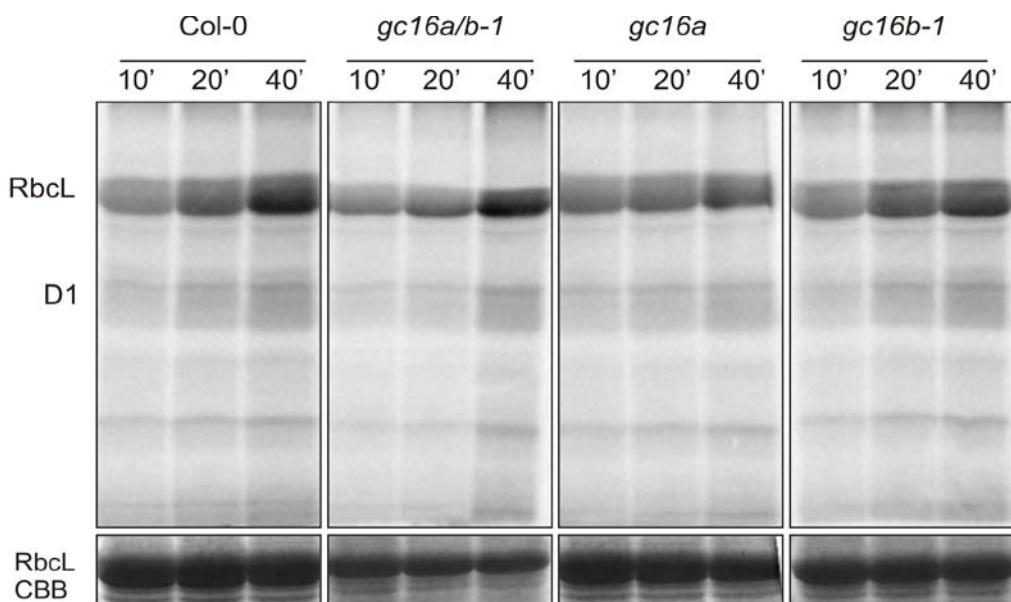


Figure 30. Synthesis rate for chloroplastic proteins.

In vivo labeling of total proteins, after blocking cytosolic translation with cycloheximide. RbcL and D1 proteins were monitored for the quantification of translation rate.

3.3.8. Chloroplast gene expression in *gc16* mutants

To check whether plastid transcription was affected in *gc16* mutants, Northern blot analyses were performed for plastome-encoded genes (Figure 31). When total RNA was resolved on a denaturing agarose gel, a different pattern of the abundant ribosomal RNAs could be observed in *gc16a/b-1* samples. Two

additional bands were detected between the 25S and 18S rRNAs. This was a first indication that a general defect in plastid gene expression could be the reason for the lowered protein content in *gc16a/b-1* plants. To identify the bands, the rRNA operon was analyzed with probes covering the rRNAs, tRNAs and also the intergenic region between 23S and 4.5S rRNA (Figure 31). The 5S (*rrn5*) and 16S (*rrn16*) transcript patterns were similar to the wild type patterns, but differed when probes were hybridized against 23S (*rrn23*), 4.5S (*rrn4.5*) and against their intergenic region. 23S and 4.5S are located on a dicistronic transcript, which is transiently formed during processing (Bollenbach et al. 2005). Interestingly, there was not only an effect on the mature transcripts, but unprocessed transcripts seem to be more abundant in *gc16a/b-1*. The amount of transcripts for the tRNAs isoleucine (*trnI*) and alanine (*trnA*) were not significantly affected, while tRNA^{Arg} revealed an increase of transcripts for *gc16b-1* and *gc16a/b-1*.

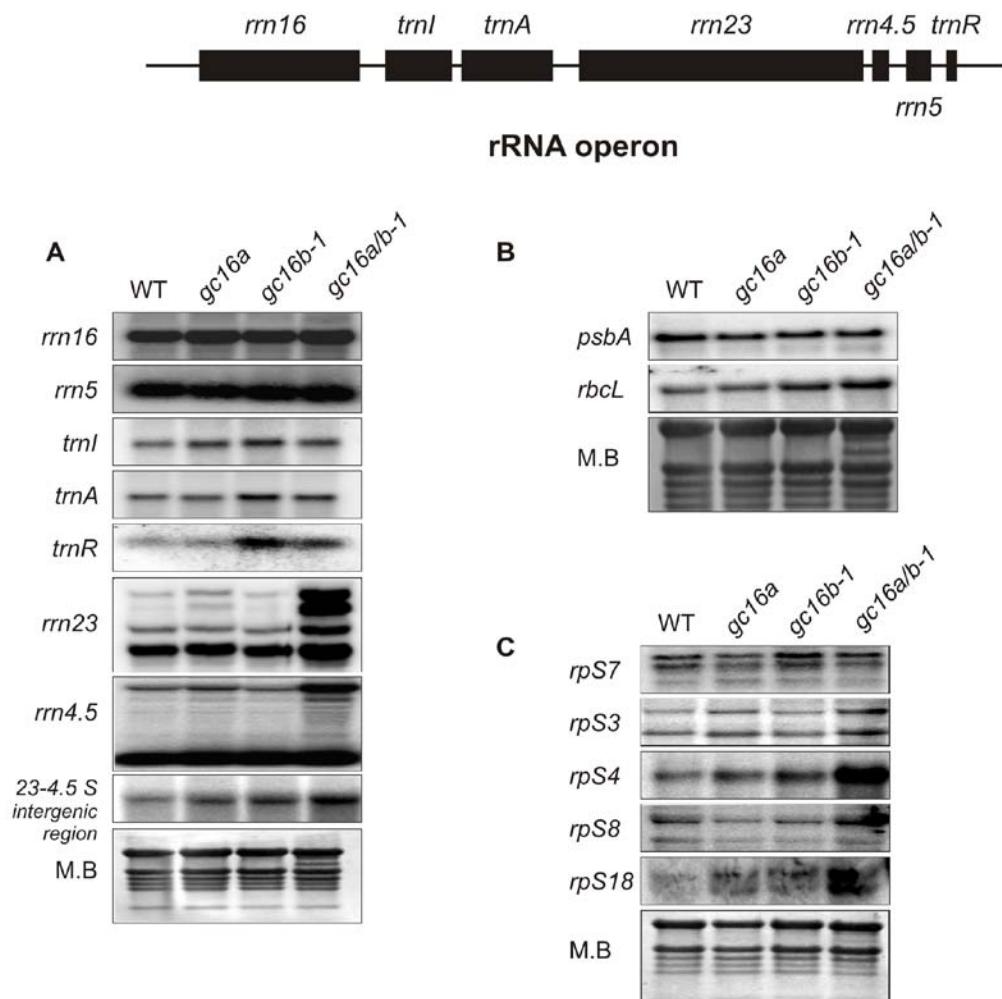


Figure 31. Transcription analyses of chloroplast genes in *gc16* mutants.

Accumulation of *rrn* (A), *psbA*, *rbcL* (B) and *rps* transcripts (C). RNAs from *gc16* and wild type plants were detected by Northern using ³²P-labeled probes. Methylene blue (M.B) staining was used as loading control.

Northern blot analysis for *psbA* transcripts revealed no difference for all the *gc16* lines compared to the wild type, but *rbcL*-transcripts accumulated more in *gc16a/b-1* (Figure 31.B). Moreover, several transcripts coding for ribosomal proteins were tested (*rpS7*, *rpS3*, *rpS4*, *rpS8* and *rpS18*) (Figure 31.C) and were increased in *gc16* lines, the most prominent effect to be for the double mutant. According to those results neither a processing defect nor downregulation of any tested transcripts could be observed, which might have provided a direct hint for Gc16 function.

3.3.9. NDH activity and photosynthetic parameters in *gc16* mutants

The lack of the PSI-NDH supercomplex formation (Figure 28) led us to the assumption that, NDH activity could be altered in the mutant lines. Hence, NDH activity measurements were performed by measuring the chlorophyll a fluorescence rise during a light to dark transition (Shikanai et al. 1998). This phenomenon observed in the wild type can be explained by a transient reduction of plastoquinone pool (PQ) mediated by the NDH complex in the dark. The increase could only be detected in wild type plants and the single mutants, but was drastically reduced in the double mutant background, *gc16a/b-1* and *gc16a/b-2* (Figure 32.A,B). This is consistent with the results from BN-PAGE, where no PSI-NDH complex was accumulated, indicating that at least one Gc16 protein is required for proper NDH biogenesis in *Arabidopsis*.

This analysis also provided a way to determine the complementation of the plants overexpressing one of the Gc16 proteins. When Gc16A or Gc16B is present in the *gc16a/b-1* double mutant background, NDH activity could be restored (Figure 32.B), confirming that the presence of Gc16 is needed for NDH biogenesis.

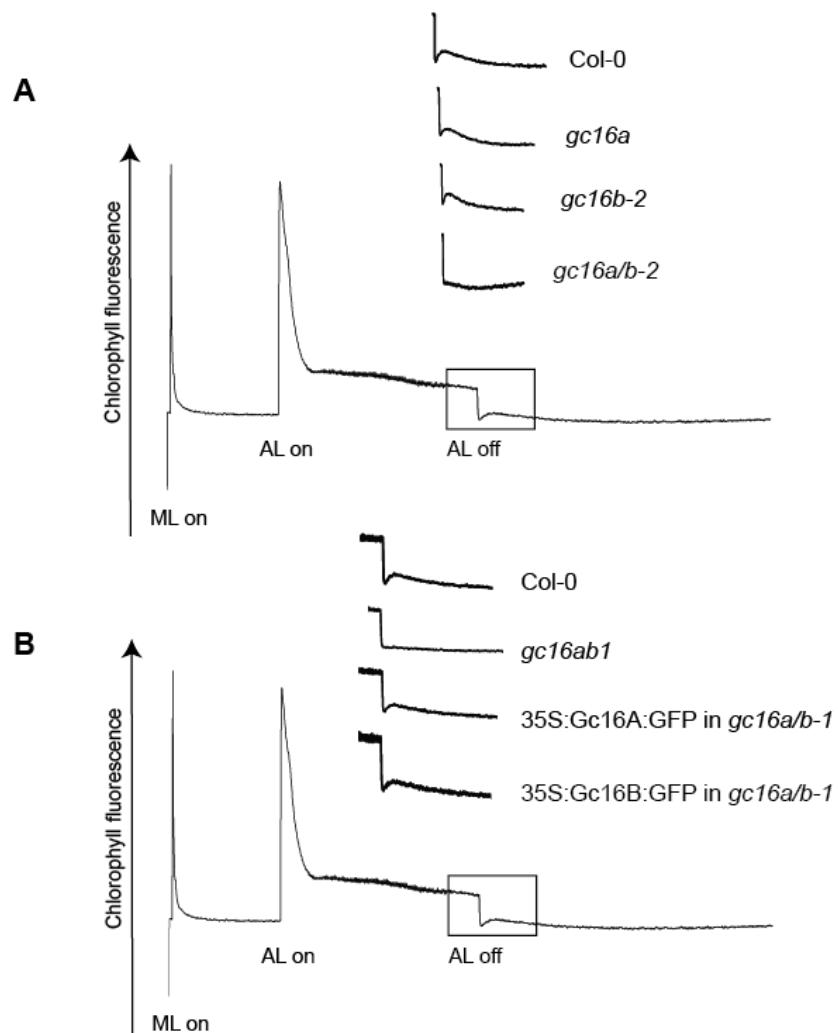


Figure 32. NDH activity in *gc16a*, *gc16b-2*, *gc16a/b-2* and *gc16a/b-1* plants and the complemented lines.

A. Four-week-old plants were used to determine the NDH activity by chlorophyll fluorescence measurements (Imaging-PAM). The leaves were dark adapted for 15 min and then exposed to a saturating light pulse to determine the maximal fluorescence yield (Fm). Actinic light (AL) was switched on for 5 min and the subsequent transient rise of fluorescence was monitored in the dark. ML: measuring light. B. Complementation of *gc16* mutants was examined by applying the same protocol to plants carrying the overexpressor construct 35S:Gc16A:GFP or 35S:Gc16B:GFP.

Photosynthetic performance was monitored for *gc16* mutants and overexpression lines using DUAL PAM. Single leaves of 4 weeks old plants grown under long day conditions were used for the measurements. Six leaves per genotype were measured. The maximum quantum efficiency (F_v/F_m) was slightly reduced in *gc16a/b-2* ($0,76 \pm 0,02$) compared to wild type ($0,81 \pm 0,00$),

indicating that PSII is affected in *gc16a/b-2*. For the PSII measurements only NPQ and qE were reduced in *gc16* mutants. Non photochemical quenching (NPQ) was significantly decreased ($0,25\pm0,02$) for the mutants from that of wild type ($0,63\pm0,05$) and restored in the overexpression lines. Photoinhibition (ql) and excitation pressure (1-qP) were not altered, but the energy dependent quenching (qE) was reduced in *gc16a/b-2* ($0,11\pm0,05$) and *gc16a* ($0,15\pm0,08$) compared to wild type ($0,49\pm0,04$). The photochemical quantum yield of PSI (Y(I)) was not affected, but the fraction of P700 that is oxidized is increased in *gc16a/b-2* ($0,53\pm0,09$) compared to wild type ($0,35\pm0,02$). This reflects a limitation of the electron transfer system at cytochrome *b₆f* site for the *gc16* mutants.

Table 8. Photosynthetic parameters of wild type, *gc16a*, *gc16b-2*, *gc16a/b-2* and overexpression lines.

Four weeks old plants grown under long day conditions were used for the measurements. Averages \pm standard deviations were calculated from 6 physiological replicates. F_v/F_m: maximal quantum yield of PSII; Y(II): photosynthetic yield; NPQ: non photochemical Chlorophyll fluorescence quenching; 1-qP: excitation pressure; qE: energy-dependent quenching of Chl fluorescence; ql: photoinhibitory quenching; Y(I): photochemical quantum yield of PSI; Y(ND) and Y(NA): non photochemical quantum yield of PSI.

	F _v /F _m	Y(II)	NPQ	1-qP	qE	ql	Y(I)	Y(ND)	Y(NA)
WT	0.81 \pm 0.00	0.57 \pm 0.01	0.63 \pm 0.05	0.20 \pm 0.02	0.49 \pm 0.04	0.23 \pm 0.05	0.65 \pm 0.02	0.35 \pm 0.02	0.05 \pm 0.03
<i>gc16a/b-2</i>	0.76 \pm 0.02	0.57 \pm 0.02	0.25 \pm 0.05	0.18 \pm 0.02	0.11 \pm 0.05	0.19 \pm 0.01	0.56 \pm 0.13	0.53 \pm 0.09	0.17 \pm 0.20
<i>gc16a</i>	0.81 \pm 0.00	0.65 \pm 0.02	0.28 \pm 0.07	0.16 \pm 0.01	0.15 \pm 0.08	0.21 \pm 0.02	0.64 \pm 0.04	0.26 \pm 0.02	0.09 \pm 0.03
<i>gc16b-2</i>	0.82 \pm 0.01	0.56 \pm 0.03	0.61 \pm 0.06	0.23 \pm 0.02	0.50 \pm 0.07	0.31 \pm 0.02	0.62 \pm 0.09	0.44 \pm 0.12	0.04 \pm 0.02
35S:GC16A:GFP.1	0.81 \pm 0.01	0.56 \pm 0.03	0.55 \pm 0.14	0.23 \pm 0.03	0.44 \pm 0.12	0.27 \pm 0.07	0.44 \pm 0.09	0.49 \pm 0.10	0.06 \pm 0.04
35S:GC16A:GFP.2	0.82 \pm 0.00	0.63 \pm 0.03	0.50 \pm 0.12	0.16 \pm 0.03	0.38 \pm 0.13	0.23 \pm 0.02	0.68 \pm 0.03	0.26 \pm 0.04	0.08 \pm 0.02
35S:GC16B:GFP.1	0.81 \pm 0.01	0.66 \pm 0.01	0.27 \pm 0.03	0.15 \pm 0.01	0.13 \pm 0.01	0.22 \pm 0.03	0.71 \pm 0.05	0.20 \pm 0.07	0.07 \pm 0.03
35S:GC16B:GFP.2	0.82 \pm 0.00	0.54 \pm 0.02	0.75 \pm 0.11	0.23 \pm 0.01	0.74 \pm 0.11	0.33 \pm 0.02	0.53 \pm 0.06	0.43 \pm 0.05	0.06 \pm 0.02

3.3.10. Gc16 is required for proper transcription in *ndhH-D* operon in the chloroplasts

11 subunits of the NDH complex are encoded by two operons and two single genes on the plastome. Hence, transcript accumulation of the chloroplast genes coding for NDH complex subunits were examined by Northern blot analysis.

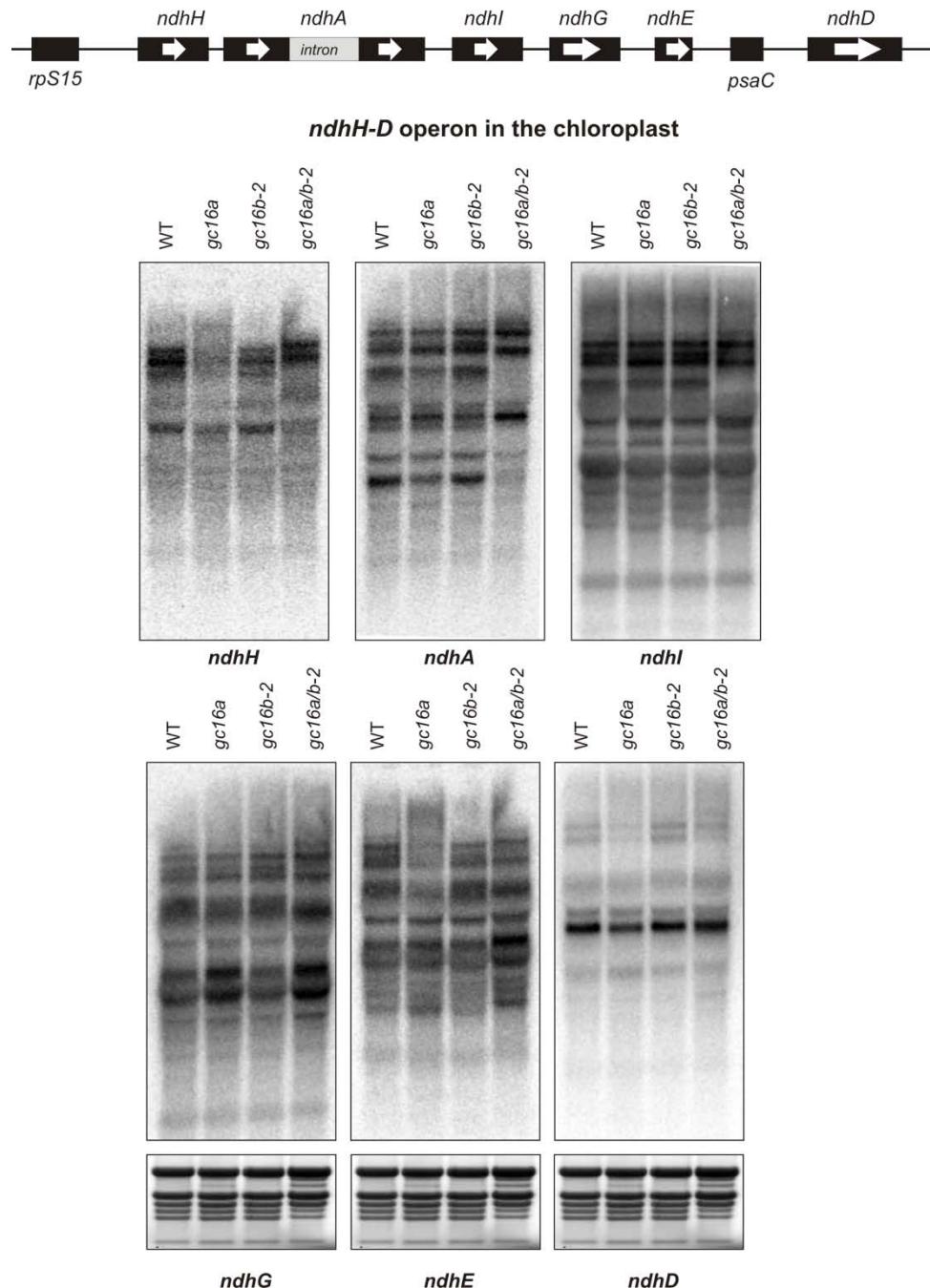


Figure 33. Transcript accumulation of chloroplast *ndh* genes, arranged in the *ndhH-D* operon, in the *gc16a/b-2* and the respective single mutants.

Total RNA was isolated from 3-week-old plants and treated with gene-specific probes for each subunit of NDH complex in this operon. Methylene blue (M.B) staining was used as loading control in each case.

Interestingly, *ndhA* is the only gene, which carries an intron in operon *ndhH-D* and requires splicing to obtain mature transcripts. Several genes of this operon were affected; *ndhG*, *ndhE* and *ndhD* containing transcripts were upregulated in the double mutant, while *ndhI* transcripts showed the same pattern compared to the wild type. However, one band was completely absent in the double mutant sample. *NdhH* transcript patterns remained the same. Even though one of three detected transcripts was less abundant, the other two did not accumulate to a higher degree, which indicated correct processing in those cases (Figure 33). Remarkably, similar effects but less pronounced were observable in the single mutant *gc16a*.

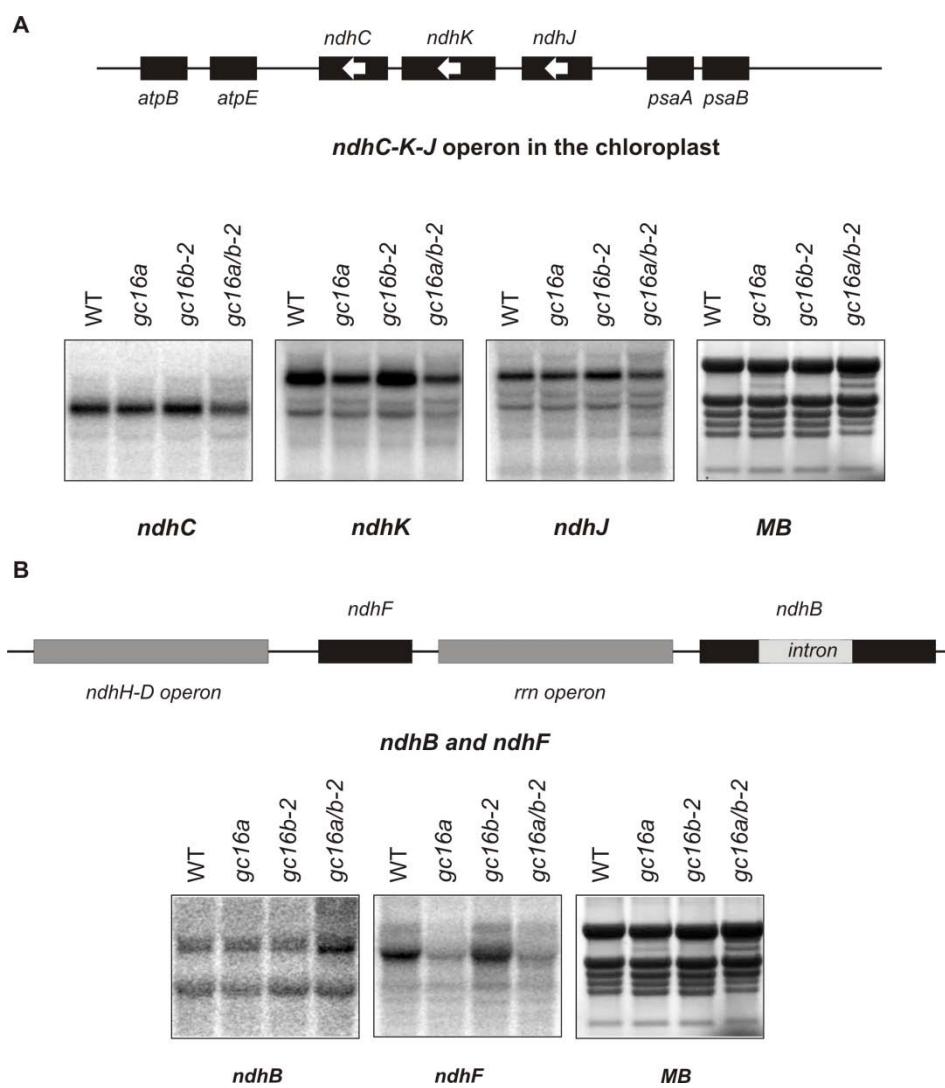


Figure 34. Transcript analysis for *ndhC-K-J* operon (A) and *ndhB* and *ndhF* genes (B).

Northern blot analysis was performed using probes against the plastid subunits C, K, J, B, F of the NDH complex. Methylene blue (M.B) staining was used as loading control for every membrane.

The next step was to clarify whether expression of the other NDH genes in the chloroplast were affected as well. According to RNA blot analysis carried out with the *ndhC-K-J* operon, transcripts of all genes accumulated in slightly reduced levels only in the double mutant background (*gc16a/b-2*) (Figure 34.A.). *NdhB*, the second intron-containing gene overaccumulated, whereas *ndhF* levels were reduced in the double mutant (Figure 34.B).

Taken those results together, the most striking difference were observed in Northern analyses with *ndhA* probes, which revealed that two transcripts were almost absent from the double mutant, but transcripts which might correspond to unprocessed transcripts were more abundant (Figure 34).

3.3.11. Gc16A/B associates with group II introns in the chloroplasts

To determine whether the defect in the two photosynthetic complexes underlies to a splicing defect in *gc16* mutants, the splicing of *ndhA* was examined. RNA blot analysis for *gc16* mutants revealed that the excised intron (1kb) was decreased in the *gc16a/b-2* and the unspliced transcripts were affected, one of the unprocessed transcripts was upregulated (Figure 35). This result in combination with the detection of *ndhA* exon shows that Gc16A and Gc16B are required for correct intron splicing of *ndhA*.

According to BN-PAGE and immunoblot analysis *Cytb₆f* was less abundant in the double mutant plants. There are two genes in this complex, located in the chloroplast, carrying an intron (*petB* and *petD* gene encoding for *Cytb₆* and subunit IV, respectively).

RNA blot assays, using intron and exon specific probes for each gene, were performed. For *petB* when the exon was detected there was an overall reduction of the transcripts, except the two high molecular weight bands, corresponding to unprocessed transcripts, which were upregulated in the *gc16a/b-2*. However, when the intron was detected, there was a decrease for the transcripts corresponding to excised intron for the double mutant and an increase for the unprocessed precursor transcripts. This effect correlates with the one that was observed for *ndhA*, when the exon and intron was detected.

petD also showed a misregulation in transcript levels, a drastic reduction of the transcripts in the double mutant but an upregulation for the spliced intron (Figure 35). So, from this analysis we could propose that the reduction of Cytb₆f in *gc16* mutants could be due also to splicing defects in *petB* and possibly *petD*.

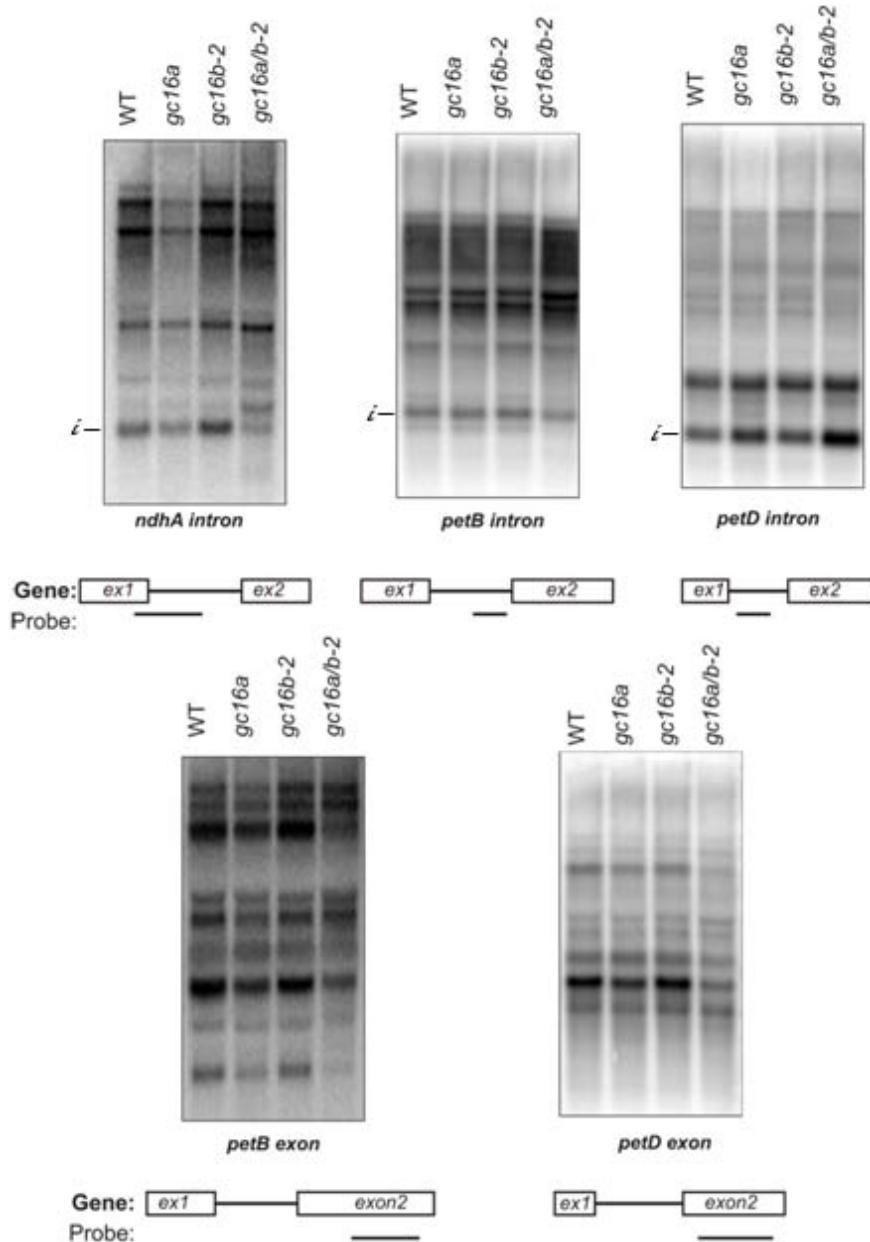


Figure 35. Expression analysis of *ndhA*, *petB* and *petD* genes.

Northern blot analysis was performed in 3 weeks old plants and gene specific probes were used for intron and exon detection. The sequences representing each probe are shown on the map below. Bands corresponding to excised introns are marked (i).

3.3.12. Gc16B is an RNA-binding protein

Since there is an indication that Gc16A/B is a splicing factor, this led to the suggestion that Gc16 could have an RNA binding domain, so it could actually act as a splicing factor. To test this possibility *in vitro*, the mature Gc16B protein lacking the chloroplast peptide was cloned into the pMAL-TEV vector (provided by Barkan A.). Gc16B was fused with its C-terminus to the Maltose Binding Protein and was overexpressed in *E.coli* cells. The protein was purified and cleaved with TEV protease. The detection of the band (52 kDa) in lane 1, Figure 36.A, shows the sample that has been used for further analysis.

To determine if the Gc16 proteins have an RNA-binding domain, Gc16B protein was used for electrophoretic mobility shift assay (EMSA). The intergenic region of *rbcL-accd* was used as radioactive probe and the proteins were Gc16B, a known RNA binding protein (ZMPPR6) and MBP for negative control. The assay was performed under non stringent conditions (0.05 mg/ml heparin). According to this assay there was an increase in the area where protein with RNA should run, that it appears as a smear in this case, at the same size where PPR6 with RNA runs and a simultaneous reduction of the free RNA amount when the protein concentration was increased (Figure 36.B). This result suggests that Gc16 could bind RNA and there is a domain in the protein, which performs this function.

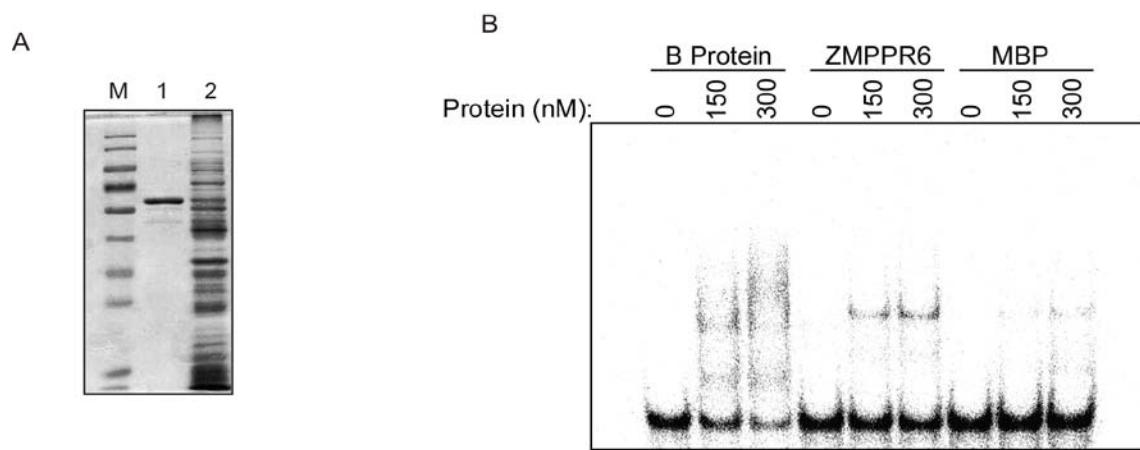


Figure 36. *In vitro* binding of Gc16B protein in *rbcL-accd* intergenic region.

A. Purification of recombinant Gc16 protein (Gc16B-MBP). The protein sample in lane 1 was used for EMSA analyses. B. Electrophoretic mobility shift assay (EMSA) was performed to

verify the RNA binding ability of recombinant Gc16B (B protein) with RNA from *rbcL-accd* intergenic region (cooperation with N. Manavski).

4. Discussion

4.1. Discussion *Synechocystis*

Recently, cyanobacteria have attracted a lot of attention as promising cell factories for products with industrial applications like biofuels. They use CO₂, H₂O and sunlight to yield organic molecules in a carbon-neutral fashion that is important for environmental reasons (Savakis & Hellingwerf 2015). *Synechocystis* can be regarded as a suitable industrial strain, because it grows fast and can be easily manipulated which is an important requirement for synthetic biology applications.

So far, several metabolic pathways have been introduced in different strains of cyanobacteria and they have successfully produced alcohols (Gao et al. 2012; Kusakabe et al. 2013), fatty acids (Liu et al. 2011) and more complex products like terpenoids (Lindberg et al. 2010a). The carotenoid pathway in *Synechocystis* has been engineered in two different approaches by overexpressing enzymes of the methylerythritol phosphate (MEP) pathway to increase the carotenoid amount and by introducing the mevalonic acid pathway (MVA), which led to isoprene overproduction, respectively (Kudoh et al. 2014; Bentley et al. 2014). Moreover, studies were conducted in *Synechocystis*, in which increased amounts of zeaxanthin and an altered carotenoid content was observable (Lagarde et al. 2000). So far, only endogenous pigment synthesis pathways have been manipulated, whereas no report is known about the introduction of genes for exogenous pigments from other organisms.

Carotenoids are important ingredients and are for that reason of biotechnological interest. Several studies aimed at engineering pigment biosynthesis pathways in plants, with the main focus on β-carotene production, a vitamin A precursor, and on xanthophyll production, like astaxanthin (Ye et al. 2000; Paine et al. 2005; Zhu et al. 2008). Lutein is an active antioxidant, presumably preventing age-related macular diseases or degenerate human diseases. Additionally, it is used in industry as food colorant (Carpentier et al. 2009; John et al. 2002). French marigold is today

the main source for lutein production, although the interest has been shifted to microalgae (like *Chlorella*) which accumulate lutein under autotrophic conditions (Fernández-Sevilla et al. 2010).

4.1.2. *Synechocystis* as a cell factory for lutein production

In this study, the enzymes from the *Arabidopsis* lutein pathway were introduced into *Synechocystis*. As a proof of concept, a colour complementation assay with *E.coli* was used to verify the functionality of the gene products that were selected. According to the HPLC analysis, lutein was produced in the lycopene-accumulating *E.coli* background strain when the four genes from *Arabidopsis* were expressed (Figure 4,5). It was evident though that β -carotene was more abundantly produced than lutein (0,8 nmol ml⁻¹ lutein and 13,8 nmol ml⁻¹ β carotene). Why does *E. coli* (+pAC-Lyc/pKS-LUT) accumulate more β -carotene? It is known that β -ring formation is a much more favorable reaction than ϵ -ring formation and that the two cyclases have different activities (Sun et al. 1996; Cunningham & Gantt 2001; Cunningham et al. 1996). Another reason why β -ring containing carotenoids are predominantly formed in *E. coli* (+pAC-Lyc/pKS-LUT) could be that the four genes are differently expressed leading to higher amounts of β -cyclases. As a conclusion, lutein can be produced by introducing the four-gene construct pKS-LUT into a bacterial host, which can provide lycopene as substrate.

When the pKS-LUT construct was used to transform *Synechocystis* wild type, no lutein was detected (Figure 7). Lutein and zeaxanthin compete for lycopene, since it is the precursor for the synthesis of both carotenoids. The ' β -carotene' branch is the endogenous pathway, which is probably the more favourable one for this reaction. Moreover, β -carotene and zeaxanthin are the two most abundant carotenoids in *Synechocystis*, which is a reflection of their importance in the metabolism of *Synechocystis*. In order to optimize lutein production, competing pathways using lycopene as a substrate were inactivated. To this end, the endogenous enzyme of lycopene cyclase, responsible for β -carotene production, was disrupted in the *Synechocystis* wild type strain.

Here, it was shown that *sll047- cruA* is responsible for the β -carotene production in this organism. According to HPLC analysis, there was a severe reduction of β -carotene in Δ *cruA* mutants, ranging from 25 to 40% compared to wild-type levels (Figure 10). The observation that β -carotene amounts were only reduced and not completely undetectable could be explained by the fact that the mutants are not fully segregated, probably because *Synechocystis* has to retain copies of this essential gene and because the gene product is the key enzyme for the synthesis of the two most abundant carotenoids. In addition, *cruA* might not be the only enzyme involved in this reaction and in its absence, other enzymes can compensate for *cruA* function. Nevertheless, upon disruption of *cruA* lycopene was found to accumulate, which was regarded as an important prerequisite for the production of the plant-specific pigment lutein (Figure 10).

Indeed, when in the strain with lutein pathway from *Arabidopsis* (*SynLUT*) *cruA* was disrupted (Δ *cruA*), lutein could be detected in low amounts. This result shows that heterologous expression of plant enzymes in *Synechocystis* was successful and that *cruA* disruption was an appropriate strategy to redirect lycopene usage to lutein production. The fact that lutein amounts were low can be explained by two possible reasons. One is that the β carotene and zeaxanthin were only reduced (around 40%) and there were still present in the cells, so there was not a need to produce new pigments in order to substitute the endogenous ones. The other is that lutein is degraded or not produced enough, because there is no endogenous cellular component in *Synechocystis* to interact with.

4.1.3. Strategies to increase lutein production in *Synechocystis*

Several strategies can be envisioned, which aim at the improvement of lutein production in *Synechocystis*. i) Disruption of *cruP*, which possibly catalyzes the same reaction as *cruA*, could lead to reduced β carotene and increased lycopene levels. ii) The enzyme with the lowest activity in the newly introduced pathway has to be identified and replaced by catalysts that are more active. The bottleneck for this pathway seems to be the epsilon cyclase *LCYe*, which acts as monoyclase and diverts the pathway to lutein

production. In maize, *LCYe* acts as a mono and a bicyclase in endosperm tissues and is not essential for its activity to form a complex with *LCYb* (Bai et al. 2009). Therefore, using *LCYe* from maize and not *LCYb* could lead to increased lutein production in *Synechocystis*. iii) Expression of endogenous or newly introduced genes can be optimized. A possibility to alter gene expression is the usage of different promoters for each gene or promoters derived from promoter libraries. It was shown that mutagenesis of constitutive promoters resulted in a variety of promoter strengths, which could alleviate the cells from stress and increase the production of the selected molecules (Jensen & Hammer 1998). Moreover, it was shown that using artificial instead of native promoters led to a significant increase of gene expression even suitable for industrial demands (Zhou et al. 2014; Albers et al. 2015). iv) Furthermore, different growth conditions, such as different light intensities, could lead to production of higher lutein levels.

4.1.4. Violaxanthin production in *Synechocystis*

Zeaxanthin epoxidase from *Arabidopsis* was also expressed in *Synechocystis* and according to pigment analysis violaxanthin was not synthesized. Instead two new carotenoids were produced in the mutants *SynZEP* that were not present in *Synechocystis* wild type (Figure 17, 18). Furthermore, when *ZEP* is using zeaxanthin, for further conversion of it to violaxanthin, it should be reduced since the pathway goes downstream, but this was not the case for the mutants. Zeaxanthin and β -carotene were increased in *SynZEP*, which could possibly mean that *ZEP* does not use zeaxanthin as substrate but another carotenoid and the increase in the amounts of them could be due to a reaction of the cells.

The two carotenoids produced in *SynZEP* could be either newly synthesized or degradation products. The function of *ZEP* is to introduce an epoxide in β rings in the ends of a polyene chain. Epoxy group addition is also a natural way for the degradation of xanthophylls, as for ABA synthesis. So, one possibility is that zeaxanthin epoxidase is using a different substrate than zeaxanthin and produces these carotenoids, or another that violaxanthin is produced but further degraded in other carotenoids. Xanthophyll cycle is

involved in thermal dissipation of excess light energy in plants and algae, but it is not present in cyanobacteria. The NPQ mechanism in cyanobacteria is triggered by the orange carotenoid protein, which acts as a photosensor when binds 3' hydroxyl echinenone. This complex goes under structural changes and induces NPQ through interaction with PBS (Jahns and Holzwarth 2011). Cells of *Synechocystis* do not need violaxanthin for energy dissipation and also there is no use of them as structural components, since there are no plant type light harvesting complexes. As a conclusion, the fact that violaxanthin was not produced could be because the enzyme does not function properly or because there is no need of it in the cells.

4.2. Discussion Gc16

4.2.1. RNA splicing mechanism in plants

The genome of plant organelles harbor introns that belong to two different groups. In plant chloroplasts, there are 20 introns divided into two groups (1 intron for group I and 19 introns in group II), which share a defined mechanism and intron structure. All introns are derived from self-splicing ribozymes, but in plant chloroplasts such a mechanism does not exist anymore. During evolution, these introns acquired the help of proteins, which are mainly nucleus encoded and play an accessory role in splicing without being part of the catalysis itself.

In the nucleus, the splicing of introns is dependent on the spliceosome, a large ribonucleoprotein complex which consists of small nuclear RNAs, small nuclear ribonucleoprotein particles and non snRNP particles (Sperling et al. 2008). The mechanism in plastids is protein dependent and each intron binds many different factors, which promote splicing. It is suggested that these factors help the intron to fold in a catalytically active structure or that they can stabilize its structure, as in the case of CRS1, which binds the *atpF* intron and induces intron folding in its proper form (Till et al. 2001; Ostheimer et al. 2003; Asakura & Barkan 2007). So far, 14 splicing factors have been identified in plastids to be involved in group II intron splicing (De Longevialle et al. 2010).

Most of these proteins belong to RNA binding protein families and are part of the organellar gene expression machinery (APO, CRM, PORR, RNC1, WTF) (Watkins et al. 2011; Asakura & Barkan 2007; Watkins et al. 2007; Kroeger et al. 2009).

Gc16A and Gc16B are proteins conserved in the entire green lineage, from green plants to algae, with the C-terminal domain to be the most conserved part among them (Figure 22,23). In this conserved part, the sequence contains a number of conserved negatively charged amino acids (glutamic acid- E), and a hydrophobic domain with multiple prolines (P). Furthermore, the alignment of the sequences with other proteins from plants revealed a close relationship with predicted proteins defined as splicing factors, containing an arginine/serine rich 19-like domain. This domain has been found in metazoan and is involved in RNA metabolism-related functions. Moreover, it is involved in nuclear intron splicing and is part of the spliceosome of these organisms (Haynes & Iakoucheva 2006). Even though Gc16A and B are targeted to the chloroplast in higher plants, the high similarity of these proteins with Gc16 proteins could suggest a possible RNA metabolism-related function in *Arabidopsis*.

4.2.2. Defects in NDH and Cytb_{6f} complexes in the absence of Gc16 proteins

Gc16 A and B seem to act synergistically and when at least one of them is present there is no obvious visible phenotype. But when both of them are absent a strong physiological effect could be observed. Gc16A seems to be the main allele, since upon its disruption a molecular phenotype was already detectable (Figure 29, 33).

Lack of Gc16 affects the abundance of the NDH and the Cytb_{6f} complex (Figure 28), as well as the activity of these complexes (Figure 32, Table 8). Although chloroplast protein levels are overall reduced in *gc16ab* mutants (adjusted to fresh weight), PSI and PSII are not as severely affected as the NDH and the Cytb_{6f} complex (Figure 29). It has previously been shown that Gc16B gene expression is up-regulated at the germination stage in *Arabidopsis* (Narsai et al. 2011). This is consistent with these results, since at

this stage the photosystems are not formed, but NDH and *Cytb₆f* complexes are already present in etioplasts and developing plastids (Fischer et al. 1997; Guéra et al. 2000). Therefore, expression studies substantiate the fact that Gc16 is present and highly active at an early developmental stage.

Previous studies for *ndh* mutants showed only subtle phenotypic differences, since cyclic electron flow is compensated by the PGR5-PGRL pathway (Shikanai 2014; Burrows et al. 1998; Endo et al. 1999; Peng & Shikanai 2011). In *gc16* mutants a retardation in growth was evident and it was more prominent in the double mutant (Figure 24.C,D). Therefore, an additional defect might explain the strong phenotype observed in *gc16* mutants. *Cytb₆f* mutants defect in splicing factors, CAF1 and CAF2, show an altered phenotype (Asakura & Barkan 2006). This could be, because *gc16* mutants show also a disruption in the cytochrome *b₆f* complex or because both of the complexes (also NDH) are affected. However, it cannot be excluded that Gc16 proteins exhibit additional functions. In mitochondria, intron-carrying genes encode also for subunits, which are part of the mitochondrial NDH (*nad1-2-4-5-7*), the cytochrome c oxidase complex (*cox2* and *ccmFc*) and two ribosomal proteins (Brown et al. 2014). So far, it cannot be excluded that Gc16 proteins are dual-targeted factors that might function in splicing of mitochondrial introns, as well. Indeed, Gc16 is present in *Chlamydomonas* with a clear mitochondrial signal peptide prediction (TargetP, <http://www.cbs.dtu.dk/services/TargetP/>). Moreover, *Chlamydomonas* misses a chloroplast NDH complex and - compared to plants - splicing mechanisms differ significantly in green algae, which argue against the fact that Gc16 proteins exhibit a conserved function exclusively located in plastids. There is one splicing factor identified in *Arabidopsis* so far, CFM3 (CRM family member 3), that in contrast with other factors was shown to be dual targeted. There are two paralogs in *Arabidopsis*, AtCFM3a is dual targeted while AtCFM3b is localized only in plastids (Asakura & Barkan 2007).

4.2.3. Gc16 - a putative RNA splicing factor?

From all plastidial genes coding for NDH complex subunits, a specific effect on transcript patterns could only be observed for *ndhA*, which leads to the

assumption that lack of Gc16 causes a splicing defect. According to Northern analysis with intron-specific probes a down-regulation of the excised intron concomitant with more precursor transcripts was observed (Figure 35). This effect was also found in *cfm-2* mutants , which are defective in *ndhA* and *ycf3* intron splicing (Asakura & Barkan 2007; Watkins et al. 2011). Hence, Gc16 proteins might be responsible for efficient intron splicing.

A similar effect could be accounted for *petB* but not for *petD* processing patterns (Figure 35). Even though processed *petD* levels were reduced, the effect was more general without the increase of precursor transcripts and the excised intron was increased. As a conclusion, disturbances in *ndhA* and *petB* processing were interpreted as a direct effect of the lack of Gc16, whereas alterations in *petD* transcript maturation and in the abundance of other transcripts (for instance *ndhF*) could be attributed to secondary effects.

Since splicing factors are not directly involved in catalytic reactions, Gc16 could be involved in intron splicing by binding to RNA, like other known splicing factors. Alternatively, it could indirectly affect this process by interacting with other splicing factors (Schmitz-Linneweber et al. 2006; Till et al. 2001; Jenkins & Barkan 2001; Watkins et al. 2011). CRS2 in maize does not bind to RNA but was found to interact with CAF1 and CAF2 to promote splicing (Ostheimer et al. 2003). However, according to preliminary EMSA results using *rbcL* transcripts, Gc16B protein is able to bind RNA (Figure 35). Thus, Gc16 proteins are putative RNA-binding proteins, which could be novel splicing factors in *Arabidopsis* chloroplasts involved in group II intron splicing. To further determine whether these two genes are the only targets of these proteins RIP-Chip analysis should follow (Keene et al. 2006). EMSA with potential targets might verify the putative interactions and also determine the sequence in the proteins that is the actual binding domain. Furthermore, an interaction of Gc16 proteins with other splicing factors, CFM2 and CRS2/CAF2 for *ndhA* and six different factors involved in *petB* splicing (chapter 5, figure 21), will clarify how these proteins function.

REFERENCES

- Albers, S.C., Gallegos, V.A. & Peebles, C.A.M., 2015. Engineering of genetic control tools in *Synechocystis* sp. PCC 6803 using rational design techniques. *Journal of biotechnology*, 216, pp.36–46
- Alonso, J.M. et al., 2003. Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science (New York, N.Y.)*, 301(5633), pp.653–657.
- Andersson, S.G.E. et al., 2003. On the origin of mitochondria: a genomics perspective. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 358(1429), pp.165–177
- Armstrong, G. a, 1997. Genetics of eubacterial carotenoid biosynthesis: a colorful tale. *Annual review of microbiology*, 51, pp.629–59
- Arteni, A. a., Ajlani, G. & Boekema, E.J., 2009. Structural organisation of phycobilisomes from *Synechocystis* sp. strain PCC6803 and their interaction with the membrane. *Biochimica et Biophysica Acta - Bioenergetics*, 1787(4), pp.272–279
- Asakura, Y. & Barkan, A., 2007. A CRM domain protein functions dually in group I and group II intron splicing in land plant chloroplasts. *The Plant cell*, 19(12), pp.3864–3875.
- Asakura, Y. & Barkan, A., 2006. Arabidopsis orthologs of maize chloroplast splicing factors promote splicing of orthologous and species-specific group II introns. *Plant physiology*, 142(4), pp.1656–1663.
- Bai, L. et al., 2009. Novel lycopene epsilon cyclase activities in maize revealed through perturbation of carotenoid biosynthesis. *The Plant Journal*, 59(4), pp.588–599.
- Barkan, a & Goldschmidt-Clermont, M., 2000. Participation of nuclear genes in chloroplast gene expression. *Biochimie*, 82(6-7), pp.559–572.
- Barros, T. & Kühlbrandt, W., 2009. Crystallisation, structure and function of plant light-harvesting Complex II. *Biochimica et biophysica acta*, 1787(6), pp.753–72.
- Beckmann, J. et al., 2009. Improvement of light to biomass conversion by de-regulation of light-harvesting protein translation in *Chlamydomonas reinhardtii*. *Journal of biotechnology*, 142(1), pp.70–7.

- Bennett-Lovsey, R. et al., 2002. The SWIB and the MDM2 domains are homologous and share a common fold. *Bioinformatics (Oxford, England)*, 18(4), pp.626–630.
- Bentley, F.K., Zurbriggen, A. & Melis, A., 2014. Heterologous expression of the mevalonic acid pathway in cyanobacteria enhances endogenous carbon partitioning to isoprene. *Molecular Plant*, 7(1), pp.71–86.
- Blankenship, R.E. et al., 2011. Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science (New York, N.Y.)*, 332(6031), pp.805–809.
- Bolle, C. et al., 2013. GABI-DUPLO: A collection of double mutants to overcome genetic redundancy in *Arabidopsis thaliana*. *Plant Journal*, 75(1), pp.157–171.
- Bollenbach, T.J. et al., 2005. RNR1, a'3-5' exoribonuclease belonging to the RNR superfamily, catalyzes 3' maturation of chloroplast ribosomal RNAs in *Arabidopsis thaliana*. *Nucleic Acids Research*, 33(8), pp.2751–2763.
- Boronowsky, U. et al., 2001. Isolation of membrane protein subunits in their native state: Evidence for selective binding of chlorophyll and carotenoid to the b6 subunit of the cytochrome b6f complex. *Biochimica et Biophysica Acta - Bioenergetics*, 1506(1), pp.55–66.
- Bouvier, F. et al., 1996. Xanthophyll Biosynthesis. *Journal of Biological Chemistry*, 271(46), pp.28861–28867.
- Brown, G.G., Colas des Francs-Small, C. & Ostersetzer-Biran, O., 2014. Group II intron splicing factors in plant mitochondria. *Frontiers in plant science*, 5(February), p.35.
- Bugos, R.C., Hieber, A.D. & Yamamoto, H.Y., 1998. Xanthophyll cycle enzymes are members of the lipocalin family, the first identified from plants. *The Journal of biological chemistry*, 273(25), pp.15321–4.
- Burrows, P. a. et al., 1998. Identification of a functional respiratory complex in chloroplasts through analysis of tobacco mutants containing disrupted plastid ndh genes. *EMBO Journal*, 17(4), pp.868–876.
- Cai, Y.P. & Wolk, C.P., 1990. Use of a conditionally lethal gene in *Anabaena* sp. strain PCC 7120 to select for double recombinants and to entrap insertion sequences. *Journal of bacteriology*, 172(6), pp.3138–45.

del Campo, E.M., 2009. Post-Transcriptional Control of Chloroplast Gene Expression. *Gene regulation and Systems Biology* (3), pp.31–47.

Carpentier, S., Knaus, M. & Suh, M., 2009. Associations between Lutein, Zeaxanthin, and Age-Related Macular Degeneration: An Overview. *Critical Reviews in Food Science and Nutrition*, 49(4), pp.313–326.

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

Collins, A.M. et al., 2012. *Photosynthetic pigment localization and thylakoid membrane morphology are altered in synechocystis 6803 phycobilisome mutants*. *Plant physiology* 158 (4), pp 1600-1609.

Croce, R. & Weiss, S., 1999. MEMBRANES AND BIOENERGETICS: Carotenoid -binding Sites of the Major Light-harvesting Complex II of Higher Plants Carotenoid-binding Sites of the Major Light-harvesting Complex II of Higher Plants, 274(42), pp.29613–29623.

Cunningham, F.X. et al., 1994. Molecular structure and enzymatic function of lycopene cyclase from the cyanobacterium *Synechococcus* sp strain PCC7942. *The Plant cell*, 6(8), pp.1107–21.

Cunningham, F.X. & Gantt, E., 1998. Genes and Enzymes of Carotenoid Biosynthesis in Plants. *Annual review of plant physiology and plant molecular biology*, 49, pp.557–583.

Cunningham, F.X. & Gantt, E., 2001. One ring or two? Determination of ring number in carotenoids by lycopene -cyclases. *PNAS*, 98(5), pp.2905–2910.

Cunningham, F.X., Pogson, B.J. & Sun, Z., 1996. Functional Analysis of the B and E Lycopene Cyclase Enzymes of *Arabidopsis* Reveals a Mechanism for Control of Cyclic Carotenoid Formation. *Society*, 8(September), pp.1613–1626.

Demmig-Adams, B. & Adams, W.W., 1992. Carotenoid Composition in Sun and Shade Leaves of Plants with Different Life Forms. *Plant Cell and Environment*, 15(4), pp.411–419.

Deng, X.W. & Grussem, W., 1987. Control of plastid gene expression during development: the limited role of transcriptional regulation. *Cell*, 49(3), pp.379–387.

Endo, T. et al., 1999. The role of chloroplastic NAD(P)H dehydrogenase in photoprotection. *FEBS Letters*, 457(1), pp.5–8.

Engler, C. et al., 2009. Golden gate shuffling: a one-pot DNA shuffling method based on type IIIs restriction enzymes. *PLoS one*, 4(5), p.e5553.

Engler, C., Kandzia, R. & Marillonnet, S., 2008. A one pot, one step, precision cloning method with high throughput capability. *PLoS one*, 3(11), p.e3647.

Farber, a. et al., 1997. Dynamics of Xanthophyll-Cycle Activity in Different Antenna Subcomplexes in the Photosynthetic Membranes of Higher Plants (The Relationship between Zeaxanthin Conversion and Nonphotochemical Fluorescence Quenching). *Plant physiology*, 115(4), pp.1609–1618.

Fernández-Sevilla, J.M., Acién Fernández, F.G. & Molina Grima, E., 2010. Biotechnological production of lutein and its applications. *Applied Microbiology and Biotechnology*, 86(1), pp.27–40.

Fischer, M., Fun K, E. & Steinmuller, K., 1997. The expression of subunits of the mitochondrial complex I-homologous NAD(P)H-plastoquinone-oxidoreductase during plastid development. *Zeitschrift für Naturforschung. C, Journal of biosciences*, 52(7-8), pp.481–486.

Gao, Z. et al., 2012. Photosynthetic production of ethanol from carbon dioxide in genetically engineered cyanobacteria. *Energy & Environmental Science*, pp.9857–9865.

Goodwin, T.W., 1980. *The Biochemistry of the Carotenoids*,

Grossman, A.R. et al., 2010. Phylogenomic analysis of the Chlamydomonas genome unmasks proteins potentially involved in photosynthetic function and regulation. *Photosynthesis research*, 106(1-2), pp.3–17.

Guéra, a, de Nova, P.G. & Sabater, B., 2000. Identification of the Ndh (NAD(P)H-plastoquinone-oxidoreductase) complex in etioplast membranes of barley: changes during photomorphogenesis of chloroplasts. *Plant & cell physiology*, 41(1), pp.49–59.

Hager, a. & Holocher, K., 1994. Localization of the xanthophyll-cycle enzyme violaxanthin de-epoxidase within the thylakoid lumen and abolition of its mobility by a (light-dependent) pH decrease. *Planta*, 192(4), pp.581–589.

Hajdukiewicz, P.T.J., Allison, L. a. & Maliga, P., 1997. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. *EMBO Journal*, 16(13), pp.4041–4048.

Harry Y. Yamamoto, Lavonne Kamite, Y.-Y.W., 1972. An ascorbate-induced absorbance change in chloroplasts from Violaxanthin De-epoxidation. , pp.224–228.

Haugen, P., Simon, D.M. & Bhattacharya, D., 2005. The natural history of group I introns. *Trends in Genetics*, 21(2), pp.111–119.

Haynes, C. & Iakoucheva, L.M., 2006. Serine/arginine-rich splicing factors belong to a class of intrinsically disordered proteins. *Nucleic Acids Research*, 34(1), pp.305–312.

He, Q. et al., 1999. Expression of a higher plant light-harvesting chlorophyll a / b-binding protein in Synechocystis sp . PCC 6803. *Culture*, 570, pp.561–570.

He, Q. et al., 1999. Expression of a higher plant light-harvesting chlorophyll a/b-binding protein in Synechocystis sp. PCC 6803. *European journal of biochemistry / FEBS*, 263(2), pp.561–70.

Horváth, E.M. et al., 2000. Targeted inactivation of the plastid ndhB gene in tobacco results in an enhanced sensitivity of photosynthesis to moderate stomatal closure. *Plant physiology*, 123(4), pp.1337–1350.

Ifuku, K. et al., 2011. Structure of the chloroplast NADH dehydrogenase-like complex: Nomenclature for nuclear-encoded subunits. *Plant and Cell Physiology*, 52(9), pp.1560–1568.

Jahns, P. & Holzwarth, A.R., 2011. The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. *BBA - Bioenergetics*, 1817(1), pp.182–193.

Jenkins, B.D. & Barkan, a., 2001. Recruitment of a peptidyl-tRNA hydrolase as a facilitator of group II intron splicing in chloroplasts. *EMBO Journal*, 20(4), pp.872–879.

Jensen, P.R. & Hammer, K., 1998. Artificial promoters for metabolic optimization. *Biotechnology and bioengineering*, 58(2-3), pp.191–5.

John, J.H. et al., 2002. Effects of fruit and vegetable consumption on plasma antioxidant concentrations and blood pressure: a randomised controlled trial. *Lancet*,

359(9322), pp.1969–1974.

Joliot, P. & Joliot, A., 2006. Cyclic electron flow in C₃ plants. *Biochimica et biophysica acta*, 1757(5-6), pp.362–368.

Kane, J.F., 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coli*. *Current Opinion in Biotechnology*, 6(5), pp.494–500

Kaneko, T. & Tabata, S., 1997. Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803. *Plant & cell physiology*, 38(11), pp.1171–1176.

Karimi, M., Inzé, D. & Depicker, A., 2002. GATEWAY™ vectors for Agrobacterium-mediated plant transformation. *Trends in Plant Science*, 7(5), pp.193–195.

Karpowicz, S.J. et al., 2011. The GreenCut2 Resource, a Phylogenomically Derived Inventory of Proteins Specific to the Plant Lineage. *The Journal of biological chemistry*, 286(24), pp.21427–39.

Keene, J.D., Komisarow, J.M. & Friedersdorf, M.B., 2006. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nature protocols*, 1(1), pp.302–307.

Kelly, J.R. et al., 2009. Measuring the activity of BioBrick promoters using an *in vivo* reference standard. *Journal of biological engineering*, 3, p.4.

Kim, J. & DellaPenna, D., 2006. Defining the primary route for lutein synthesis in plants: the role of *Arabidopsis* carotenoid beta-ring hydroxylase CYP97A3. *Proceedings of the National Academy of Sciences of the United States of America*, 103(9), pp.3474–9.

Kleine, T., Maier, U.G. & Leister, D., 2009. DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. *Annual review of plant biology*, 60, pp.115–138.

Kofer, W. et al., 1998. Mutagenesis of the genes encoding subunits A, C, H, I, J and K of the plastid NAD(P)H-plastoquinone-oxidoreductase in tobacco by polyethylene glycol-mediated plastome transformation. *Molecular and General Genetics MGG*, 258(1-2), pp.166–173.

- Kogata, N. et al., 1999. Involvement of a chloroplast homologue of the signal recognition particle receptor protein, FtsY, in protein targeting to thylakoids. *FEBS Letters*, 447(2-3), pp.329–333.
- Kroeger, T.S. et al., 2009. A plant-specific RNA-binding domain revealed through analysis of chloroplast group II intron splicing. *Proceedings of the National Academy of Sciences of the United States of America*, 106(11), pp.4537–4542.
- Kudoh, K. et al., 2014. Prerequisite for highly efficient isoprenoid production by cyanobacteria discovered through the over-expression of 1-deoxy-d-xylulose 5-phosphate synthase and carbon allocation analysis. *Journal of bioscience and bioengineering*, 118(1), pp.20–8.
- Kunert, A, Hagemann, M. & Erdmann, N., 2000. Construction of promoter probe vectors for Synechocystis sp. PCC 6803 using the light-emitting reporter systems Gfp and LuxAB. *Journal of microbiological methods*, 41(3), pp.185–94.
- Kusakabe, T. et al., 2013. Engineering a synthetic pathway in cyanobacteria for isopropanol production directly from carbon dioxide and light. *Metabolic Engineering*, 20, pp.101–108.
- Lagarde, D., Beuf, L. & Vermaas, W., 2000. Increased production of zeaxanthin and other pigments by application of genetic engineering techniques to Synechocystis sp. strain PCC 6803. *Applied and environmental microbiology*, 66(1), pp.64–72.
- Li, X. et al., 1995. A chloroplast homologue of the signal recognition particle subunit SRP54 is involved in the posttranslational integration of a protein into thylakoid membranes. *Proceedings of the National Academy of Sciences of the United States of America*, 92(9), pp.3789–3793.
- Li, Y. et al., 2003. GABI-Kat SimpleSearch: A flanking sequence tag (FST) database for the identification of T-DNA insertion mutants in *Arabidopsis thaliana*. *Bioinformatics*, 19(11), pp.1441–1442.
- Liang, C. et al., 2006. Carotenoid biosynthesis in cyanobacteria: structural and evolutionary scenarios based on comparative genomics. *International journal of biological sciences*, 2(4), pp.197–207.
- Lindberg, P., Park, S. & Melis, A., 2010a. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism.

Metabolic Engineering, 12(1), pp.70–79.

Lindberg, P., Park, S. & Melis, A., 2010b. Engineering a platform for photosynthetic isoprene production in cyanobacteria, using *Synechocystis* as the model organism. *Metabolic engineering*, 12(1), pp.70–9.

Liu, X., Sheng, J. & Curtiss, R., 2011. Fatty acid production in genetically modified cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 108(17), pp.6899–904.

Liu, Z. et al., 2004. Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature*, 428(6980), pp.287–292.

De Longevialle, A.F. et al., 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 Journal*, 56(1), pp.157–168.

De Longevialle, A.F., Small, I.D. & Lurin, C., 2010. Nuclearly encoded splicing factors implicated in RNA splicing in higher plant organelles. *Molecular Plant*, 3(4), pp.691–705.

MacColl, R., 1998. Cyanobacterial phycobilisomes. *Journal of structural biology*, 124(2-3), pp.311–34.

Maier, R.M. et al., 1995. Complete sequence of the maize chloroplast genome: gene content, hotspots of divergence and fine tuning of genetic information by transcript editing. *Journal of molecular biology*, 251(5), pp.614–628.

Manavski, N. et al., 2012. An Essential Pentatricopeptide Repeat Protein Facilitates 5' Maturation and Translation Initiation of rps3 mRNA in Maize Mitochondria. *The Plant Cell*, 24(7), pp.3087–3105.

Maresca, J.A. et al., 2007. Identification of a fourth family of lycopene cyclases in photosynthetic bacteria. *PNAS*, 104(28), pp.11784–11789.

Melis, A., 2012. Photosynthesis-to-fuels: from sunlight to hydrogen, isoprene, and botryococcene production. *Energy & Environmental Science*, 5(2), p.5531.

Melis, A., 2009. Solar energy conversion efficiencies in photosynthesis: Minimizing the chlorophyll antennae to maximize efficiency. *Plant Science*, 177(4), pp.272–280.

Merchant, S.S. et al., 2007. The *Chlamydomonas* genome reveals the evolution of

- key animal and plant functions. *Science (New York, N.Y.)*, 318(5848), pp.245–50.
- Michel, F. & Ferat, J.L., 1995. Structure and activities of group II introns. *Annual review of biochemistry*, 64, pp.435–461.
- Mohamed, A. et al., 1993. Differential expression of the psbA genes in the cyanobacterium Synechocystis 6803. *Methods in Enzymology*, 3, pp.161–168.
- Mohamed, H.E. et al., 2005. Myxoxanthophyll is required for normal cell wall structure and thylakoid organization in the cyanobacterium Synechocystis sp. strain PCC 6803. *Journal of bacteriology*, 187(20), pp.6883–6892.
- Munekage, Y. et al., 2004. Cyclic electron flow around photosystem I is essential for photosynthesis. *Nature*, 429(6991), pp.579–582.
- Narsai, R. et al., 2011. In-Depth Temporal Transcriptome Profiling Reveals a Crucial Developmental Switch with Roles for RNA Processing and Organelle Metabolism That Are Essential for Germination in Arabidopsis. *Plant Physiology*, 157(3), pp.1342–1362.
- Oliver, J.W.K. & Atsumi, S., 2014. Metabolic design for cyanobacterial chemical synthesis. *Photosynthesis Research*, 120(3), pp.249–261.
- Ostheimer, G.J. et al., 2003. Group II intron splicing factors derived by diversification of an ancient RNA-binding domain. *EMBO Journal*, 22(15), pp.3919–3929.
- Page, L.E., Liberton, M. & Pakrasi, H.B., 2012. Reduction of photoautotrophic productivity in the cyanobacterium Synechocystis sp. strain PCC 6803 by phycobilisome antenna truncation. *Applied and Environmental Microbiology*, 78(17), pp.6349–6351.
- Paine, J.A. et al., 2005. Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nature Biotechnology*, 23(4), pp.482–487.
- Peng, L. & Shikanai, T., 2011. Supercomplex formation with photosystem I is required for the stabilization of the chloroplast NADH dehydrogenase-like complex in Arabidopsis. *Plant physiology*, 155(4), pp.1629–1639.
- Pesaresi, P., 2011. Studying Translation in Arabidopsis Chloroplasts. In R. P. Jarvis, ed. *Chloroplast Research in Arabidopsis SE - 14. Methods in Molecular Biology*. Humana Press, pp. 209–224.

- Porra, R.J., Thompson, W. a. & Kriedemann, P.E., 1989. Determination of accurate extinction coefficients and simultaneous equations for assaying chlorophylls a and b extracted with four different solvents: verification of the concentration of chlorophyll standards by atomic absorption spectroscopy. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 975(3), pp.384–394.
- Quinlan, R.F. et al., 2012. Synergistic Interactions between Carotene Ring Hydroxylases Drive Lutein Formation in Plant Carotenoid Biosynthesis. *Plant Physiology*, 160(1), pp.204–214.
- Raven, J. a & Allen, J.F., 2003. Genomics and chloroplast evolution: what did cyanobacteria do for plants? *Genome biology*, 4(3), p.209.
- Rippka, R. & Deruelles, J., 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Journal of General*, 110(2), pp.1–61.
- Rittmann, B.E., 2008. Opportunities for renewable bioenergy using microorganisms. *Biotechnology and Bioengineering*, 100(2), pp.203–212.
- Rühle, T. et al., 2014. The Arabidopsis protein CONSERVED ONLY IN THE GREEN LINEAGE160 promotes the assembly of the membranous part of the chloroplast ATP synthase. *Plant physiology*, 165(1), pp.207–26.
- Ruiz-Sola, M., Rodriguez-Concepcion, M., 2012. Carotenoid Biosynthesis in Arabidopsis: A Colorful Pathway. *The Arabidopsis Book*, pp.1–28.
- Sambrook, J. and Russel, D.W., 2001. *Molecular Cloning: A laboratory manual*,
- Sandmann, G., 1994. Carotenoid biosynthesis in microorganisms and plants. *European journal of biochemistry / FEBS*, 223(1), pp.7–24.
- Savakis, P. & Hellingwerf, K.J., 2015. Engineering cyanobacteria for direct biofuel production from CO₂. *Current opinion in biotechnology*, 33, pp.8–14.
- Schäfer, L., Vioque, A. & Sandmann, G., 2005. Functional in situ evaluation of photosynthesis-protecting carotenoids in mutants of the cyanobacterium *Synechocystis* PCC6803. *Journal of photochemistry and photobiology. B, Biology*, 78(3), pp.195–201.
- Schägger, H., 2006. Tricine-SDS-PAGE. *Nature protocols*, 1(1), pp.16–22.
- Schägger, H., Cramer, W. a & 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. *Analytical biochemistry*, 217(2), pp.220–230.

Schägger, H. & von Jagow, G., 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical biochemistry*, 166(2), pp.368–379.

Schmid, V.H.R., 2008. Light-harvesting complexes of vascular plants. *Cellular and molecular life sciences : CMLS*, 65(22), pp.3619–39.

Schmitz-Linneweber, C. et al., 2006. A pentatricopeptide repeat protein facilitates the trans-splicing of the maize chloroplast rps12 pre-mRNA. *The Plant cell*, 18(10), pp.2650–2663.

Schuenemann, D. et al., 1998. A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. *Proceedings of the National Academy of Sciences*, 95(17), pp.10312–01316.

Sessions, A., Burke, E. & Presting, G., 2002. High-Throughput Arabidopsis Reverse Genetics System. *The Plant Cell*, 14(December), pp.2985–2994.

Shikanai, T., 2014. Central role of cyclic electron transport around photosystem I in the regulation of photosynthesis. *Current Opinion in Biotechnology*, 26(Figure 1), pp.25–30.

Shikanai, T. et al., 1998. Directed disruption of the tobacco *ndhB* gene impairs cyclic electron flow around photosystem I. *Proceedings of the National Academy of Sciences of the United States of America*, 95(16), pp.9705–9709.

Siefermann, D. & Yamamoto, H.Y., 1975. Properties of NADPH and oxygen-dependent zeaxanthin epoxidation in isolated chloroplasts. A transmembrane model for the violaxanthin cycle. *Archives of biochemistry and biophysics*, 171(1), pp.70–77.

Sperling, J., Azubel, M. & Sperling, R., 2008. Structure and function of the Pre-mRNA splicing machine. *Structure (London, England : 1993)*, 16(11), pp.1605–15.

Stoppel, R. et al., 2012. RHON1 is a novel ribonucleic acid-binding protein that supports RNase e function in the Arabidopsis chloroplast. *Nucleic Acids Research*, 40(17), pp.8593–8606.

- Stoppel, R. & Meurer, J., 2013. Complex RNA metabolism in the chloroplast: An update on the psbB operon. *Planta*, 237(2), pp.441–449.
- Sun, Z., Gantt, E. & Cunningham, F.X., 1996. Cloning and functional analysis of the beta-carotene hydroxylase of *Arabidopsis thaliana*. *The Journal of biological chemistry*, 271(40), pp.24349–52.
- Sundberg, E. et al., 1997. ALBINO3, an *Arabidopsis* nuclear gene essential for chloroplast differentiation, encodes a chloroplast protein that shows homology to proteins present in bacterial membranes and yeast mitochondria. *The Plant cell*, 9(5), pp.717–730.
- Takaichi, S., Maoka, T. & Masamoto, K., 2001. Myoxanthophyll in *Synechocystis* sp. PCC 6803 is myxol 2'-dimethyl-fucoside, (3R,2'S)-myxol 2'-(2,4-di-O-methyl-alpha-L-fucoside), not rhamnoside. *Plant & cell physiology*, 42(7), pp.756–762.
- Takaichi, S. & Mochimaru, M., 2007. Carotenoids and carotenogenesis in cyanobacteria: unique ketocarotenoids and carotenoid glycosides. *Cellular and molecular life sciences*, 64(19-20), pp.2607 – 2619.
- Tian, L. & DellaPenna, D., 2001. Characterization of a second carotenoid beta-hydroxylase gene from *Arabidopsis* and its relationship to the LUT1 locus. *Plant molecular biology*, 47(3), pp.379–88.
- Till, B. et al., 2001. CRS1 is a novel group II intron splicing factor that was derived from a domain of ancient origin. *RNA (New York, N.Y.)*, 7(9), pp.1227–1238.
- Tillett, D. & Neilan, B.A., 2000. Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. *J. Phycol.*, 36, pp.251–258.
- Timmis, J.N. et al., 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature reviews. Genetics*, 5(2), pp.123–135.
- Viola, S., Röhle, T. & Leister, D., 2014. A single vector-based strategy for marker-less gene replacement in *Synechocystis* sp. PCC 6803. *Microbial cell factories*, 13, p.4.
- Vogel, J., Börner, T. & Hess, W.R., 1999. Comparative analysis of splicing of the complete set of chloroplast group II introns in three higher plant mutants. *Nucleic Acids Research*, 27(19), pp.3866–3874.

- Watkins, K.P. et al., 2007. A ribonuclease III domain protein functions in group II intron splicing in maize chloroplasts. *The Plant cell*, 19(8), pp.2606–2623.
- Watkins, K.P. et al., 2011. APO1 promotes the splicing of chloroplast group II introns and harbors a plant-specific zinc-dependent RNA binding domain. *The Plant cell*, 23(3), pp.1082–1092.
- Werck-Reichhart, D., Bak, S. & Paquette, S., 2002. Cytochromes P450. *The Arabidopsis Book*, 2(1), p.1.
- Williams, J.G.K.K., 1988. Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in Synechocystis 6803. *Methods in Enzymology*, 167(1985), pp.766–778.
- Wu, F.-H. et al., 2009. Tape-Arabidopsis Sandwich - a simpler Arabidopsis protoplast isolation method. *Plant methods*, 5, p.16.
- Xu, H., Vavilin, D. & Vermaas, W., 2001. Chlorophyll b can serve as the major pigment in functional photosystem II complexes of cyanobacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 98(24), pp.14168–73.
- Ye, X. et al., 2000. Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science (New York, N.Y.)*, 287(5451), pp.303–305.
- Zhou, J. et al., 2014. Discovery of a super-strong promoter enables efficient production of heterologous proteins in cyanobacteria. *Scientific Reports*, 4, pp.1–6.
- Zhu, C. et al., 2008. Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proceedings of the National Academy of Sciences of the United States of America*, 105(47), pp.18232–7.
- Zoschke, R. et al., 2010. An organellar maturase associates with multiple group II introns. *Proceedings of the National Academy of Sciences of the United States of America*, 107(7), pp.3245–3250.

Acknowledgements

I would like to thank Prof. Dr. Dario Leister for giving me the opportunity and the funding to accomplish my Ph.D. in his research group.

Thanks to Prof. Dr. Peter Geigenberger for taking over the “Zweitgutachter”.

I dearly want to thank Dr. Thilo Rühle for the useful supervision and the motivation, support and good ideas he gave me throughout my entire Ph.D.

A special thanks to Nico for being my friend, colleague, translator and many more. I consider you as a lifelong friend.

Thanks to Francesca for the friendship and all the nice memories.

Thanks to Mathias, my office mate, for helping me with all the German authorities and also sharing many nice moments.

Thanks to Luca for all the help and interesting conversations.

Thanks to all the people in the lab for the years spent together with the bad and good moments.

A special thank to my mum and dad for everything they have done for me.

To Manolis, my better half, who supports me through everything.

Curriculum vitae

Personal details:

Name: Evgenia Vamvaka

Date of birth: 21.03.1981

Place of birth: Ptolemaida, Greece

Citizenship: Greek

Education:

February 2005- July 2008

Bachelor degree in Agriculture. Final grade: 7,29/10

University of Thessaloniki, Greece

October 2008- January 2011

Master degree in Molecular Biology and Biotechnology. Final grade: 9,49/10

University of Crete, Greece

Research experience:

February 2008-July 2008

Research project: "*Diurnal changes in photosynthesis and phenolics in Laurel and Oleander leaves*".

Supervisor Prof. E.I. Konstantinidou. University of Thessaloniki, Greece

October 2009- January 2011

Research project: "*In vivo analysis of the plant 3'-5' exonuclease ERL-1*".
Supervisor Prof. K. Kalantidis. University of Crete, Greece.

Publications:

S. Grahl, B. Reiter, I.L. Gügel, E. Vamvaka, C. Gandini, P.Jahns, J.Soll, D. Leister, T. Röhle. (2016) The Arabidopsis protein CGLD11 is required for chloroplast ATP synthase accumulation. *Molecular Plant*

Röhle T., Razeghi J., Vamvaka E., Viola S., Gandini C., Kleine T., Schünemann D., Barbato R., Jahns P., Leister D. (2014) The Arabidopsis Protein CONSERVED ONLY IN THE GREEN LINEAGE160 Promotes the Assembly of the Membranous Part of the Chloroplast ATP Synthase. *Plant Phys.*

Declaration / Eidesstattliche Erklärung

Ich versichere hiermit an Eides statt, dass die vorgelegte Dissertation von mir selbstständig und ohne unerlaubte Hilfe angefertigt ist.

München, den

Evgenia Vamvaka