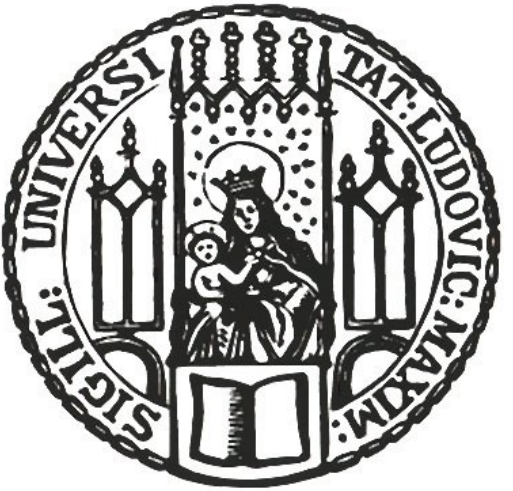
Molecular analysis of nucleus-encoded regulators of plastid protein synthesis



Dissertation

zur Erlangung des Grades eines Doktors der Naturwissenschaften
an der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

vorgelegt von
Laura Kleinknecht

München, 15.02.2018

Erstgutachter: Prof. Dr. Jörg Nickelsen

Zweitgutachter: Prof. Dr. Peter Geigenberger

Tag der Abgabe: 15.02.2018

Tag der mündlichen Prüfung: 16.05.2018

Table of Content

Abbreviations

2D-BN-PAGE	2 dimensional blue native polyacrylamide gel electrophoresis
ADP	adenosine diphosphate
A. thaliana	*Arabidopsis thaliana*
ATP	adenosine triphosphate
bp	base pairs
cDNA	complementary deoxyribonucleic acid
CES	control by epistasy of synthesis
cpm	counts per minute
cpPDC	pyruvate dehydrogenase complex of the chloroplast
Da	dalton
DNA	deoxyribonucleic acid
HMW	high molecular weight
kb	kilobase(s)
kDa	kilo dalton
knt	kilonucleotide(s)
mRNA	messenger RNA
$NAD^+(H)$	nicotinamide adenine dinucleotide (oxidized and reduced form)
$NADP^+(H)$	Nicotinamide adenine dinucleotide phosphate (oxidized and reduced form)
NEP	nuclear encoded (plastidial) RNA-Polymerase
nt	nucleotide(s)
(d)NTP	(Deoxy) nuclesidetriphosphate
OD	optical Density
OPR	octotricopeptide repeat
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PEP	plastid encoded (plastidial) RNA-Polymerase
P_i	phosphate
PPR	pentatricopeptide repeat
PSI	photosystem I
PSII	photosystem II
PDH	Pyruvatdehydrogenase
PDC	pyruvate dehydrogenase complex

RBP	RNA binding protein
RNA	ribonucleic acid
RNase	ribonuclease
rrn operon	ribosomal RNA operon
rRNA	ribosomal RNA
sRNA	small RNA
T-DNA	transfer DNA
TPR	tetratricopeptide repeat
tRNA	transfer RNA
UTR	untranslated region
WT	wild type
µ	Micro

Zusammenfassung

Zahlreiche kernkodierte Proteine regulieren die Synthese und Assemblierung plastidärer Proteine und Proteinkomplexe. Hierbei dominieren besonders die Prozesse auf post-transkriptioneller Ebene zur Kontrolle der Genexpression. Das Ziel dieser Arbeit war es, zwei dieser Faktoren, RAP und DLA2, aus der höheren Pflanze *Arabidopsis thaliana* beziehungsweise der Grünalge *Chlamydomonas reinhardtii* näher zu analysieren.

Das Protein RAP gehört zur Superfamilie der *helical repeat* Proteine, die einen Großteil der RNA-bindenden Faktoren in Pflanzenorganellen ausmachen. Interessanterweise ist RAP das einzige identifizierte OPR (*octotricopeptide repeat*) Protein in Arabidopsis, das bisher nur im Zusammenhang mit der Abwehr von pflanzlichen Pathogenen beschrieben wurde. Mit Hilfe einer T-DNA Insertionslinie wurde aufgedeckt, dass RAP eine Rolle in der korrekten und effizienten Prozessierung der *16S* rRNA spielt. Während die Menge an reifer *16S* rRNA deutlich reduziert ist, kommt es zu einer Anreicherung von Vorläufern. Dieser Defekt führt zu einer verminderten Synthese und Anreicherung von chloroplastenkodierten Proteinen, die z.B. für Prozesse wie die Photosynthese wichtig sind. Zusammenfassend kann man sagen, dass das einzige OPR Protein aus Arabidopsis eine wichtige Rolle in einem grundlegenden Prozess der Chloroplastenbiogenese spielt.

Das *moonlighting* Enzym DLA2 ist zum einen eine aktive Untereinheit des plastidären Pyruvatdehydrogenase (PDH) Komplexes und zum anderen ein RNA-bindendes Protein, das in die Genexpression des Chloroplasten involviert ist. Diese zweite Funktion tritt nur unter mixotrophen Wachstumsbedingungen, das heißt in der Anwesenheit von Licht und Acetat, zu Tage. Hierfür wurde gezeigt, dass die PDH Komplex Untereinheiten E1 und E3 nicht Teil des DLA2-RNA Komplexes sind. Die Tatsache, dass sich das Verhältnis zwischen den Untereinheiten unter verschiedenen Wachstumsbedingungen nicht ändert, unterstützt die Vermutung, dass der PDH Komplex unter mixotrophen Wachstumsbedingungen zumindest teilweise disassembliert und somit DLA2 für seine zweite Funktion freigibt. Bemerkenswerterweise ist die Bindung von E3 und der RNA an DLA2 kompetitiv. Eine Analyse des Acetyloms von *C. reinhardtii* offenbarte, dass die Acetylierung eines Lysines in der putativen RNA-Binderegion von DLA2 unter mixotrophen Bedingungen hochreguliert ist. Diese Erkenntnisse unterstützen die Hypothese, dass die zwei Funktionen von DLA2 miteinander konkurrieren und möglicherweise durch eine post-translationale Modifikation reguliert werden. Abschließend kann man sagen, dass die Funktion zweier kernkodierter

Proteine, die an der Synthese von plastidären Proteinen beteiligt sind, genauer untersucht wurde und somit dazu beiträgt das komplexe System der Regulation der Genexpression in Chloroplasten besser zu verstehen.

Summary

Synthesis and assembly of chloroplast proteins and protein complexes are mainly regulated by nucleus-encoded factors which act on various steps of gene expression including mRNA processing, splicing, stabilization, and translation initiation. The present study was aimed to analyze two such factors, RAP and DLA2, in the higher plant *Arabidopsis thaliana* and the green algae *Chlamydomonas reinhardtii*, respectively.

The protein RAP belongs to the superfamily of helical repeat proteins, which account for the majority of RNA-binding factors in plant organelles. To be more precise, RAP is the only identified OPR (octotricopeptide repeat) protein in *A. thaliana*, which has previously been described to be involved in plant pathogen defense. With the help of a T-DNA knockout-line a function of RAP in the correct and efficient processing of the 5' leader of the *16S* rRNA was uncovered. The level of mature *16S* rRNA is severely decreased, while precursors accumulate. This defect led in turn to a reduced synthesis and accumulation of chloroplast-encoded proteins. In summary, our data suggest an important role of the single OPR protein in Arabidopsis in a basic process in chloroplast biogenesis.

It was previously reported that the moonlighting enzyme DLA2 functions as a subunit of the chloroplast pyruvate dehydrogenase complex (PDC) as well as an RNA-binding protein involved in chloroplast gene expression. Its second function occurs only under mixotrophic growth conditions, i.e. in the presence of light and acetate. Here, it was shown that the cpPDC subunits E1 and E3 are not part of the DLA2-RNA complex. Further results, revealing no change in ratio between the subunits under different growth conditions, support the idea that under mixotrophic growth conditions the cpPDC is at least partially disassembled, thus releasing DLA2 for its second function. Remarkably, the *psbA* mRNA and the E3 protein bind to DLA in a competitive fashion. Moreover, analysis of the lysine acetylome revealed that the acetylation of a lysine residue within the proposed RNA-binding is upregulated under mixotrophic conditions. These findings support the hypothesis that the two functions of DLA2 are carried out in a competitive way regulated possibly by a post-translational modification.

In conclusion, the function of two nuclear-encoded factors involved in plastid protein synthesis was uncovered in more detail, helping to understand the complex system of regulation of gene expression in the chloroplast.

1 Introduction

1.1 Origin of chloroplasts

The uptake of a cyanobacteria-like prokaryote by a eukaryotic host cell about 1.5 billion years ago was the first step towards the development of the chloroplast as plant organelle. This event is called primary endosymbiosis. Prior to this, the same host cell had already taken up an α-proteobacterium-like prokaryote that led to the development of mitochondria. Today it is thought that the three main autotrophic lineages of glaucophytes, green algae and read algae emerged from one single endosymbiotic event. The photosynthetic cyanobacteria-like endosymbiont is therefore considered to be the common ancestor of all plastids (reviewed in Gould et al., 2008; McFadden and van Dooren, 2004).

Over time this endosymbiont evolved into an organelle. Its functions depend completely on the host cell, since it is not self-sufficient any longer. Most of the original bacterial genes were transferred to the nucleus of the host cell or even entirely lost, leaving the chloroplast genome of higher plants encoding only about 100 genes. The chloroplast genome of *Arabidopsis thaliana* (*A. thaliana*) for example encodes only 87 potential protein-encoding genes and 41 genes for structural RNAs (Sato et al., 1999). The situation in the green algae *Chlamydomonas reinhardtii* (*C. reinhardtii*) with 99 expressed sequences in total is very similar (Maul et al., 2002). The chloroplast genome encodes a wide variety of proteins involved in photosynthesis, fatty acid synthesis, gene expression and many other pathways. Chloroplasts contain their own gene expression machinery, even though most of the proteins in this chimeric system are encoded in the nuclear genome of the cell (reviewed in Bock and Timmis, 2008; Reyes-Prieto et al., 2007; Sato et al., 1999). More precisely, more than 90% of the proteins in the chloroplast are encoded in the nuclear genome, translated in the cytosol and thus need to be transported into the chloroplast (reviewed in Jarvis and Soll, 2001). In order to achieve this, transport mechanisms as well as the N-terminal transit peptide, a signal sequence for plastid targeting of proteins, evolved. The best characterized import system is a multi-protein translocon complex – the TOC (**t**ranslocon at the **o**uter membrane of **c**hloroplasts) and TIC (TIC—**t**ranslocon at the **i**nner membrane of **c**hloroplasts) complexes – that facilitates transport of proteins across both envelope membranes (reviewed in Li and Chiu, 2010; Schwenkert et al., 2011). Summing up, one can see that the uptake of the chloroplast and its

evolution into an organelle have brought drastic changes to the host cell, which has found numerous ways to deal with this sophisticated set-up as described in the following sections.

1.2 Functions of the chloroplast

Essential pathways like synthesis of fatty acids and amino acids, assimilation if nitrogen and sulfur and most characteristically photosynthesis take place in the chloroplast. In plants, fatty acid synthesis is mostly carried out within in the chloroplast and followed by their export to other cell compartments (reviewed in Bresinsky et al., 2008). The conversion of acetyl-CoA to malonyl-CoA by the enzyme acetyl-CoA carboxylase is the first committed step in fatty acid synthesis, which is followed by numerous sequential reactions carried out by the fatty acid synthetase, desaturases, elongases and other enzymes (Gunstone et al., 1994). In green leave tissue and microalgae the required acetyl-CoA is mainly provided by the chloroplast pyruvate dehydrogenase complex (reviewed in Harwood, 2009).

1.2.1 Photosynthesis

As mentioned above photosynthesis is the most commonly known process that takes place in chloroplasts. This process, converting sunlight into energy, evolved about 1 billion years before the primary endosymbiotic event took place, thereafter changing earth's atmosphere drastically by releasing oxygen and enabling the evolution of life as we know it today. During this reaction, light energy, water and carbon dioxide are used to turn inorganic substances into organic sugars (Hohmann-Marriott and Blankenship, 2011; Nelson, 2011).

Oxygenic photosynthesis can be divided into the so called light-dependent and light-independent reactions. During the light-dependent reaction a photon is channeled by the light-harvesting complexes towards a chlorophyll molecule, which is attached to the reaction center of photosystem II (PSII). The chlorophyll takes up this photon releasing an electron in return. This electron is travelling through the photosynthetic electron transport chain that consists of several thylakoid membrane embedded pigment protein complexes. In short, the released electron is transferred from PSII through the electron transport chain passing the cytochrome b_6f (Cyt b_6f) complex to a chlorophyll molecule embedded in PSI. From there, it finally reaches ferredoxin-$NADP^+$ reductase, an iron-sulfur protein, and $NADP^+$ is reduced to NAPDH (Nelson and Yocum, 2006; Xiong and Bauer, 2002). Figure 1 depicts this process as well as the connection to the light-independent reactions.

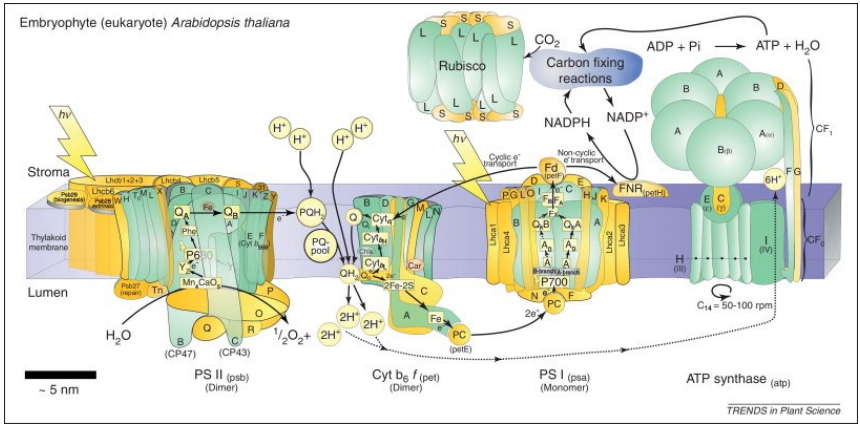


Figure 1: An overview of the photosynthetic reactions in *A. thaliana* (Allen et al., 2011). Major complexes of the light-dependent and light-independent reactions are depicted in this scheme. Plastid-encoded proteins are labeled in green, nuclear-encoded subunits in yellow. For further details, see text.

PSII consists of about 20 subunits, the number differs slightly between organisms, in which the proteins D1 and D2 always build the core of the reaction center. Antennae proteins like CP43 and CP47 or the proteins PsbO, PsbQ and PsbP, composing the oxygen evolving complex, are attached to it. Additionally, many accessory proteins and cofactors are involved in the described reactions (Dekker and Boekema, 2005). In order to recover the chlorophyll molecule in the reaction center of PSII, a water molecule is immediately split by a Mn_4CaO_5 cluster in the oxygen evolving complex in a process called photolysis (Suga et al., 2015). This process releases, as the name of the complex already indicates, oxygen as well as protons. These protons build, together with protons released during the transport of the electrons along the electron transport chain, a proton gradient across the thylakoid membrane. Subsequently, this gradient is used by the ATP synthase to convert ADP + P_i to ATP (Xiong and Bauer, 2002).

The light-independent reaction is named Calvin–Benson–Bassham cycle after the scientists, who discovered it in 1950. Briefly, an enzyme named ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) fixates carbon dioxide from the atmosphere to generate, with the aid of several additional enzymes, a three-carbon sugar product, namely glycerate-3-phosphate. The process can be divided into three steps: the carbon fixation, the reduction reactions, and the ribulose-1,5-bisphosphate regeneration. During this process NADPH and ATP generated in the light reaction are utilized (Bassham et al., 1950; Buchanan, 2016). In

summary, the absorbed light energy is converted to reduction equivalents and ATP used to drive the organism's numerous catabolic reactions.

1.3 The model organisms Arabidopsis thaliana and Chlamydomonas reinhardtii

The flowering plant Arabidopsis became the leading plant model organism over the last 30 years. In comparison to other angiosperms, it shows a short generation time of about 6-8 weeks and can – due to its small size of about 20 cm – be cultivated easily in restricted space such as in green houses and cultivation chambers. It's relatively small nuclear genome of about 157 Mb was completely sequenced in the year 2000 and annotated thoroughly since then. Most importantly a large number of T-DNA mutant lines is available, facilitating research with Arabidopsis. Such and other advantages have made Arabidopsis a widely used model organism for studies of cellular and molecular biology of higher plants (reviewed in Koornneef and Meinke, 2010).

On the other hand, *Chlamydomonas reinhardtii*, a small unicellular green algae with the ability to grow heterotrophically in acetate-containing media, has become an outstanding algal model organism (reviewed in Nickelsen and Kück, 2000). It is about 10 µm in size and contains a single cup-shaped chloroplast. Moreover, it reacts to light conditions of the environment with the help of a photoreceptive eye spot in the chloroplast and two flagella in a process called phototaxis (Harris, 1989). In 2007, the complete nuclear genome sequence of *C. reinhardtii* was published (Merchant et al., 2007). Furthermore, all three genomes (nucleus, plastid and mitochondrium) can be transformed efficiently. A recently created mutant insertion line library makes molecular analysis of many factors in Chlamydomonas even easier (Li et al., 2016). Taken together both here described and used organisms are well established models in many fields of plant research.

1.4 Chloroplast gene expression

1.4.1 Transcription

In higher plant chloroplasts two kinds of RNA polymerases, the plastid-encoded RNA polymerase (PEP) and the nucleus-encoded RNA polymerase (NEP), carry out the first step of gene expression: the transcription (Hess et al., 1993).

PEP is a multi-enzyme complex consisting in general of 4 subunits (rpoA, rpoB, rpoC1 and rpoC2), that derived directly from its cyanobacterial ancestor. Its activity in higher plants is regulated by at least six sigma-like transcription factors (SFLs), whereas in genomes of green algae only one SFL, namely RPOD, was identified (Allison, 2000; Carter et al., 2004). Some of the SFLs in *A. thaliana* are working in general mode, whereas others are at least partially gene specific. Their expression depends on light, circadian rhythm, tissue and developmental stage (Allison, 2000; Favory et al., 2005; Zghidi et al., 2007). This interesting increase of factors in land plants compared to algae is explained by a degeneration and diversification of promoters that needs to be compensated for (Allison, 2000; Maier et al., 2008). Moreover, additional proteins are involved in transcription of certain genes (Jalal et al., 2015; Pfannschmidt et al., 2009; Sutoh et al., 1999).

The second type of polymerase in the chloroplast of higher plants, the NEP, is a single subunit enzyme homologous to RNA polymerases of T3 and T7 phages. Up to now two nucleus-encoded polymerases targeted to the chloroplast were identified in *A. thaliana*. RpoTp is exclusively localized to the chloroplast while RpoTmp can also be found in mitochondria (Liere et al., 2011). This type of RNA polymerase could not be identified in some green algae including *C. reinhardtii*, *Osteococcus tauria* as well as *Thalassiosira pseudomona*, although genome-wide searches were conducted (Armbrust et al., 2004; Derelle et al., 2006; Smith and Purton, 2002). Interestingly, many transcripts in Arabidopsis possess promoters for both polymerases. It was proposed that PEP dominates the transcription of photosynthetic genes (Börner et al., 2015).

1.4.2 RNA Processing

Many genes in the chloroplast, transcribed by the above mentioned RNA polymerases PEP and NEP, are organized in gene clusters. This results in polycistronic transcripts, that subsequently are further processed into monocistronic transcripts. It is postulated, that translation of monocistronic forms is more effective than that of polycistronic transcripts (Barkan et al., 1994; Hirose and Sugiura, 1997).

The RNA maturation processes include endonucleolytic cleavage, maturation of 5' and 3' ends and splicing, but also other post-transcriptional modifications like editing and RNA stability as depicted in figure 2 (reviewed in Barnes and Mayfield, 2003; Bollenbach et al., 2004; Gray et al., 2003; Stern et al., 2010). So far RNA editing has only been found in vascular plants; with an average of 35 editing sites per chloroplast genome (Shikanai, 2006;

Tillich et al., 2006). In this process specific cytidines are replaced by uridines. Interestingly, most editing sites are found in coding mRNA regions, leading to an alteration of the amino acid sequence and often affecting amino acids that play a role in protein function. Evidence indicates that editing is a mechanism to counterbalance deleterious effect of point mutations (reviewed in del Campo, 2009).

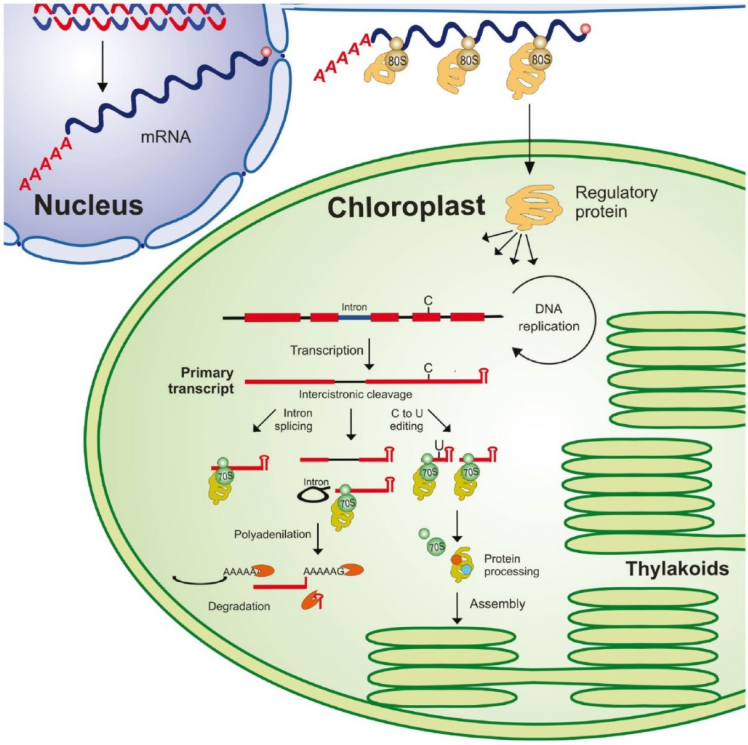


Figure 2: Overview of chloroplast post-transcriptional mechanisms involved in the control of chloroplast gene expression (del Campo, 2009). Several regulatory proteins, encoded in the nucleus, are involved in post-transcriptional processes like intron splicing, RNA editing, nucleolytic activity, 5'- and 3'-end maturation, stabilization and degradation. For further information, see text.

Introns can be found in about 17% of angiosperm chloroplast genes, most of which resemble *cis*-acting group II introns (reviewed in Herrin and Nickelsen, 2004; Plant and Gray, 1988). Although these introns are auto-spliced *in vitro*, data from several experiments suggest that other factors are required for sufficient splicing *in vivo* (Stern et al., 2010). Several nuclear-encoded proteins involved in splicing have been identified in higher plants, among

them OTP70, an Arabidopsis pentatricopeptide repeat protein, that is involved in splicing of the *rpoC1* intron (Chateigner-Boutin et al., 2011). It belongs to a family of helical repeat-proteins that will be discussed later (see section 1.5).

Processing of the earlier mentioned polycistronic transcripts requires endo- as well as exonucleolytic activities. In *Arabidopsis thaliana* only 17 of more than 180 annotated ribonucleases are predicted to be localized in the chloroplast (Stoppel and Meurer, 2011). Several of these have been characterized, e.g. the PPR protein CRR2 is involved in *rps7*/*ndhB* RNA processing (Hashimoto et al., 2003). Another well-known example is the PNPase (polynucleotide phosphorylase), which is involved in 3' to 5' processing/degradation of RNA, in eubacteria as well as in chloroplasts. Additionally, it plays a role in polyadenylating 3'-ends of plastid transcripts, which promotes a rapid RNA turnover by exonucleolytic factors like the PNPase itself (Walter et al., 2002; Yehudai-Resheff et al., 2001). Its function as an exonuclease is required in processing of the *rrn* (ribosomal RNA) operon, to mention and elucidate just one example.

The transcripts of ribosomal RNAs are organized in one of the biggest operons of the Arabidopsis chloroplast genome, which includes genes encoding the four ribosomal RNAs present in chloroplasts, *rrn16*, *rrn23*, *rrn4.5* and *rrn5*, as well as three tRNA genes, *trnI*, *trnA* and *trnR* (Fig. 3). Nomenclature of *rrn* promoters is used according to Lerbs-Mache (2000). Chloroplast ribosomes consist of more than 50 ribosomal proteins and the just mention ribosomal RNAs, *16S*, *23S*, *4.5S* and *5S*. The resulting 70S ribosomes are composed of 30S and 50S subunits, similar to the ribosomes of their prokaryotic ancestors (reviewed in Harris et al., 1994). The presence of ribosomal RNAs in the ribosomes is required for translational activity as will be discussed in more detail in section 1.4.3 (Nissen et al., 2000).

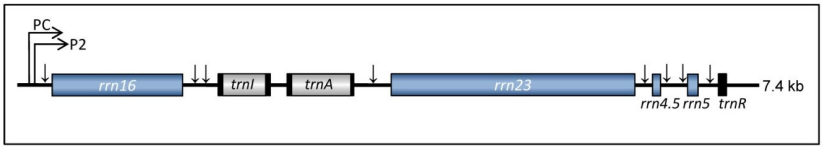


Figure 3: Schematic representation of the chloroplast *rrn* operon in *A. thaliana*. Positions of a NEP-transcribed promoter (PC) and a PEP-transcribed promoter (P2) are indicated by curved arrows. Positions of processing sites are indicated by vertical arrows. Introns in tRNA genes are illustrated in grey, exons in black. Genes encoding ribosomal RNAs are illustrated in blue.

After transcription of the operon, either starting from the PEP- or the NEP-transcribed promoter, named PC and P2 respectively, the 7.4 kb rRNA polycistronic transcript is cleaved by not yet identified endonucleases, resulting in pre-tRNAs for isoleucine and alanine, *16S*

rRNAs precursors, as well as dicistronic *23S-4.5S* and *5S*-tRNA^{ARG} transcripts. Analyses of several mutants lacking nucleus-encoded factors revealed their role in following processing and maturation steps. Processing of the monocistronic and dicistronic transcripts as well as the maturation of 5' and 3' ends of ribosomal RNAs seems to be dependent on the presence of the factors RNR1, DCL, CSP41a/b, DAL, WCO, BPG2, PAC and the already mentioned PNPase (Beligni and Mayfield, 2008; Bellaoui et al., 2003; Bisanz et al., 2003; Bollenbach et al., 2005; Komatsu et al., 2010; Meurer et al., 2017; Walter et al., 2002; Yamamoto et al., 2000). To gain a profound understanding of the complex processing of the ribosomal RNA operon other factors involved must be identified and the function of all involved proteins must be revealed.

Transcript stability does not only depend on the just mentioned ribonucleases, but mostly on the presence of 5' untranslated regions (UTRs) and 3' UTRs. Deletion of either of these, leads to reduced transcript accumulation and translation (Stern et al., 2010). Many transcripts contain short inverted repeats close to their 3'-end that can form stable stem-loops, thus preventing 3' to 5' exonucleolytic decay. Experimental evidence shows that transcript stabilization is also ensured by specific protection through protein binding to 5'- and 3'-ends of transcripts (Hammani et al., 2012; Pfalz et al., 2009; Prikryl et al., 2011; Zhelyazkova et al., 2011).

1.4.3 Translation

Protein translation is the next major step for control of gene expression in the chloroplast, which is mainly regulated through initiation of translation although elongation steps can also be adjusted. Chloroplast ribosomes arose during the evolution of the former free living endosymbiont to an organelle and are clearly different from their eukaryotic counterparts. They share certain characteristics with ribosomes of their bacterial ancestor, which are obvious seen when comparing the structure of the 70S ribosomes. Nevertheless, they acquired additional subunits and domains that might have a function in organelle-specific processes (reviewed in Marin-Navarro et al., 2007).

In contrast to translation in prokaryotes, the ribosome binding site, called Shine-Dalgarno (SD) sequence, is more dispensable in plastidial translation. It is proposed that its lack can be compensated by sequence specific factors guiding the ribosomes to their site of action (Hirose and Sugiura, 1996). In *C. reinhardtii* e.g. the SD sequence is required for synthesis of D1, while its removal leads only to slight changes in translation of D2 (Mayfield

et al., 1994; Nickelsen et al., 1999). Remarkably, a similar sequence is even a negative regulator element for translation of the tobacco *rps2* transcript (Plader and Sugiura, 2003).

Moreover, the localization of mRNAs within the chloroplast can play a role in their translation. In *C. reinhardtii* a specific zone of translation (T-zone) was identified. It is a site of protein synthesis for *de novo* biogenesis of PSII and possibly other complexes and was defined by co-localization of the chloroplast ribosome, mRNAs encoding PSII subunits as well as the RBP40 translation marker (Schottkowski et al., 2012; Uniacke and Zerges, 2009). More recently, Bohne et al. (2013) reported that a nucleus-encoded factor, DLA2, tethers the *psbA* mRNA to the T-zone only under certain growth conditions. Future research will show which other factors are involved in spatial organization within the chloroplast.

Several nuclear-encoded factors influencing chloroplast translation were identified. In higher plants for instance, CRP1 from maize, and HCF107 and HCF173 from Arabidopsis are required for translation of *petA/petD*, *psbB* and *psbA* mRNAs, respectively (Sane et al., 2005; Schmitz-Linneweber et al., 2005; Schult et al., 2007). The most prominent example in *C. reinhardtii* is the regulation of the D1 protein synthesis, in which at least one multi-subunit complex is involved. This complex consists of the RB60 protein, a protein disulfide isomerase, the RB47 protein, a poly(A)-binding protein with an intrinsic RNA binding activity, the RB38 protein, a poly(U)-binding protein that showed interaction with the target *in vitro* and last the RB55 protein, whose molecular characterization remains elusive so far (Barnes et al., 2004; Danon and Mayfield, 1991; Kim and Mayfield, 1997; Yohn et al., 1998). An independent approach identified the already mentioned RNA-binding protein DLA2, binding specifically upstream of the *psbA* start codon, whose function will be explained in more detail in section 1.6.1 (Ossenbühl et al., 2002). An increased level of D1 synthesis in light, even though a high level of *psbA* transcripts is already present in the dark, could be explained by regulation and an increased binding activity of the described factors (Danon and Mayfield, 1991; Malnöe et al., 1988). There is evidence that RB60 is controlled by phosphorylation, while RB47 is inactivated in the dark through oxidation (Danon and Mayfield, 1994a, b). The exact working mode of this complex and its regulation is controversially discussed and needs to be elucidated in the future.

In the green algae *C. reinhardtii* another remarkable level of translational regulation, the principle of control by epistasy of synthesis (CES) controlling the synthesis and assembly of photosynthetic complexes, was described. Complex assembly starts with the insertion of a

dominant subunit, which acts as a scaffold for subsequent assembly steps. For instance, D2 is the anchor protein of PSII, PetB of the $Cytb_6f$ complex and PsaB of PSI. In the absence or reduction of these dominant proteins, translation of the next protein to be inserted into the complex is inhibited, which inhibits translation of the next protein and so on, thus leading to a cascade of control of protein synthesis (Choquet and Vallon, 2000). Control of translation plays an important role in regulation of gene expression.

It is clearly evident, that the RNA metabolism of chloroplasts has gained additional complexity compared to that of their cyanobacterial progenitors. In plastid gene expression, a predominance of post-transcriptional mechanisms exists. Part of this new complexity might be due to the reason that chloroplasts propagate asexually and are inherited uniparental. Without a meiotic event no recombination takes place; deleterious mutations have to be neutralized by other ways. Presumably for this reason as well as for purposes of regulation and coordination of nucleus-encoded and plastid-encoded subunits of protein complexes, a high number of nuclear factors involved in chloroplast gene expression evolved (Maier et al., 2008; Tillich et al., 2010).

1.5 Helical repeat proteins involved in regulation of chloroplast gene expression

In both vascular plants and algae the biogenesis and activity of chloroplasts is controlled by an intracellular network of nucleus encoded factors. Several families of helical repeat proteins are involved in the processes described in earlier sections. The most prominent member of this superfamily is the family of pentatricopeptide repeat (PPR) proteins, consisting of RNA-binding proteins that contain up to 30 tandem repeats of loosely conserved 35 amino acids. They were predicted to build an array of α-helices and therefore belong to the α-solenoid superfamily together with tetratricopeptide repeat (TPR) proteins (Small and Peeters, 2000). This structure was confirmed for the PPR10 protein from maize by crystallization and X-ray scattering (Gully et al., 2015; Yin et al., 2013).

The first described PPR protein was reported in *Saccharomyces cerevisiae* (Manthey and McEwen, 1995). However, members of this family can be found in all eukaryotes, but not in prokaryotic organisms. Interestingly, there is a striking difference in numbers between higher plants and algae. While in *A. thaliana* about 450 PPR proteins are predicted, the green alga *C. reinhardtii* possesses only 11 PPR proteins (Jalal et al., 2015; Schmitz-Linneweber

and Small, 2008). Domain swap experiments with CRR21 and CRR4 as well as crystal structures of PPR10 binding one of its RNA targets proved the RNA binding function of the PPR domain (Gully et al., 2015; Okuda et al., 2007; Yin et al., 2013). In *A. thaliana* many chloroplast localized PPR proteins are well studied and show a function in chloroplast gene expression.

The second well-known family of helical repeat proteins is the tetratricopeptide repeat (TPR) protein family. As the name indicates, the number of amino acids in the 3-16 tandem repeats is 34 instead of 35 as in PPR proteins (Blatch and Lassle, 1999). The structure of several TPR proteins, arrays of anti-parallel α-helices that generate an α-solenoid helical structure, was revealed by crystallization (D'Andrea and Regan, 2003). Similar to PPR proteins, TPR domains are highly conserved among prokaryotic as well eukaryotic organisms. In contrast to RNA-binding PPR proteins, TPR proteins participate in various processes via protein-protein-interaction. In *A. thaliana* chloroplast's LPA1 is a well characterized example; it interacts with the photosystem II reaction center protein D1 and is required for photosystem II assembly (Peng et al., 2006). A prominent example in the chloroplast of *C. reinhardtii* is the protein Nac2, whose binding is absolutely necessary for transcript accumulation of *psbD* (Boudreau et al., 2000). Subsequent interaction with the RBP40 protein initiates translation of D2 (Ossenbühl and Nickelsen, 2000).

The last family of helical repeat proteins introduced here is the more recently described family of octotricopeptide repeat (OPR) proteins. This protein family is named after its 38 - 40 amino acid degenerate motif. Secondary structure predictions have foretold paired α-helices that generate an α-solenoid helical structure similar to the proposed PPR and TPR protein structure (Eberhard et al., 2011). Interestingly, there are more than 100 OPR proteins predicted in *C. reinhardtii*, compared to only a single one in higher plants (O. Vallon, A. Bohne, L. Cerutti, J. D. Rochaix, unpublished data).

The characterization of a number of *C. reinhardtii* OPR proteins sheds light on their function. The MCD1 protein is involved in *petD* RNA stability, which encodes a subunit of the Cyt b_6f complex (Murakami et al., 2005). The factor RAA1 is needed for *psaA* RNA processing, which is a PSI reaction center protein (Merendino et al., 2006; Perron et al., 2004). TBC2 and TDA1 are both translation factors involved in *psbC*, which encodes the CP43 antennae protein of PSII, and *atpA*, which codes for the α-subunit of the ATPase, translation, respectively (Auchincloss et al., 2002; Eberhard et al., 2011). MGC1 and MBI1

are two more recently identified mRNA stabilization factors, which bind to *petG*, encoding a small subunit of the Cyt b_6f complex and *psbI*, encoding a small PSII subunit, respectively (Wang et al., 2015). All so far characterized OPR proteins play a role in chloroplast gene expression, and are supposedly RNA-binding proteins, which leads to the assumption that this family plays a similar role as helical repeat proteins mentioned earlier.

1.6 Moonlighting proteins

Unlike helical repeat proteins, which provide supposedly the majority of RNA binding proteins in the organelles of plants and algae, the so called moonlighting enzymes do not all belong to one family. Rather this group of proteins is linked by their acquisition of an additional function next to their first described function, which is often enzymatic (Huberts and van der Klei, 2010). Moonlighting proteins need to be distinguished from the use of a single gene to generate different proteins by post-translational processing, alternative splicing and DNA rearrangement as well as from multifunctional proteins, which harbor multiple domains each serving a different function (reviewed in Jeffery, 2018). It was postulated that moonlighting proteins started out as uni-functional proteins, whose secondary function evolved with time (Jeffery, 2003).

In 1988, the first such proteins, crystallins in the lenses of eyes, were described (Piatigorsky et al., 1988). In ducks, the protein delta-crystallin plays a structural role in transparency and cataract in the lens, but additionally shows an argininosuccinate lyase activity. Interestingly, later it was found that several lens crystallins are identical to metabolic enzymes and stress proteins found in numerous tissues (Piatigorsky, 1998; Wistow and Piatigorsky, 1988). Soon after the first description of this kind of gene sharing, many more proteins with two independent functions were discovered and the term moonlighting proteins was introduced, referring to workers working in two jobs whereas the second additional job would often happen at night thereby in moonlight (Huberts and van der Klei, 2010; Jeffery, 1999).

There are several ways how moonlighting proteins can evolve, but it is in general greatly promoted by the fact that there is unused space in many proteins since the active site typically only takes up a small part of the protein. Alterations in this unused space can lead to the acquisition of an additional function. Interestingly, this is a mechanism that can sometimes explain the resistance to certain antibiotic in bacteria (Sengupta et al., 2008). Moreover, many

moonlighting proteins originate from the fusion of two single function genes resulting in one protein harboring two functions (Gancedo and Flores, 2008). Even other moonlighting proteins feature only one active site that carries out the protein's two functions. In these cases, mutations in the active site led to the evolvement of the second function. It is postulated that these bifunctional proteins are evolutionary favorable due to the conservation of amino acids and energy otherwise required to build multiple proteins (Jeffery, 1999).

1.6.1 The moonlighting enzyme DLA2

One example of a moonlighting protein discovered by chance in *C. reinhardtii* is the RNA-binding protein RBP63 described by Ossenbühl et al. (2002). A screen for RNA-binding proteins associated with the thylakoid membrane of the chloroplast was performed and this protein with a size of 63 kDa was detected. Further experiments revealed a preference for the 5' UTR of the *psbA* mRNA compared to other chloroplast mRNAs. The *PsbA* gene codes for the D1 subunit in the reaction center of PSII as described earlier (section 1.2.1). The *psbA* mRNA binding activity of RBP63 was not only confirmed with additional methods by Bohne et al. (2013), but remarkably it became evident that RBP63 is actually the previously described protein DLA2 (dihydrolipoamid acetyltranserase), the E2 subunit of the chloroplast pyruvate dehydrogenase complex (cpPDC) (Mooney et al., 1999; Reid et al., 1977).

The multi-enzyme pyruvate dehydrogenase complex is ubiquitously present among organisms converting pyruvate to acetyl-CoA in a process called pyruvate decarboxylation, thereby connecting the citric acid cycle and subsequent oxidative phosphorylation to the glycolysis, gluconeogenesis and lipid as well as amino acid metabolism pathways. Plant cells possess two copies of this complex, one located in mitochondria and in in chloroplasts (Reid et al., 1977). It consists in general of many copies of its three subunits that conjointly execute the complex's function (Fig 4). The E1 subunit, the pyruvate dehydrogenase, transfers an acetyl group of the pyruvate onto a lipoamide attached to the E2 subunit aided by its cofactor a thiamine pyrophosphate. In turn the E2 subunit transfers the acetyl by a transacetylation reaction from the swinging arm of lipoyl to coenzyme A generating acetyl-CoA, which is released from the enzyme complex at this point. Lastly, the dihydrolipoate, still bound to a lysine residue of the E2, migrates to the dihydrolipoyl dehydrogenase (E3) active site where it undergoes a flavin-mediated oxidation. Hereby, dihydrolipoate is oxidized by FAD back to its lipoate resting state, producing $FADH_2$. Subsequently, a NAD^+ cofactor oxidizes $FADH_2$ back to FAD producing NADH (Camp and Randall, 1985; Mooney et al., 2002; Reid et al.,

1977). It should be noted that the E2 subunit of the cpPDC complex in chloroplasts of *C. reinhardtii* DLA2 contains a single predicted lipoamide attachment site, similar to several other photosynthetic organisms as the cyanobacterium *Synechocystis sp.* PCC 6803 protein or the chloroplast E2 subunit from *A. thaliana* as opposed to up to three sites observed in *E.coli* and other organisms (Bohne et al., 2013; Camp and Randall, 1985; Guest et al., 1985; Tovar-Méndez et al., 2003).

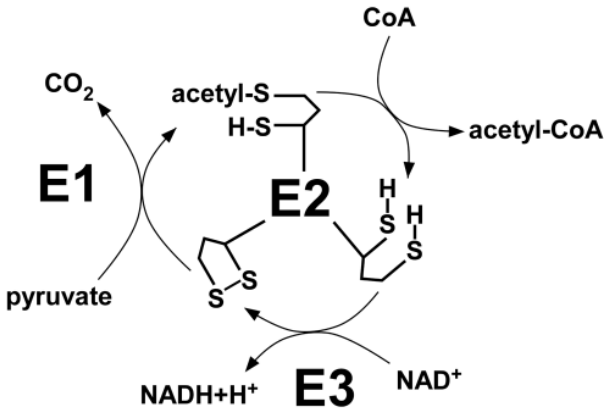


Figure 4: Simplified working mode of the pyruvate dehydrogenase complex (Foth et al., 2005). The conversion of pyruvate to acetyl-CoA and the subsequent recovery of the subunits are shown. In this case the E2 subunit has three lipoamids attached to it, typical for bacterial PDCs. For further explanations, see text.

Remarkably, DLA2 plays an additional role in the chloroplast of *C. reinhardtii* as mentioned above. When the cells are grown under mixotrophic conditions, i.e. in the presence of light and acetate, DLA2 is part of an RNA containing high molecular weight (HMW) complex with a size of more than 1 MDa. Interestingly the DLA2-RNA complex was not detected in a mutant lacking the *psbA* mRNA. Furthermore, it was suggested that by binding to the *psbA* mRNA, DLA2 tethers this mRNA to the T-zone in the chloroplast (Bohne et al., 2013). By localizing the *psbA* mRNA in this zone DLA2 can thereby promote the translation of the D1 protein at the thylakoid membrane (Bohne and Nickelsen, 2017; Bohne et al., 2013).

A proposed regulation of *psbA* gene expression by the presence of acetate in the medium is illustrated in figure 5. In the absence of acetate, i.e. under photoautotrophic conditions, DLA2 is required as a cpPDC subunit to produce acetyl-CoA in the chloroplast. A

normal level of D1 translation is taking place. In contrast, under mixotrophic growth conditions acetate is converted into acetyl-CoA by the acetate synthetase (ACS) and/or by the acetyl-kinase/ phosphate acetyltransferase (ACK/PAT) system of *C. reinhardtii* leading to accumulating acetyl-CoA levels (Bohne et al., 2013; Spalding, 2009). This in turn signals substrate availability for fatty acid synthesis in the chloroplast and can lead to a product inhibition of the cpPDC and putatively the partial disassembly of the cpPDC (Tovar-Méndez et al., 2003; Wellen and Thompson, 2012; Xing and Poirier, 2012).

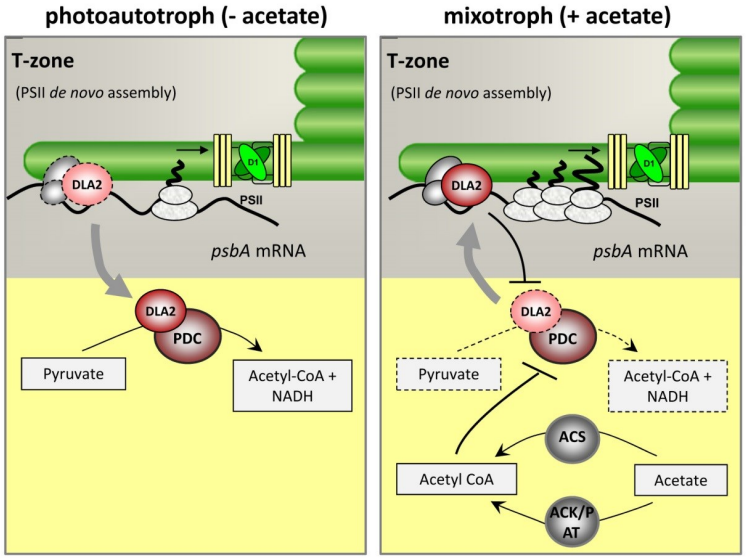


Figure 5: Acetate-dependent regulation of psbA gene expression in *C. reinhardtii* by DLA2. Under photoautotrophic conditions DLA2 is required as a cpPDC subunit to produce acetyl-CoA (left panel). In contrast, under mixotrophic growth conditions (right panel), acetate is converted into acetyl-CoA by ACS and/or by the ACK/PAT system of *C. reinhardtii*, leading to binding of DLA2 to the psbA mRNA (adapted from Bohne et al., 2013). For further explanation, see text.

Free DLA2 subunit might bind increasingly to the *psbA* mRNA and thus lead to its localization to the T-zone favoring D1 synthesis for *de novo* assembly of PSII. Reciprocally, the cpPDC activity is decreased by *psbA* mRNA binding to DLA2 as shown by Bohne et al. (2013). In summary, DLA2 is a multifunctional moonlighting protein connecting carbon metabolism and gene expression in the chloroplast of *C. reinhardtii*.

2 Aims of this work

Plants need to react to changes in environmental conditions, which they often they do by adapting their gene expression. In the chloroplast, most regulation of gene expression happens on post-transcriptional level. The main aim of this work was to further elucidate the role of two such factors, DLA2 and RAP, two nucleus-encoded proteins localized in the chloroplasts from *C. reinhardtii* and *A. thaliana*, respectively.

RAP, the single OPR protein in Arabidopsis was reported as a factor in plant defense so far, without any molecular characterization of its working mode. The fact that several OPR proteins from the green algae *C. reinhardtii* play a role in gene expression, made it even more desirable to elucidate the function of RAP. Therefore, we analyzed the protein with the help of a T-DNA insertion line in detail. Phenotypical and molecular characterization helped to uncover its function (section 3.1).

To further understand the regulatory function of DLA2 as a connector of gene expression and carbon metabolism, further experiments were required. The mechanism behind the regulation of DLA2's two function was still unknown. We concentrated our research on the composition of the RNA-DLA2 complex and the elucidation of the RNA-binding region of DLA2 (section 3.2).

Overall, the data in this thesis aims to shed further light on the regulation of protein synthesis by nucleus-encoded factors in chloroplasts of higher plants and green algae.

3 Results

The following section includes two studies (section 3.1 and 3.3) that are published in international peer-reviewed journals and one manuscript that is still in preparation (section 3.2). At the beginning of each chapter the main conclusions of the respective study as well as the contributions of the author are summarized.

3.1 RAP, the sole octotricopeptide repeat protein in Arabidopsis, is required for chloroplast 16S rRNA maturation

Kleinknecht, L., Wang, F., Stübe, R., Philippar, K., Nickelsen, J., and Bohne, A.-V. (2014). Plant Cell 26: 777-787

The focus of this study was the characterization of RAP, the sole OPR protein in *A. thaliana*, which has previously been described to be involved in plant pathogen defense. The analyzed knockout mutant line *rap-1* showed retarded growth and a defect in photosynthetic activity based on measurements of the PS II efficiency. Also, immunoblot analysis and *in vivo* labeling of chloroplast proteins pointed to an impaired chloroplast translation. Furthermore, Northern blots were able to reveal a reduction of most analyzed chloroplast transcripts and a dramatic reduction of the plastid encoded *16S* rRNA. Detailed analysis of the plastid rRNA operon showed no changes in levels or pattern of *23S*, *4.5S* and *5S* rRNA. In stark contrast *16S* rRNA precursors accumulated, while the mature transcript was lacking entirely. Importantly, *in vitro* RNA binding experiments revealed that RAP has an intrinsic RNA binding capacity. A number of experiments point to the fact that RAP binds to the 5' leader sequence of the *16S* precursor rRNA. Northern blots of small RNAs were performed to detect previously described footprints - small RNAs that accumulate often as the result of a protein binding to a stretch of RNA and thereby protecting the fragment. Interestingly, a footprint in the 5' end of the *16S* rRNA precursor detected in wild type and the *rfb1-1* control, which shows general defects in rRNA accumulation, was not detectable in the *rap-1* mutant.

In fact, when RNA binding curves were plotted from filter binding assays, RAP showed a higher affinity to a probe spanning this footprint compared to two other reported footprint sequences in the same region. Since the affinity was only slightly higher, we cannot exclude that RAP binds to another sequence within the *16S* rRNA precursor additionally. Nevertheless primer extension analyses and circular RT-PCR revealed a severe defect in correct processing of the *16S* precursor rRNA, supporting the role of RAP in correct trimming of the mature *16S* rRNA. Additionally, GFP import assays showed a nucleoid localization of RAP, marking the nucleoid as the site of chloroplast rRNA processing. In summary, our data suggest an important role of the single OPR protein in Arabidopsis in a basic process in chloroplast biogenesis.

I contributed to this study by conducting all the experiments except for the initial western blots, *in vivo* labeling and the GFP import studies, which were performed by F. Wang and R. Stübe, respectively. Bioinformatic protein characterization was conducted by A.-V. Bohne. The manuscript was written by A.-V. Bohne, J. Nickelsen, K. Philippar and me and revised by all co-authors.

In 2016, a correction to the original paper was published in Plant Cell. It is included in the presented manuscript and figure denominations in the following sections are according to the final paper i.e. after the correction was published.

The Plant Cell, Vol. 26: 777–787, February 2014, www.plantcell.org

RAP, the Sole Octotricopeptide Repeat Protein in *Arabidopsis*, Is Required for Chloroplast *16S* rRNA Maturation[W]

Laura Kleinknecht,[a] Fei Wang,[a,1] Roland Stübe,[b] Katrin Philippar,[b] Jörg Nickelsen,[a,2] and Alexandra-Viola Bohne[a]

[a] Molecular Plant Sciences, Ludwig-Maximilians-University, 82152 Planegg-Martinsried, Germany
[b] Plant Biochemistry and Physiology, Ludwig-Maximillians-University, 82152 Planegg-Martinsried, Germany

The biogenesis and activity of chloroplasts in both vascular plants and algae depends on an intracellular network of nucleus-encoded, *trans*-acting factors that control almost all aspects of organellar gene expression. Most of these regulatory factors belong to the helical repeat protein superfamily, which includes tetratricopeptide repeat, pentatricopeptide repeat, and the recently identified octotricopeptide repeat (OPR) proteins. Whereas green algae express many different OPR proteins, only a single orthologous OPR protein is encoded in the genomes of most land plants. Here, we report the characterization of the only OPR protein in *Arabidopsis thaliana*, RAP, which has previously been implicated in plant pathogen defense. Loss of RAP led to a severe defect in processing of chloroplast *16S* rRNA resulting in impaired chloroplast translation and photosynthesis. In vitro RNA binding and RNase protection assays revealed that RAP has an intrinsic and specific RNA binding capacity, and the RAP binding site was mapped to the 5′ region of the *16S* rRNA precursor. Nucleoid localization of RAP was shown by transient green fluorescent protein import assays, implicating the nucleoid as the site of chloroplast rRNA processing. Taken together, our data indicate that the single OPR protein in *Arabidopsis* is important for a basic process of chloroplast biogenesis.

INTRODUCTION

Chloroplasts, the photosynthetic organelles of plants and algae, derive from the integration of a photosynthetic cyanobacterium-like prokaryote into a eukaryotic host cell (Timmis et al., 2004). During evolution, the endosymbiotic organism was converted into an organelle that still possesses a reduced genome and its own gene expression machinery. The process was accompanied by the development of a set of nucleus-encoded, *trans*-acting factors that must be imported into the chloroplast. These factors form part of an intracellular network that coordinates organellar and nuclear gene expression, mainly at the posttranscriptional level (Del Campo, 2009; Barkan, 2011). Among the processes affected are the maturation of chloroplast RNAs, such as intergenic cleavage of polycistronic transcripts, RNA editing, and the generation of mature 5′ and 3′ RNA ends, as well as their regulated translation on eubacterial-like *70S* ribosomes (Bollenbach et al., 2007; Barkan, 2011). Chloroplast ribosomes are composed of more than 50 proteins and four rRNAs (*16S*, *23S*, *4.5S*, and *5S*), which are encoded in a cotranscribed gene cluster and undergo complex maturation processes in a ribosome assembly-assisted manner (Shajani et al., 2011; Stoppel and Meurer, 2012).

The identification of numerous components of the intracellular communication network between nucleus and chloroplasts has revealed that the overwhelming majority belong to the helical repeat protein superfamily, including tetratricopeptide repeat (TPR) and pentatricopeptide repeat (PPR) proteins (Stern et al., 2010; Shikanai and Fujii, 2013). TPR and PPR domains are repetitive units formed by two antiparallel α-helices with characteristic consensus motifs and have been reported to mediate protein–protein or RNA–protein interactions, respectively (Das et al., 1998; Schmitz-Linneweber and Small, 2008; Ringel et al., 2011). Whereas genomes of higher plants, like *Arabidopsis thaliana*, encode more than 450 members of the PPR protein family, the unicellular green alga *Chlamydomonas reinhardtii* possesses only 12 *PPR* genes (Schmitz-Linneweber and Small, 2008).

However, recently, a novel class of helical repeat proteins, named octotricopeptide repeat (OPR) proteins, has been described in *C. reinhardtii*, and its members are characterized by tandemly repeated, degenerate 38–amino acid units (Eberhard et al., 2011; Rahire et al., 2012). Based on secondary structure predictions and in vitro RNA binding experiments, these repeats, like PPR repeats, are postulated to form α-helical RNA binding domains (Eberhard et al., 2011; Rahire et al., 2012). This is further supported by the functions of characterized OPRs, all of which act as RNA stabilization/processing and translation factors (Auchincloss et al., 2002; Perron et al., 2004; Murakami et al., 2005; Merendino et al., 2006; Eberhard et al., 2011; Rahire et al., 2012). Interestingly, the OPR gene family seems to have undergone a marked expansion in green algae, with dozens of members in *C. reinhardtii* (Eberhard et al., 2011; Rahire et al., 2012). However, in stark contrast to the large numbers of PPR proteins, most land plant genomes contain a single *OPR* gene, including those of representative model organisms such as *Arabidopsis*, tobacco (*Nicotiana tabacum*), rice (*Oryza sativa*), maize (*Zea mays*), and *Physcomitrella patens* (Olivier Vallon, personal communication).

Here, we report the functional characterization of the sole OPR protein found in *Arabidopsis*, RAP, which has previously been

[1] Current address: Institut de Biologie Physico-Chimique, CNRS, 13 Rue Pierre et Marie Curie, 75005 Paris, France.
[2] Address correspondence to joerg.nickelsen@lrz.uni-muenchen.de.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Jörg Nickelsen (joerg.nickelsen@lrz.uni-muenchen.de).
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.114.122853

implicated in the negative regulation of plant disease resistance (Katiyar-Agarwal et al., 2007). Our data now reveal a role of RAP in a fundamental process of chloroplast gene expression (i.e., rRNA maturation).

RESULTS

Arabidopsis Has Only a Single OPR Protein

A few putative OPR proteins have been reported in *Arabidopsis* (Eberhard et al., 2011). However, reevaluation of available genomic data has revealed only a single OPR protein, RAP (Olivier Vallon, personal communication). *Arabidopsis* RAP exhibits a putative plastid transit peptide of 78 amino acids (Figure 1A; Supplemental Figure 1A). The mature protein has a molecular mass of 67 kD. Its C-terminal half comprises four OPR repeats followed by a RAP (RNA binding domain abundant in apicomplexans) domain (Figure 1; Supplemental Figure 1A; Lee and Hong, 2004), which is probably related to OPR repeats (Eberhard et al., 2011). Secondary structure analysis with the Jpred algorithm (www.compbio.dundee.ac.uk/www.jpred; Cole et al. 2008) predicted the presence of two α-helices in each of the OPR repeats identified (Figure 1B), as in the case of PPR and TPR repeats (Das et al., 1998; Ban et al., 2013). This α-helical structure of the OPR repeats is further supported by the prediction of the 3D structure of the region representing OPR repeats 1 to 3 (Figure 1C).

Interestingly, similarity searches revealed also only a single orthologous *OPR* gene in representative land plant genomes investigated, including the moss *P. patens* (Supplemental Figure 1A). The analysis of these OPR proteins showed clear conservation at the C terminus, including the OPR repeats and the RAP domain, indicating a monophyletic origin, whereas the N-terminal region is more variable (Supplemental Figure 1A). Like RAP, all analyzed orthologs are predicted to possess an organellar targeting signal (Supplemental Figure 1B).

Loss of RAP Impairs Translation in Chloroplasts

To characterize the function of RAP, we analyzed the *Arabidopsis* mutant line *rap-1*, which carries a T-DNA insertion in the third exon of the *RAP* gene (Figure 2A). Homozygous mutants were obtained from the T3 generation (Supplemental Figure 2B). The *Arabidopsis rap-1* mutant (previously called *atrap-1*) was reported by Katiyar-Agarwal et al. (2007) to lack the full-length *RAP* mRNA and to exhibit retarded growth and a photobleached phenotype. We confirmed this phenotype, and we also uncovered a defect in photosynthetic activity in *rap-1* based on our measurements of the maximal efficiency of photosystem II (PSII) photochemistry (Figure 2B).

The phenotype of *rap-1* was complemented by introducing an *RAP* cDNA (Figure 2B). Even though 3-week-old complemented plants displayed slightly variegated and more serrated leaves than the wild type, their photosynthetic performance (as indicated by ratios of variable to maximum chlorophyll fluorescence [Fv/Fm]) was restored (Figure 2B). Except for a slightly retarded growth, 5-week-old plants displayed an almost completely wild-type phenotype. Because the introduced *RAP* sequence was expressed under control of the strong constitutive cauliflower mosaic virus *35S* promoter, the variegated leaf phenotype in younger plants suggests a dose-dependent function of RAP during early developmental stages. However, even though unlikely, we cannot formally exclude a second mutation in *rap-1* that might be responsible for the incomplete restoration of the wild-type phenotype in young complemented plants.

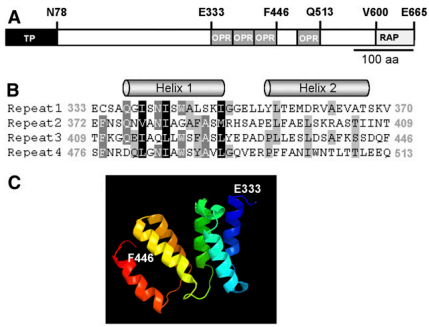


Figure 1. Structural Features of the *Arabidopsis* RAP Protein.

(A) RAP protein structure. The plastid transit sequence predicted by TargetP (Emanuelsson et al., 1999) is shown as a black box. OPR repeats are depicted as gray boxes and the C-terminal RAP domain as a light-gray box. Several characteristic amino acids are indicated above the diagram. aa, amino acids.

(B) Sequence alignment of the four OPR repeats in RAP displayed with GeneDoc (Nicholas and Nicholas, 1997). Amino acid positions describing the beginning and end of the respective repeat are indicated in gray. The two α-helices in each repeat predicted by Jpred (www.compbio.dundee.ac.uk/www.jpred; Cole et al., 2008) are depicted above the sequences.

(C) 3D protein structure prediction. The structure model for the consecutive OPR tract including repeats 1 to 3 (amino acids 333 to 446) was predicted with the Phyre2 server (http://www.sbg.bio.ic.ac.uk/phyre2/; Kelley and Sternberg, 2009). The first (E333) and last (F446) amino acids are indicated. Rainbow coloring is shaded from blue (N-terminal site) to red (C-terminal site).

Because photosynthesis was clearly affected in *rap-1*, we analyzed the levels of core proteins of photosynthetic complexes in the *rap-1* mutant line (Figure 2C). Whereas amounts of the nucleus-encoded light-harvesting complex II (LHCII) proteins were found to be unaffected in *rap-1*, amounts of all chloroplast-encoded proteins tested, including the PSII reaction center protein D1, the photosystem I reaction center protein PsaA, and the large subunit of Rubisco (RbcL), were clearly reduced.

We next investigated the de novo synthesis rates of chloroplast-encoded proteins by performing in vivo ^{35}S pulse-labeling experiments. As shown in Figure 2D, the overall protein synthesis rate was markedly lower in *rap-1* relative to the wild type. The analysis of chloroplast transcripts revealed reductions in the steady state levels of most of the analyzed mRNAs, like *rbcL*, *atpA*, and *petA*, although the *psbA* transcript was present in wild-type amounts (Figure 2E). Most strikingly, inspection of the ethidium bromide–stained rRNAs used as a loading control uncovered a dramatic reduction in plastid *16S* rRNA in mutant plants, whereas

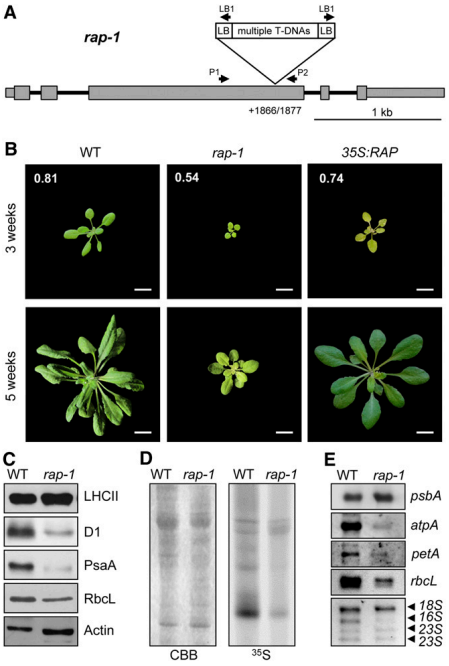


Figure 2. Characterization of the *rap-1* Mutant.

(A) Schematic depiction of the T-DNA insertion site in *rap-1*. Exons are shown as gray boxes, 5′ and 3′ UTRs as thinner gray boxes, and introns as black lines. The exact position of the T-DNA insertion site within the *RAP* gene in *rap-1*, identified by sequencing the DNA flanking the insertion site (Supplemental Figure 2A), is indicated (position +1866/1877 with respect to the translation initiation site). Primers used to identify homozygous mutants are indicated by the arrows above the gene model (Supplemental Figure 2B). The T-DNA insert is not drawn to scale.

(B) Growth phenotype of *rap-1* and its complementation. The wild type, the *rap-1* mutant, and *rap-1* complemented with *RAP* cDNA (*35S:RAP*) were grown for 3 and 5 weeks as indicated. Fv/Fm values for 3-week-old plants are shown in white numbers on the top three photographs. Bars = 1 cm.

(C) Accumulation of chloroplast-encoded proteins in *rap-1*. Total protein extracts (30 μg) from 3-week-old wild-type and *rap-1* plants were subjected to immunoblot analysis using antibodies against the proteins indicated on the right. β-Actin was used as the loading control. The RbcL protein was detected with an antiserum raised against the spinach Rubisco holoenzyme on a parallel, identical blot.

(D) In vivo translation assay. ^{35}S-labeled thylakoid proteins from wild-type and *rap-1* plants were separated by SDS-PAGE. The Coomassie blue–stained gel (CBB) and the autoradiograph (^{35}S) are shown.

(E) Accumulation of chloroplast transcripts in *rap-1*. Total RNAs from 3-week-old wild-type and *rap-1* plants were subjected to RNA gel blot analysis using the gene-specific probes indicated on the right. rRNAs on the ethidium bromide–stained gel were used as loading control (bottom panel).

the cytoplasmic *18S* rRNA and plastid *23S* rRNAs accumulated normally (Figure 2E, bottom panel). Because *16S* rRNA is required for ribosome assembly, and, therefore, for translation in chloroplasts, these data are consistent with the observed general decrease in chloroplast protein synthesis (Figure 2D) and explain the growth-retarded, chlorotic phenotype of *rap-1* mutants (Figure 2B).

RAP Is Required for Maturation of *16S* rRNA

Chloroplast *16S* rRNA is cotranscribed with *23S*, *4.5S*, and *5S* rRNAs, as well as two tRNAs, yielding a single RNA precursor that undergoes a complex series of processing events (Figure 3A). To investigate these events in detail, RNA gel blot analyses were performed on total leaf RNAs from 3-week-old wild-type and *rap-1* plants using rRNA-specific probes (Figure 3B).

We detected equal amounts of the full-length (7.4 kilonucleotides [knt]) rRNA precursor in *rap-1* and the wild type, indicating that accumulation of the full-length precursor rRNA is not significantly altered in the mutant (Figure 3B, probes A, C, and D). However, as suggested by the ethidium bromide–stained gels (Figures 2E and 3B), much less of the mature *16S* rRNA of ~1.5 knt is present in *rap-1* (Figure 3B, probe B), whereas probes specific for *16S* rRNAs retaining unprocessed 5′ and 3′ ends revealed a dramatic accumulation of such *16S* precursors in *rap-1* (Figure 3B, probes A and C). Both 5′- and 3′-specific probes detected a *16S* rRNA precursor of ~1.9 knt. In addition, the 5′ probe detected a less abundant ~1.7-knt precursor that was not identified with the 3′ probe. This indicates that some, albeit incomplete, processing of the 3′ end of the immature 1.9-knt *16S* species occurs, leading to the accumulation of 5′ unprocessed but 3′ processed *16S* rRNAs. Thus, failure to process the 5′ end does not preclude trimming of the 3′ end. In contrast with *16S* rRNA maturation, processing and accumulation of *23S*, *4.5S*, and *5S* rRNAs were not affected in *rap-1* (Figure 3B, probes D to F). Moreover, wild-type levels of mature *16S* rRNA were restored in *rap-1* mutants transformed with *RAP* cDNA, confirming that the lack of RAP is responsible for the defect in the maturation of *16S* rRNA (Figure 3C).

To identify the 5′ ends of the *16S*-related transcripts that accumulate in the *rap-1* mutant, a primer extension analysis was performed (Figures 4A and 4B). In agreement with the RNA gel blot analysis, the total amount of correctly 5′ processed, mature *16S* RNA was appreciably reduced in the *rap-1* mutant relative to the wild type (Figure 4B). By contrast, *pre-16S* rRNA 5′ ends originating from initiation at the P2 promoter at position −112 (transcribed from the plastid-encoded RNA polymerase PEP), as well as from processing at position −31, were clearly more abundant in the *rap-1* mutant. In addition to these known *16S* rRNA 5′ ends, we observed some transcript ends downstream of P2 that are more abundant in *rap-1* than in the wild type and are likely to represent unspecific processing and/or degradation products (Figure 4B). Taken together, these data indicate that maturation of *16S* rRNA is inefficient in the absence of RAP.

RAP Functions by Binding to the *16S* rRNA Precursor

A recent analysis of RNA deep-sequencing data sets identified 50 small RNAs (sRNAs) in the chloroplast of *Arabidopsis*, which are hypothesized to represent footprints of RNA binding proteins

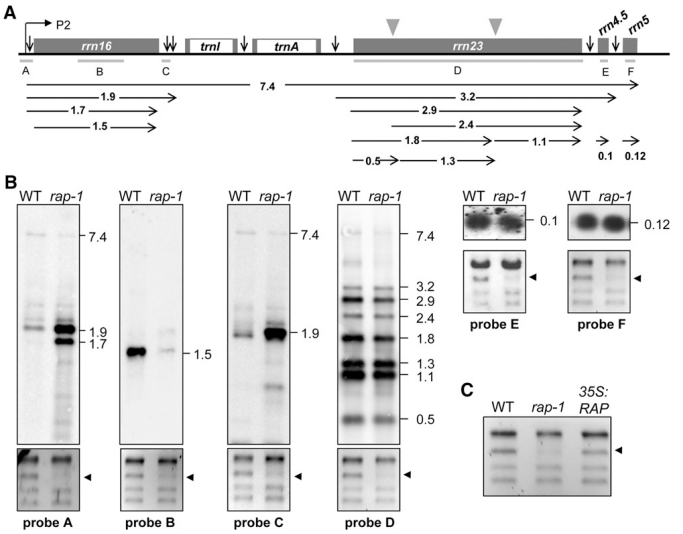


Figure 3. Accumulation of *16S* rRNA Precursors in *rap-1* Plants.

(A) Schematic representation of the chloroplast *rrn* operon in *Arabidopsis*. Gray boxes indicate exons and white boxes introns. The P2 promoter is represented by the bent arrow. Vertical arrows indicate processing sites in the primary transcript of the *rrn* operon. The locations of the probes used in **(B)** are marked by gray lines under the operon (A to F). Positions of internal cleavage sites (hidden breaks) in the *23S* rRNA are shown as gray triangles. Black horizontal arrows below the operon indicate locations and sizes (in knt) of the primary transcript and the various processing products.

(B) RNA gel blot analyses of chloroplast rRNAs from the wild type and *rap-1*. Mature rRNAs and precursors were detected with probes A to F shown in **(A)**. Transcript sizes are indicated in knt to the right of each panel. Results for probe F were obtained by reprobing the filter shown for probe D. Ethidium bromide–stained gels of rRNAs were used as loading controls, and the *16S* rRNA is indicated by black arrowheads (bottom panels).

(C) Wild-type levels of mature *16S* rRNA are restored in complemented *rap-1* mutants. Total RNAs from 3-week-old wild-type, *rap-1*, and *rap-1* plants complemented with *RAP* cDNA under control of the *35S* promoter (*35S:RAP*) were fractionated on a denaturing agarose gel and stained with ethidium bromide.

that protect them from degradation (Ruwe and Schmitz-Linneweber, 2012). Three such putative footprints had been identified within the *16S* precursor (Figure 4A; Supplemental Figure 3; Ruwe and Schmitz-Linneweber, 2012). RNA probes spanning these possible RAP footprints were hybridized to total RNAs prepared from wild-type or *rap-1* plants and analyzed in an RNase protection assay (Figure 4C). With probes 2 and 3, protected fragments were obtained in both the wild type and *rap-1*, indicating that these two sRNAs accumulate independently of RAP (Figure 4C). However, probe 1, spanning an 18-nucleotide footprint downstream of the P2 promoter (FP1), detected two sRNAs of 18 to 20 nucleotides in length that are protected by total RNA from the wild type but not from *rap-1* plants (Figure 4C; Supplemental Figure 3). Weaker signals in this size range obtained for *rap-1* were also detected when yeast tRNA was used as a negative control, and these were therefore not considered as protected fragments. Hence, these data strongly suggest that RAP mediates its function by binding ~100 nucleotides upstream of the 5′ end of the mature *16S* rRNA. As RAP seems not to be involved in the protection of the other two putative footprints within the *16S* precursor rRNA, other RNA binding proteins, like PPRs, might play a complementary or independent role in the *16S* maturation process.

RAP Interacts Directly with RNA in Vitro

To test for an intrinsic RNA binding activity of the recombinant RAP (rRAP) protein, which would support its direct involvement in sRNA protection, in vitro RNA binding assays were performed (Figure 5). As expression of RAP in fusion with a glutathione *S*-transferase tag resulted in very low overall expression levels, the protein was fused to the maltose binding protein (MBP), which had previously been used to successfully express PPR proteins (Beick et al., 2008; Barkan et al., 2012). Using this system, reasonably high expression levels and sufficient amounts of soluble RAP protein were obtained. To exclude an interference of the MBP tag with the RNA binding capacity of RAP, we proteolytically removed the tag prior to RNA binding assays (Figure 5A).

An in vitro–transcribed RNA probe spanning the putative binding site for RAP in the *16S* 5′ region (position −117 to −68

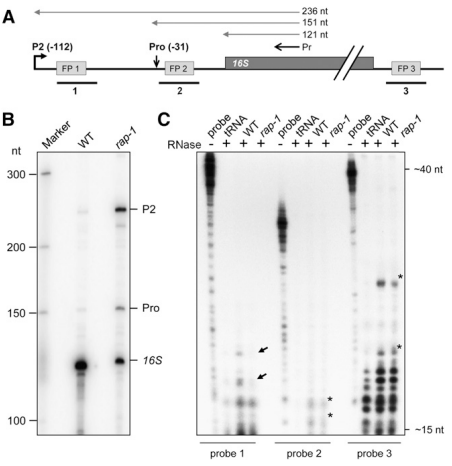


Figure 4. RAP Is Involved in 5′ End Processing of *16S* rRNA.

(A) Schematic representation of the *16S* rRNA precursor in *Arabidopsis*. Positions of the P2 (−112) promoter and the precursor processing site (Pro, −31) are indicated with respect to the start of the mature transcript (Lerbs-Mache, 2000). The primer (Pr) used for the primer extension analysis shown in **(B)** as well as expected extension products are depicted as black or gray vertical arrows, respectively, above the gene model. Positions of footprints (FP1-3) described by Ruwe and Schmitz-Linneweber (2012) are shown as gray boxes (for sequences, see Supplemental Figure 3). Probes used for the RNase protection assay in **(C)**, which span these footprints, are indicated as black lines below the model (1 to 3). nt, nucleotides.

(B) Primer extension analysis of *16S* rRNA 5′ ends. Total RNAs from wild-type and *rap-1* plants were subjected to primer extension analysis using the primer depicted in **(A)**. Known 5′ ends are indicated on the right. Sizes of bands of single-stranded DNA markers are indicated on the left.

(C) RNase protection assay. Total RNAs from wild-type or *rap-1* plants were hybridized with the respective radiolabeled probe indicated below the panel (cf. **[A]**) and treated with single-strand specific RNases A and T1. Protected fragments were analyzed on a sequencing gel alongside 1/30 of the respective undigested hybridization probe (probe). Probes incubated with yeast tRNA before RNase digestion (lane "tRNA") were used as a control. Black arrows mark fragments that are less abundant in *rap-1* and asterisks major fragments protected in both the wild type and *rap-1*. Expected sizes of fragments were estimated from the running fronts of xylene cyanol (~40 nucleotides) and bromophenol blue (~15 nucleotides) indicated on the right.

with respect to the start of the mature *16S* rRNA), was cross-linked to rRAP by irradiation with UV light. As positive control, we used the RNA binding protein RBP40, which had previously been shown to bind unspecifically to RNA in vitro (Schwarz et al., 2012; Bohne et al., 2013). Purified MBP served as a negative control. A single signal in the expected size range for rRAP (67 kD) was observed, indicating that rRAP directly interacts with RNA and that the protein preparation contains no substantial contaminating RNA binding activities from *Escherichia coli* (Figure 5B). Since UV cross-linking of RNA and protein leads to their covalent linkage, this assay is not suitable for detection of RNA binding activities under noncompetitive conditions.

We therefore employed filter binding assays that leave RAP in its native state to determine the equilibrium constant (K_d) for the binding reactions of RAP to different RNAs (Figure 6A). Besides the putative target RNA (*pre-16S* 5′ region), we included its complementary sequence (as *pre-16S* 5′ region) as well as sequences of the *psbD* 5′ untranslated region (UTR) and the noncoding *trnN* 5′ region. All probes were similar in length and GC content and exhibited a similar or lower propensity to form secondary structures than the specific probe (determined by calculation of the free energy ΔG of the thermodynamic ensemble of RNA structures). Prior to the binding reactions, the integrity and concentration of RNA probes was verified by gel electrophoresis (Supplemental Figure 4A).

Whereas the K_d value obtained for the putative target RNA was ~101 nM, and therefore similar to those measured for other chloroplast RNA–protein interactions (Ostersetzer et al., 2005; Hammani et al., 2012; Bohne et al., 2013), K_d values for the *psbD*, *trnN*, and antisense probes were considerably higher and could not be determined under these conditions. This supports the specific binding of RAP to the footprint RNA identified in vivo (Figure 4C) and suggests that RAP itself carries the main determinants required for specific recognition of its binding site within the 5′ region of the precursor of *16S* rRNA. However, a slightly increased affinity of rRAP was observed for the *16S* antisense probe compared with the other nontarget probes.

In order to further substantiate the RNA binding specificity of RAP, competition experiments using a similar filter binding assay were performed. Same amounts of rRAP were incubated with the radiolabeled RNA probe containing its putative binding site in the presence of either homologous or heterologous competitor RNAs. The concentration and integrity of RNA probes was again verified by gel electrophoresis (Supplemental Figure 4B). As shown in Figure 6B, the competing effect was strongest when the

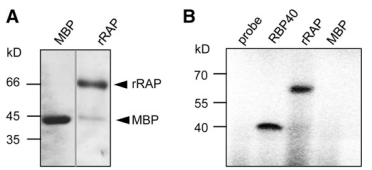


Figure 5. rRAP Exhibits an Intrinsic RNA Binding Capacity.

(A) Purification of rRAP protein. Coomassie blue–stained SDS-PAGE gel showing the affinity-purified rRAP protein after removal of the maltose binding protein tag that was electrophoresed alongside authentic MBP. Mobilities of size markers are indicated on the left. Note that the two samples were electrophoresed on the same gel but not in adjacent lanes.

(B) UV cross-linking experiment. Purified rRAP protein, together with two control proteins (MBP and the RNA binding protein RBP40), was analyzed after UV cross-linking in the presence of a radiolabeled RNA probe corresponding to the *16S* region spanning FP1 (*pre-16S* 5′ region). Sizes of marker bands are given in kilodaltons on the left.

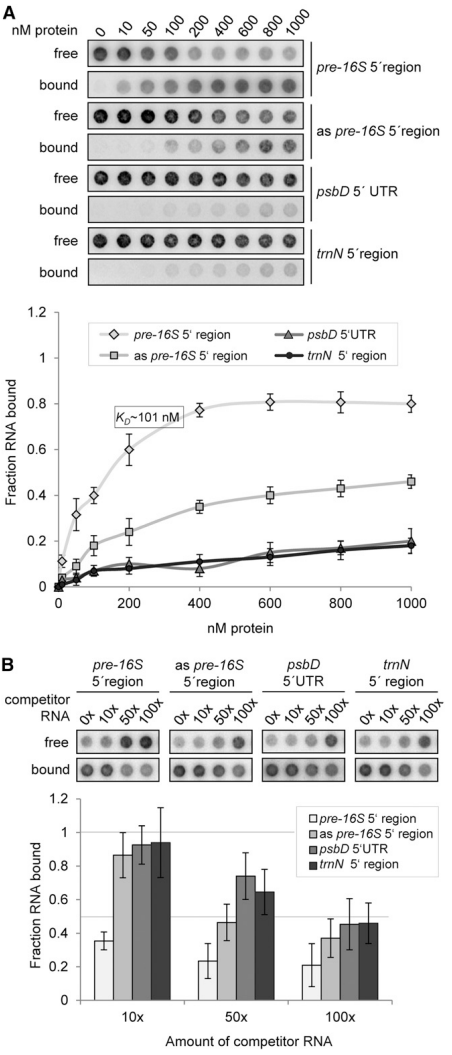


Figure 6. rRAP Binds Preferentially to the 5′ Region of the *16S* Precursor Transcript.

(A) Determination of RNA binding curves. Binding reactions containing 6 pM ^{32}P-labeled RNA of each indicated RNA and increasing molarities of rRAP were filtered through stacked nitrocellulose and nylon membranes using a dot-blot apparatus (top panel). Signal intensities for nitrocellulose-bound protein-RNA complexes (bound) as well as nylon membrane–bound free RNAs (free) were quantified by phosphor imaging. The binding curves were determined from three experiments performed as triplicates with the same rRAP preparation (bottom panel). Calculated means are shown with standard deviations indicated by error bars. The equilibrium binding constant (K_d) of rRAP and the *pre-16S* 5′ region probe was determined to be 101 nM as indicated.

(B) Competition experiments. Binding reactions containing rRAP protein, ^{32}P-labeled RNA of the *pre-16S* 5′ region, and the indicated molar excess of competitor RNAs representing the homologous RNA, sequences of the *psbD* 5′ UTR, the *trnN* 5′ noncoding region, or the antisense sequence of the radiolabeled *pre-16S* 5′ region (as *pre-16S* 5′ region), respectively, were treated as described in **(A)**. Signal intensities obtained for each reaction without competitor RNA were set to 1. Three independent experiments were performed as triplicates for each reaction, and calculated means are shown with standard deviations indicated by error bars (bottom panel).

homologous RNA was used, thereby confirming a specific binding of RAP to the *16S* 5′ probe. In agreement with data from the binding curves, the competition experiments revealed an increased competition effect of the antisense probe compared with the *trnN* and *psbD* probes. Interestingly, a comparison of sense and antisense sequences elucidated a sequence of eight identical nucleotides corresponding to the 3′ end of the identified footprint (Supplemental Figure 5A). Hence, the antisense probe includes approximately half of the putative RAP binding site, and this most likely accounts for the somewhat higher affinity of RAP to this RNA compared with the other nonspecific RNAs.

RAP Associates with Chloroplast Nucleoids

Little is known about the spatial organization of the process of rRNA maturation in chloroplasts. So far, evidence derives from recent proteomic data from maize, which suggest that the nucleoid, the site of the chloroplast genome, and a region of DNA-RNA-protein assembly is the major location of ribosome assembly and rRNA processing in the chloroplast (Majeran et al., 2012). We therefore attempted to determine whether RAP is targeted to chloroplasts and to determine its suborganellar localization. For this purpose, green fluorescent protein (GFP) was fused to the C-terminal end of RAP and transiently expressed in tobacco (*Nicotiana benthamiana*) protoplasts under the control of a cauliflower mosaic virus *35S* promoter. As shown in Figure 7, the fusion protein accumulated in distinct spots overlapping the chlorophyll autofluorescence of the chloroplasts. Furthermore, the RAP-GFP signal was colocalized with 4′,6-diamidino-2-phenylindole (DAPI)–stained nucleoid DNA, indicating its association with the nucleoids and thus supporting the idea that the nucleoid is the site of *16S* rRNA maturation.

DISCUSSION

RAP Is an RNA Binding Protein Assisting in Maturation of *16S* rRNA

Because of the importance of OPR proteins in chloroplast RNA metabolism in algae, the precise molecular role of RAP, a negative

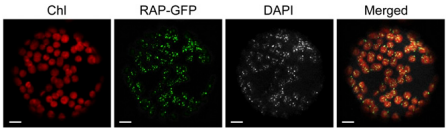


Figure 7. RAP Is Associated with Chloroplast Nucleoids.

Confocal laser scanning microscopy of tobacco protoplasts transiently expressing RAP fused to GFP (RAP-GFP). The autofluorescence of chloroplasts is shown in red (Chl). To visualize chloroplast nucleoid DNA, the protoplasts were stained with DAPI. The merged images reveal colocalization of the RAP-GFP and DAPI signals in the chloroplasts (Merged). Bars = 10 μm.

regulator of plant defense and the only OPR in *Arabidopsis*, is of particular interest. We investigated the RAP knockout line *rap-1* (Katiyar-Agarwal et al., 2007), which exhibits slow growth and impaired photosynthesis (Figure 2B). We found that protein synthesis in the chloroplast was severely affected as the consequence of a specific defect in the trimming/processing of the chloroplast *16S* rRNA precursor transcript (Figures 2C to 2E, 3B, and 4B). Similar to *E. coli*, rRNA maturation and ribosome assembly in the chloroplast are closely linked processes, as indicated by altered ribosome biogenesis in rRNA maturation mutants (Bisanz et al., 2003; Williams and Barkan, 2003; Schmitz-Linneweber et al., 2006; Bollenbach et al., 2007). Defects in rRNA maturation can lead to reduced polysomal loading in these mutants and consequently can result in reduced chloroplast translation (Barkan, 1993; Bellaoui et al., 2003; Beligni and Mayfield, 2008; Sharwood et al., 2011). As suggested by Bisanz et al. (2003) for the *Arabidopsis Dal* mutant, which similar to *rap-1* accumulates *16S* rRNA precursors, we also propose that the accumulation of these precursors in the *rap-1* mutant diminishes translational efficiency by preventing the formation of active ribosomes. In *E. coli*, it has been reported that complete processing of the *16S* rRNA requires the association with the ribosomal *30S* subunit, and the final maturation steps are thought to take place in polysomes (Shajani et al., 2011). Moreover, in vitro reconstitution assays revealed that the bacterial *30S* subunits containing *16S* precursors are inactive, suggesting that the processing of the *16S* rRNA is required for protein synthesis (Wireman and Sypherd, 1974). However, recent data reveal that extensions at the *16S* 5′ end have little effect on ribosome assembly itself, and it is rather assumed that the *16S* leader sequences preclude appropriate *16S* folding required for translational fidelity of the ribosome (Roy-Chaudhuri et al., 2010; Gutgsell and Jain, 2012).

Most chloroplast mRNAs investigated in *rap-1* were also less abundant than in the wild type (Figure 2E). Consistently, lower steady state levels of several chloroplast mRNAs have previously been found in other mutants defective in rRNA maturation, suggesting that such transcripts require ribosomal loading and/or translation for efficient stabilization (Barkan, 1993; Yamamoto et al., 2000; Bisanz et al., 2003). RAP probably mediates its function by directly binding, via its OPR domain, to the 5′ end of the *16S* rRNA precursor (Figures 4C, 5, and 6). RNA binding activity of OPR repeats has recently been demonstrated for the translation initiation factor Tab1 from *C. reinhardtii* (Rahire et al., 2012). However, more detailed molecular work will be required to define the individual contributions of each of the repeats in RAP to the recognition of its target RNA.

Since plastid RNases that are thought to participate in *16S* rRNA maturation (e.g., RNase J and RNase R) are assumed to have little or no intrinsic sequence specificity, one may speculate that RAP might facilitate *16S* maturation by conferring sequence specificity on these enzymes (Stoppel and Meurer, 2012; Germain et al., 2013). If so, it would functionally resemble the RHON1 protein from *Arabidopsis*, which has been suggested to bind to the 5′ end of chloroplast *23S* rRNA and confer sequence specificity on endonuclease RNase E (Stoppel et al., 2012). A possible joint function of RNase J and RAP in *16S* 5′ end maturation is supported by an analysis of RNase J–deficient *Arabidopsis* and tobacco plants, which reveal, similar to *rap-1* plants, a decreased accumulation of mature *16S* rRNA as well as a series of 5′ extensions of *16S* rRNA (Sharwood et al., 2011). Moreover, an interactive role of RNase J and helical repeat proteins has been recently reported by Luro et al. (2013). Here, the authors postulated that RNase J trims chloroplast mRNA 5′ ends to mature forms defined by bound PPR proteins. However, if and how RAP and RNase J interact for *16S* 5′ end maturation remains elusive.

Alternatively, it is also conceivable that binding of RAP facilitates the accessibility of sequence-specific RNase(s) and/or accessory factors by modifying secondary structures in the *16S* precursor 5′ region (compare with Supplemental Figure 5) or that RAP itself reveals an intrinsic RNase activity.

Furthermore, our finding that RAP is associated with nucleoids provides cytological evidence for a sublocalized *16S* rRNA processing in the chloroplast (Figure 7). This is further sustained by the identification of the maize ortholog in the nucleoid proteome (Majeran et al., 2012) and strongly supports the idea that chloroplast transcription and ribosome assembly are tightly coupled (Majeran et al., 2012; Germain et al., 2013). Interestingly, only recently, the *E. coli pre-16S* rRNA 5′ leader region has been shown to associate with the nucleoid, and this association depends on RNase III, which participates in the maturation of pre-rRNAs

```
                                *          20
A.thaliana      : GAATA---TGAAGCGCATGGA : 18
P.trichocarpa   : GAATA---TGAAGCGCATGGA : 18
S.oleracea      : GAATA---TGAAGCGCATGGA : 18
V.vinifera      : GAATA---TGAAGCGCATGGA : 18
Z.mays          : TAATAATCTGAAGCGCATGGA : 21
O.sativa        : TAATAATCTGAAGCGCATGGA : 21
H.vulgare       : TAATAATCTGAAGCGCATGGA : 21
B.distachyon    : TAATAATCTGAAGCGCATGGA : 21
P.patens        : TAATA---TGAAGCAAATGAA : 18
C.reinhardtii   : AAATT---AAAACACTGGGA : 18
```

Figure 8. Conservation of the Putative RAP Binding Site.

Alignment of the *16S* 5′ region corresponding to footprint 1 in *Arabidopsis* (Supplemental Figure 3) with respective segments of the *16S* 5′ region of indicated species. Black shading represents 100% conservation and dark gray and gray 80 and 60%, respectively. For sequence accession numbers, see Methods.

(Malagon, 2013). One might therefore also speculate that RAP plays a role in nucleoid localization of the chloroplast *16S* precursor by binding to its 5′ end.

Evolutionary Conservation of RAP Function

The conservation of a single orthologous OPR in most land plants, the identification of the maize OPR in the nucleoid proteome, and the predicted presence of an organellar transit sequence in all OPR orthologs analyzed prompted us to speculate that, like RAP, they too function in plastid *16S* rRNA metabolism. In the dicots so far analyzed, this possibility is supported by the perfect conservation of the putative RAP footprint in the corresponding *16S* precursors, whereas sequences from monocotyledonous plants reveal an insertion of three nucleotides (Figure 8). Interestingly, the sequence is least conserved in *P. patens* and reveals three nucleotide exchanges compared with the analyzed dicotyledonous plant sequences (Figure 8). As the *P. patens* OPR protein also exhibits the lowest degree of conservation relative to other plant OPRs examined (Supplemental Figure 1A), this might reflect coevolution of the OPR proteins with their respective binding sites. However, whether and how these OPRs are involved in plastid rRNA metabolism in organisms other than *Arabidopsis* remains to be elucidated. Nonetheless, based on the function of RAP in the basic process of rRNA maturation and its conservation in land plants, it is tempting to speculate that RAP played an important role during the early evolutionary development of the chloroplast.

Remarkably, no obvious ortholog of RAP can be identified by sequence similarity searches of the *C. reinhardtii* genome. Consistently, there is no obvious conservation of the putative footprint of RAP in the sequence upstream of the mature *16S* rRNA in the alga (Figure 8). In this context, it is especially noteworthy that transcription of the *16S* precursor in *C. reinhardtii* has been reported to initiate downstream of the region in which the footprint in *Arabidopsis* is located (Schneider et al., 1985).

A further question that arises is why, in contrast with algae, Streptophyte genomes generally encode only a single OPR protein. Given that a certain number of RNA binding proteins is required to guarantee the specific and precise processing of diverse organellar transcripts, including tRNA and rRNA precursors, it appears likely that members of the PPR protein family, which is highly diverse in Streptophytes, have assumed many of the roles performed by OPRs in algae.

RAP as a Negative Regulator of Plant Defense

Interestingly, a role of RAP as a negative regulator of plant disease resistance was previously noted because its expression is downregulated via a small interfering RNA (*Arabidopsis lsiRNA-1*) upon infection with *Pseudomonas syringae* (Katiyar-Agarwal et al., 2007). Accordingly the *rap-1* mutant, also used in this study, was found to show higher resistance to this pathogen than the wild type (Katiyar-Agarwal et al., 2007). However, the molecular function of RAP was not investigated further. Our analysis of RAP now suggests a possible explanation: Induced downregulation of *RAP* expression by *Arabidopsis lsiRNA-1* upon pathogen infection could contribute to a downregulation of chloroplast protein synthesis and therefore reduce photosynthetic activity at the infection site. This would agree with the local decrease in photosynthesis observed in *P. syringae*–infected *Arabidopsis* leaves (Bonfig et al., 2006). How this downregulation of photosynthesis then triggers further defense measures against the pathogen remains to be clarified.

METHODS

Plant Material and Growth Conditions

Arabidopsis thaliana Columbia-0 (wild-type) and the *rap-1* T-DNA line (SAIL_1223_C10; Syngenta *Arabidopsis* Insertion Library T-DNA collection; Sessions et al., 2002) were grown on soil under controlled greenhouse conditions (70 to 90 μmol m^{-2} s^{-1}, 16-h-light/8-h-dark cycles). T-DNA insertion lines homozygous for *rap-1* were identified by PCR using gene-specific (P1m 5′-TTAAGGGTCAAGAGATTGCTC-3′; P2, 5′-AATCAAGCCCTGTACTTATAAGAA-3′) and T-DNA–specific (LB1, 5′-GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC-3′) primers. For mutant complementation, the *RAP* cDNA was cloned into the vector pH2GW7 using Gateway technology (Invitrogen) to create the construct *RAP*/pH2GW7 (Karimi et al., 2002). Homozygous *rap-1* plants were transformed with *Agrobacterium tumefaciens* strain GV3101 containing the construct *RAP*/pH2GW7 by the floral dip method (Holsters et al., 1980; Zhang et al., 2006).

Chlorophyll Fluorescence Measurements

The maximum quantum yield of PSII of single leaves was calculated from the Fv/Fm measured with a FluorCam 800 MF (Photon Systems Instruments) according to the manufacturer's instructions.

In Vivo Translation Assay of Thylakoid Proteins

In vivo radioactive ^{35}S labeling of thylakoid proteins was performed as described by Armbruster et al. (2010) using five *Arabidopsis* leaves each from the wild type or the *rap-1* mutant, harvested at the 12-leaf rosette stage (of the wild type).

RNA Preparation and Transcript Analysis

Frozen leaves from 3-week-old plants were ground in liquid nitrogen, and RNA was extracted using TriReagent (Sigma-Aldrich) according to the manufacturer's instructions. RNA gel blot analysis of total RNA from *rap-1* and wild-type plants was performed using standard methods. Specific transcripts were detected with digoxigenin-labeled PCR products.

In Vitro Synthesis of RNA and UV Cross-Linking to rRAP

To express the recombinant *Arabidopsis* RAP protein (rRAP), a cDNA sequence encoding amino acids 79 to 671 was inserted into the plasmid pMAL-c5x (New England Biolabs). Expression was performed in *Escherichia coli* Rosetta cells (Novagen) by induction with 0.5 mM isopropyl β-D-1-thiogalactopyranoside for 3 h at 30°C. Purification of the recombinant protein was performed according to the New England Biolabs protocol for purification of MBP-tagged recombinant proteins, including the removal of the MBP tag by proteolytic digestion with factor Xa. Recombinant RBP40 was expressed as previously described by Bohne et al. (2013). UV cross-linking experiments were performed essentially as described by Zerges and Rochaix (1998). The primers T7 top strand (5′-atgtaatacgactcactatagggg-3′) and rm16 5′ bottom (5′-tacattatgctgagtgatatcccTCGCTTGAGGTACGCTTTATACTTCGCGTACCTATGTTCAATACTGAAC-3′) were annealed to create a DNA template for in vitro synthesis of 5′ *pre-16S* rRNA. The template contained the T7 promoter (sequence in lowercase letters). Hybridized primers were transcribed in vitro using T7 RNA polymerase and digested with DNase I (Promega) according to the manufacturer's protocol. Reactions

were extracted with phenol-chloroform and ethanol precipitated. Binding reactions were performed at room temperature for 5 min and contained 20 mM HEPES/KOH, pH 7.8, 5 mM $MgCl_2$, 60 mM KCl, 500 ng protein, and 100 kcpm of ^{32}P-labeled RNA probe. Protein-bound RNA probes were UV cross-linked (1 J/cm^2), and nonbound ^{32}P-RNA probes were digested with 10 units of RNase One (Promega) for 20 min at 37°C. Samples were fractionated by SDS-PAGE and analyzed by phosphor imaging.

Primer Extension and RNase Protection Assays

Aliquots (2 μg) of total leaf RNA were used for primer extension reactions according to standard protocols (Sambrook and Russell, 2001). The oligonucleotide (PE rrn16 coding 5′-GGGCAGGTTCTTACGCGT-3′) and the marker (GeneRuler Low Range DNA Ladder; Thermo Scientific) were end-labeled with [γ-^{32}P]ATP (Hartmann Analytic) using T4 polynucleotide kinase (New England Biolabs). Unincorporated nucleotides were removed with the QIAquick nucleotide removal kit (Qiagen) according to the manufacturer's instructions. Primer extensions were performed at 55°C with Superscript III reverse transcriptase (Invitrogen), and the products were fractionated on 6% polyacrylamide sequencing gels and analyzed by phosphor imaging.

For RNase protection assays, probes were transcribed and radiolabeled in vitro as described above using the following annealed primer pairs: probe 1 sRNA 16rrn5′-1neu forward 5′-taatacgactcactatagggTCATTCCAAGTCGTGGCTTGTATCCATGCGCTTCATATTC-3′/sRNA 16rrn5′-1neu reverse 5′-attatgctgagtgatatcccAGTAAGGTTCAGTATTGAACATAGGTACGCGAAGTATAAG-3′; probe 2 sRNA 16rrn5′-2 forward 5′-taatacgactcactatagggCAGATGCTTCTTCCTTCGATATTCATTACGTTGATACTTA-3′/sRNA 16rrn5′-2 reverse 5′-attatgctgagtgatatcccGTCTACGAAGAAGGAAGCTATAAGTAATGCAACTATGAAT-3′; probe 3 sRNA 16rrn3′ forward 5′-taatacgactcactatagggGAAAAGTCCCTCTCGATTACGAAGAACCCATAAATCCAAA-3′/sRNA 16rrn3′ reverse 5′-attatgctgagtgatatcccCTTTTCAGGGAGAGCTAATGCTTCTTGGGTATTTAGGTTT-3′. The T7 promoter sequence is given in lowercase letters. Probes were gel purified and amounts equivalent to 1 to 5 × 10^4 cpm were hybridized to 2-μg aliquots of total RNA in hybridization buffer (1.5 M KCl, 0.1 M Tris, pH 8.3, and 10 mM EDTA) at 42°C. RNase A and RNase T1 were added to the hybridization reactions to final concentrations of 200 μg/mL or 5000 units/mL, respectively, and incubated for 45 min at 37°C. Nucleic acids were ethanol precipitated, electrophoresed on 12% sequencing gels, and analyzed by phosphor imaging.

Determination of RNA Binding Curves and Competition Experiments

The RNA binding curves and the K_d value for the specific RNA were determined as described by Bohne et al. (2013). Binding reactions were performed at room temperature for 15 min and contained 20 mM HEPES/KOH, pH 7.8, 5 mM $MgCl_2$, 60 mM KCl, 0.5 mg/mL heparin, and 6 pM of the indicated ^{32}P-labeled RNA probe. To generate templates for in vitro transcription of RNAs the primer T7 top strand (5′-atgtaatacgactcactataggg-3′) was annealed with rrn16 5′ bottom (5′-tacattatgctgagtgatatcccTCGCTTGAGGTACGCTTATACTTCGCGTACCTATGTTCAATACTGAAC-3′), rrn16 5′ antisense (5′-tacattatgctgagtgatatcccAGCGAACTCCATGCGAATATGAAGCGCATGGATACAAGTATGACTTG-3′), psbD (5′-tacattatgctgagtgatatccTTGTAATTCCACAAGCCTTTACCAACTTCATCTACTTATCCTCCTAGC-3′), or trnN (5′-tacattatgctgagtgatatcccGTACCCAACTCTTGCCCTTAACTTGAGATACTCTAGATTAGAGGGCAA-3′). RNA was in vitro transcribed as described above and probes were gel purified according to Ostersetzer et al. (2005). Molarities of RNA probes were calculated based on the quantitation of incorporated ^{32}P-labeled UTP using a Mini Monitor G-M tube (Mini Instruments). Further steps of the filter binding assays were performed as described for the K_d value determination by Bohne et al. (2013). Results were quantified using ImageQuantTL (GE Healthcare). For competition experiments, reactions containing rRAP (600 nM) and a ^{32}P-labeled fragment of the *pre-16S* 5′ region (6 pM) premixed with increasing amounts of cold competitor RNA were incubated in binding buffer (20 mM HEPES/KOH, pH 7.8, 5 mM $MgCl_2$, and 60 mM KCl) at room temperature for 15 min. RNA in vitro transcription (competitor RNAs with 1/1000 of [^{32}P]UTP compared with the labeled RNA probe), purification, quantification, and subsequent steps were performed as described for the binding curves and above.

Agrobacterium-Mediated Transient Expression in Tobacco (*Nicotiana benthamiana*)

The *Arabidopsis RAP* cDNA was cloned into the vector pK7FWG2 (Karimi et al., 2002) using the Gateway technology (Invitrogen). Transient expression of the corresponding RAP-GFP construct was achieved by *Agrobacterium*-mediated infiltration of 4-week-old tobacco leaves. To this end, 30 mL of cultures of AGL-1 agrobacteria, previously transformed with the RAP-GFP construct, were harvested by centrifugation and resuspended in induction medium (10 mM MES/KOH, pH 6, 10 mM $MgCl_2$, and 200 μM acetosyringone). Following incubation at 28°C for 2 h at 75 rpm, cells were resuspended in 5% Suc containing 200 μM acetosyringone, and tobacco leaves were infiltrated with the cell suspension at OD_{600} = 0.7. Afterwards, plants were kept in the greenhouse for 3 d, and protoplasts were isolated according to Koop et al. (1996). GFP fluorescence was detected at 672 to 750 nm and chlorophyll autofluorescence monitored at 503 to 542 nm by laser scanning microscopy (Leica TCS SP5/DM 6000B, argon laser, excitation wavelength of 488 nm). For DNA staining, protoplasts were incubated for 10 min with 1 μg/mL DAPI and directly examined with a UV laser (excitation wavelength 405 nm/ detection at 423 to 490 nm). All images were processed with Leica SAF Lite software (Leica).

Accession Numbers

The Arabidopsis Genome Initiative locus identifier for RAP is At2g31890. DNA sequence data from alignment in Figure 8 can be found in the GenBank data library under the following accession numbers: *Arabidopsis* (AP000423.1), *Physcomitrella patens* (AP005672), *Oryza sativa* (JN861110), *Populus trichocarpa* (AC208093), *Zea mays* (AY928077), *Spinacia oleracea* (AJ400848), *Hordeum vulgare* (EF115541), *Brachypodium distachyon* (EU325680), *Vitis vinifera* (DQ424856), and *Chlamydomonas reinhardtii* (BK000554.2). Protein sequence data from alignment in Supplemental Figure 1 can be found in the GenBank data library under the following accession numbers: *Arabidopsis* (AEC08600.1), *P. trichocarpa* (XP_002331644), *Z. mays* (DAA52984.1), and *O. sativa* (NP_001050400.1). The *P. patens* sequence was obtained from the cosmoss genome browser (Pp1s157_38G2, www.cosmoss.org).

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Sequence Alignment and Targeting Predictions for RAP and Its Orthologs in Higher Plants and Moss.

Supplemental Figure 2. PCR Analysis of *rap-1* Mutants.

Supplemental Figure 3. Distribution of Footprints within the *16S* rRNA Precursor.

Supplemental Figure 4. Integrity of Probes Used for RNA Binding Assays in Figure 6.

Supplemental Figure 5. Formation of a Potential Stem Loop Structure at the *16S* rRNA Precursor 5′ End.

Supplemental References.

ACKNOWLEDGMENTS

This work was supported by grants from the Deutsche Forschungsgemeinschaft to J.N. (Grant Ni390/4-2) and K.P. (Grant PH73/3-3). We

thank Jürgen Soll for providing the LHCII antibody. The antiserum against the spinach Rubisco holoenzyme was kindly provided by Günther Wildner. We thank Olivier Vallon for helpful discussion and for sharing data on the distribution of the OPR family.

AUTHOR CONTRIBUTIONS

A.-V.B. and J.N. designed the research. L.K., F.W., and A.-V.B. conducted the analysis of the mutant, recombinant proteins, and bioinformatic protein characteristics. R.S. performed GFP import experiments. A.-V.B., J.N., K.P., and L.K. wrote the article. J.N. and K.P. contributed reagents/materials/analysis tools.

Received January 10, 2014; revised January 10, 2014; accepted February 3, 2014; published February 28, 2014.

REFERENCES

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

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

Ban, T., Ke, J., Chen, R., Gu, X., Tan, M.H.E., Zhou, X.E., Kang, Y., Melcher, K., Zhu, J.-K., and Xu, H.E. (2013). Structure of a PLS-class pentatricopeptide repeat protein provides insights into mechanism of RNA recognition. J. Biol. Chem. **288:** 31540–31548.

Barkan, A. (1993). Nuclear mutants of maize with defects in chloroplast polysome assembly have altered chloroplast RNA metabolism. Plant Cell **5:** 389–402.

Barkan, A. (2011). Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. Plant Physiol. **155:** 1520–1532.

Barkan, A., Rojas, M., Fujii, S., Yap, A., Chong, Y.S., Bond, C.S., and Small, I. (2012). A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. PLoS Genet. **8:** e1002910.

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

Beligni, M.V., and Mayfield, S.P. (2008). *Arabidopsis thaliana* mutants reveal a role for CSP41a and CSP41b, two ribosome-associated endonucleases, in chloroplast ribosomal RNA metabolism. Plant Mol. Biol. **67:** 389–401.

Bellaoui, M., Keddie, J.S., and Gruissem, W. (2003). DCL is a plant-specific protein required for plastid ribosomal RNA processing and embryo development. Plant Mol. Biol. **53:** 531–543.

Bisanz, C., Bégot, L., Carol, P., Perez, P., Bligny, M., Pesey, H., Gallois, J.-L., Lerbs-Mache, S., and Mache, R. (2003). The Arabidopsis nuclear DAL gene encodes a chloroplast protein which is required for the maturation of the plastid ribosomal RNAs and is essential for chloroplast differentiation. Plant Mol. Biol. **51:** 651–663.

Bohne, A.-V., Schwarz, C., Schottkowski, M., Lidschreiber, M., Piotrowski, M., Zerges, W., and Nickelsen, J. (2013). Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. PLoS Biol. **11:** e1001482.

Bollenbach, T., Schuster, G., Portnoy, V., and Stern, D.B. (2007). Processing, degradation, and polyadenylation of chloroplast transcripts. In Cell and Molecular Biology of Plastids, R. Bock, ed (Berlin, Heidelberg: Springer), pp. 175–211.

Bonfig, K.B., Schreiber, U., Gabler, A., Roitsch, T., and Berger, S. (2006). Infection with virulent and avirulent *P. syringae* strains differentially affects photosynthesis and sink metabolism in *Arabidopsis* leaves. Planta **225:** 1–12.

Cole, C., Barber, J.D., and Barton, G.J. (2008). The Jpred 3 secondary structure prediction server. Nucleic Acids Res. **36:** W197-201.

Das, A.K., Cohen, P.W., and Barford, D. (1998). The structure of the tetratricopeptide repeats of protein phosphatase 5: Implications for TPR-mediated protein-protein interactions. EMBO J. **17:** 1192–1199.

Del Campo, E.M. (2009). Post-transcriptional control of chloroplast gene expression. Gene Regul. Syst. Biol. **3:** 31–47.

Eberhard, S., Loiselay, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., Choquet, Y., and Wollman, F.-A. (2011). Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of *atpA* transcripts in *Chlamydomonas reinhardtii* chloroplasts. Plant J. **67:** 1055–1066.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. Protein Sci. **8:** 978–984.

Germain, A., Hotto, A.M., Barkan, A., and Stern, D.B. (2013). RNA processing and decay in plastids. Wiley Interdiscip. Rev RNA **4:** 295–316.

Gutgsell, N.S., and Jain, C. (2012). Gateway role for rRNA precursors in ribosome assembly. J. Bacteriol. **194:** 6875–6882.

Hammani, K., Cook, W.B., and Barkan, A. (2012). RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. Proc. Natl. Acad. Sci. USA **109:** 5651–5656.

Holsters, M., et al. (1980). The functional organization of the nopaline *A. tumefaciens* plasmid pTiC58. Plasmid **3:** 212–230.

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends Plant Sci. **7:** 193–195.

Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A., and Jin, H. (2007). A novel class of bacteria-induced small RNAs in *Arabidopsis*. Genes Dev. **21:** 3123–3134.

Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: A case study using the Phyre server. Nat. Protoc. **4:** 363–371.

Koop, H.U., Steinmüller, K., Wagner, H., Rössler, C., Eibl, C., and Sacher, L. (1996). Integration of foreign sequences into the tobacco plastome via polyethylene glycol-mediated protoplast transformation. Planta **199:** 193–201.

Lee, I., and Hong, W. (2004). RAP—A putative RNA-binding domain. Trends Biochem. Sci. **29:** 567–570.

Lerbs-Mache, S. (2000). Regulation of rDNA transcription in plastids of higher plants. Biochimie **82:** 525–535.

Luro, S., Germain, A., Sharwood, R.E., and Stern, D.B. (2013). RNase J participates in a pentatricopeptide repeat protein-mediated 5′ end maturation of chloroplast mRNAs. Nucleic Acids Res. **41:** 9141–9151.

Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., Watkins, K.P., Barkan, A., and van Wijk, K.J. (2012). Nucleoid-enriched proteomes in developing plastids and chloroplasts from maize leaves: A new conceptual framework for nucleoid functions. Plant Physiol. **158:** 156–189.

Malagon, F. (2013). RNase III is required for localization to the nucleoid of the 5′ pre-rRNA leader and for optimal induction of rRNA synthesis in *E. coli*. RNA **19:** 1200–1207.

Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J.-D., and Goldschmidt-Clermont, M. (2006). A novel multifunctional factor involved in trans-splicing of chloroplast introns in Chlamydomonas. Nucleic Acids Res. **34:** 262–274.

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

Nicholas, K.B., and Nicholas, H.B.J. (1997). GeneDoc: A Tool for Editing and Annotating Multiple Sequence Alignments. (Distributed by the author).

Ostersetzer, O., Cooke, A.M., Watkins, K.P., and Barkan, A. (2005). CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. Plant Cell **17:** 241–255.

Perron, K., Goldschmidt-Clermont, M., and Rochaix, J.-D. (2004). A multiprotein complex involved in chloroplast group II intron splicing. RNA **10:** 704–711.

Rahire, M., Laroche, F., Cerutti, L., and Rochaix, J.-D. (2012). Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. Plant J. **72:** 652–661.

Ringel, R., Sologub, M., Morozov, Y.I., Litonin, D., Cramer, P., and Temiakov, D. (2011). Structure of human mitochondrial RNA polymerase. Nature **478:** 269–273.

Roy-Chaudhuri, B., Kirthi, N., and Culver, G.M. (2010). Appropriate maturation and folding of 16S rRNA during 30S subunit biogenesis are critical for translational fidelity. Proc. Natl. Acad. Sci. USA **107:** 4567–4572.

Ruwe, H., and Schmitz-Linneweber, C. (2012). Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? Nucleic Acids Res. **40:** 3106–3116.

Sambrook, J., and Russell, D. (2001). Molecular Cloning: A Laboratory Manual. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

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

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

Schneider, M., Darlix, J.L., Erickson, J., and Rochaix, J.-D. (1985). Sequence organization of repetitive elements in the flanking regions of the chloroplast ribosomal unit of *Chlamydomonas reinhardii*. Nucleic Acids Res. **13:** 8531–8541.

Schwarz, C., Bohne, A.-V., Wang, F., Cejudo, F.J., and Nickelsen, J. (2012). An intermolecular disulfide-based light switch for chloroplast *psbD* gene expression in *Chlamydomonas reinhardtii*. Plant J. **72:** 378–389.

Sessions, A., et al. (2002). A high-throughput *Arabidopsis* reverse genetics system. Plant Cell **14:** 2985–2994.

Shajani, Z., Sykes, M.T., and Williamson, J.R. (2011). Assembly of bacterial ribosomes. Annu. Rev. Biochem. **80:** 501–526.

Sharwood, R.E., Halpert, M., Luro, S., Schuster, G., and Stern, D.B. (2011). Chloroplast RNase J compensates for inefficient transcription termination by removal of antisense RNA. RNA **17:** 2165–2176.

Shikanai, T., and Fujii, S. (2013). Function of PPR proteins in plastid gene expression. RNA Biol. **10:** 1446–1456.

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

Stoppel, R., and Meurer, J. (2012). The cutting crew - Ribonucleases are key players in the control of plastid gene expression. J. Exp. Bot. **63:** 1663–1673.

Stoppel, R., Manavski, N., Schein, A., Schuster, G., Teubner, M., Schmitz-Linneweber, C., and Meurer, J. (2012). RHON1 is a novel ribonucleic acid-binding protein that supports RNase E function in the *Arabidopsis* chloroplast. Nucleic Acids Res. **40:** 8593–8606.

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

Williams, P.M., and Barkan, A. (2003). A chloroplast-localized PPR protein required for plastid ribosome accumulation. Plant J. **36:** 675–686.

Wireman, J.W., and Sypherd, P.S. (1974). *In vitro* assembly of *30S* ribosomal particles from precursor *16S* RNA of *Escherichia coli*. Nature **247:** 552–554.

Yamamoto, Y.Y., Puente, P., and Deng, X.W. (2000). An *Arabidopsis* cotyledon-specific albino locus: A possible role in *16S* rRNA maturation. Plant Cell Physiol. **41:** 68–76.

Zerges, W., and Rochaix, J.-D. (1998). Low density membranes are associated with RNA-binding proteins and thylakoids in the chloroplast of *Chlamydomonas reinhardtii*. J. Cell Biol. **140:** 101–110.

Zhang, X., Henriques, R., Lin, S.-S., Niu, Q.-W., and Chua, N.-H. (2006). Agrobacterium-mediated transformation of *Arabidopsis thaliana* using the floral dip method. Nat. Protoc. **1:** 641–646.

The Plant Cell, Vol. 28: 984–989, April 2016, www.plantcell.org

CORRECTION[OPEN]

Kleinknecht, L., Wang, F., Stübe, R., Philippar, K., Nickelsen, J., and Bohne, A.-V. (2014). RAP, the sole octotricopeptide repeat protein in *Arabidopsis*, is required for chloroplast *16S* rRNA maturation. Plant Cell **26:** 777–787.

In the course of on-going work, the authors realized that there were mistakes in the design of primers used to generate templates for in vitro transcription of RNA probes by the T7 RNA polymerase. Templates were generated by annealing primers with incorrectly positioned T7 promoter sequence elements in reverse primers. Therefore, no RNA synthesis should have occurred. However, as observed in native agarose gels as well as in the analysis of synthesized RNAs by RNase T1 digestion, misdesigned primers had a strong self-annealing capacity leading to undefined RNAs of expected sizes. As even correctly designed primers showed self-annealing, new experiments were performed either with PCR products used as templates for in vitro transcription or synthetic RNA oligos.

While the general conclusion on the function of RAP in *16S* rRNA maturation is not affected by these errors, their consequence is that the determination of the RAP binding site within the *16S* precursor RNA (Figure 4C) as well as in vitro RAP binding affinities to RNAs (Figure 6) were not correctly resolved, for which the authors apologize. The corrected experiments do not support binding of RAP to FP1 as stated before. Instead, rRAP showed a higher affinity to the FP2 probe compared with the two other reported footprint sequences. However, the affinity of rRAP for FP2 seems to be only moderately increased compared with FP1 and FP3, for which no distinct footprint was detected (corrected Figures 4C and 6). Therefore, it is also possible that RAP binds to another sequence within the *16S* rRNA precursor or that additional determinants like overall rRNA structure or other *trans*-acting factors enhance selective binding of RAP to FP2 in vivo. Nonetheless, additional data provided in Figure 9 support a role of RAP in precise trimming of the mature *16S* 5′ end.

A brief description of the problems associated with each figure and corrections made is provided here, followed by side-by-side presentation of the original and corrected figures and the new methods (and associated references) used to prepare the corrected figures.

Figure 4B. The previously shown primer extension analysis in Figure 4B is correct and only replaced because an additional control mutant defective in *16S* rRNA processing, *rbf1-1* (Fristedt et al., 2014), is included in the analysis. Note the apparent extension of the "mature" *16S* transcript in *rap-1* compared with the wild-type and *rbf1-1*.

Figure 4C. Due to the high self-annealing capacity of primers, an annealing strategy of in vitro transcription templates was considered unsuitable for the generation of specific RNA probes. Consequently, the RNase protection experiment has been replaced by a RNA gel blot analysis of respective footprints. These new data suggest that FP2 instead of the formerly described FP1 region represents the RAP-dependent RNA footprint.

Figure 5. The experiment was repeated using a PCR product as template for the generation of in vitro transcribed RNA. The results obtained are identical to those in the original figure and reveal an unspecific intrinsic RNA binding by rRAP.

Figure 6. The experiment was repeated using synthetic FP1-FP3-specific RNA oligos. rRAP showed a slightly higher affinity to FP2 compared with the other tested RNA oligos.

Figure 8. This figure is correct but represents an alignment of footprint 1 (FP1) sequences. As it is now possible that FP2 is the RAP binding site, a new alignment of FP2 related sequences is provided.

Figure 9. To confirm the 5′ extension of "mature" *16S* transcripts in *rap-1* observed in Figure 4C, we additionally mapped precise 5′ and 3′ ends of *16S*-related transcripts by circular RT-PCR (cRT-PCR). In contrast to *rbf1-1* and the wild type, we could not detect any transcript in *rap-1* with a correct mature 5′ end. All transcripts started either at P2, Pro-29, or had a 1-nucleotide extension (starting at −1). In addition, we found many transcripts in *rap-1* with truncated 5′ and 3′ ends. While *16S* precursors starting at −112 (P2) or −29 (Pro) and with longer 3′ extensions/truncations were occasionally observed also in the wild type or *rbf1-1*, we never detected any 5′ 1-nucleotide extensions in these plants.

www.plantcell.org/cgi/doi/10.1105/tpc.16.00094

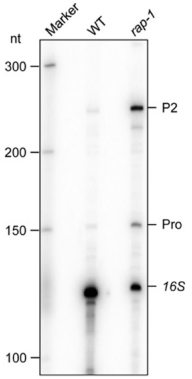


Figure 4B. Original: Primer Extension Analysis of *16S* rRNA 5′ Ends.

Total RNAs from wild-type and *rap-1* plants were subjected to primer extension analysis using the primer depicted in **(A)**. Known 5′ ends are indicated on the right. Sizes of bands of single-stranded DNA markers are indicated on the left.

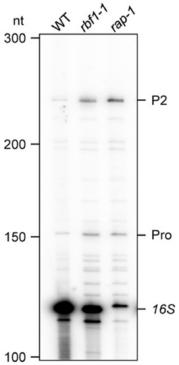


Figure 4B. Corrected: Primer Extension Analysis of *16S* rRNA 5′ Ends.

Total RNAs from wild-type, *rbf1-1*, and *rap-1* plants were subjected to primer extension analysis using the primer depicted in **(A)**. Known 5′ ends are indicated on the right. Sizes of bands of single-stranded DNA markers are indicated on the left.

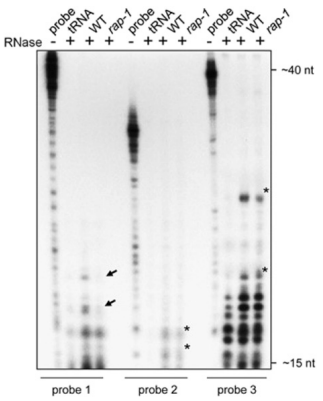


Figure 4C. Original: RNase Protection Assay.

Total RNAs from wild-type or *rap-1* plants were hybridized with the respective radiolabeled probe indicated below the panel (cf. **[A]**) and treated with single-strand specific RNases A and T1. Protected fragments were analyzed on a sequencing gel alongside 1/30 of the respective undigested hybridization probe (probe). Probes incubated with yeast tRNA before RNase digestion (lane "tRNA") were used as a control. Black arrows mark fragments that are less abundant in *rap-1* and asterisks major fragments protected in both the wild type and *rap-1*. Expected sizes of fragments were estimated from the running fronts of xylene cyanol (~40 nucleotides) and bromophenol blue (~15 nucleotides) indicated on the right.

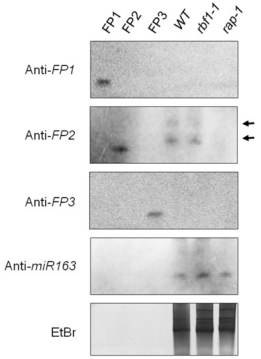


Figure 4C. Corrected: Analysis of Small RNAs in *rap-1*.

RNA gel blot analyses of small RNAs from the wild type, *rap-1*, and an additional control RNA from the *Arabidopsis rbf1-1* mutant, described to also reveal a defect in *16S* rRNA processing (Fristedt et al., 2014). Thirty micrograms of total leaf RNA was fractionated in denaturing polyacrylamide gels and transferred to a charged nylon membrane. DNA oligonucleotides that mimic each sRNA (FP1-FP3) were run in adjacent lanes. DNA probes used, which are complementary to the respective small RNA, are indicated on the left. Note that single-stranded DNA migrates slightly faster than single-stranded RNA. As a positive control, we used microRNA *miR163* (Ha et al., 2009). Two small RNAs specific to FP2 that were only detected in the wild type and *rbf-1*, but not in the *rap-1* mutant, are indicated by black arrows. A representative ethidium bromide-stained gel is shown to demonstrate equal loading.

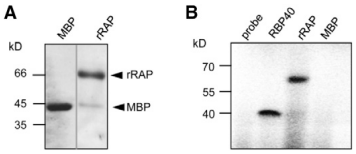


Figure 5. Original: rRAP Exhibits an Intrinsic RNA Binding Capacity.

(A) Purification of rRAP protein. Coomassie blue–stained SDS-PAGE gel showing the affinity-purified rRAP protein after removal of the maltose binding protein tag that was electrophoresed alongside authentic MBP. Mobilities of size markers are indicated on the left. Note that the two samples were electrophoresed on the same gel but not in adjacent lanes.
(B) UV cross-linking experiment. Purified rRAP protein, together with two control proteins (MBP and the RNA binding protein RBP40), was analyzed after UV cross-linking in the presence of a radiolabeled RNA probe corresponding to the *16S* region spanning FP1 (*pre-16S* 5′ region). Sizes of marker bands are given in kilodaltons on the left.

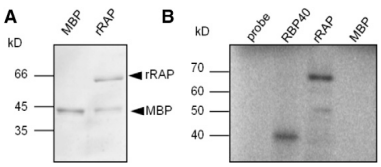


Figure 5. Corrected: rRAP Exhibits an Intrinsic RNA Binding Capacity.

(A) Purification of rRAP protein. Coomassie blue–stained SDS-PAGE gel showing the affinity-purified rRAP protein after removal of the maltose binding protein tag that was electrophoresed alongside authentic MBP. Mobilities of size markers are indicated on the left.
(B) UV cross-linking experiment. Purified rRAP protein, together with two control proteins (MBP and the RNA binding protein RBP40), was analyzed after UV cross-linking in the presence of a radiolabeled RNA probe corresponding to the 5′ *pre-16S* region. Sizes of marker bands are given in kilodaltons on the left.
The PCR product used as DNA template for in vitro synthesis of the 5′ *pre-16S* rRNA region was amplified using the following set of primers: 16S 5′ (−139) T7 forward (5′-taatacgactcactatagggGGTAGGGGTAGCTATATTTCTG-3′) and *16S* 5′ (+57) reverse (5′-ATGTGTTAAGCATGCCGC-3′).

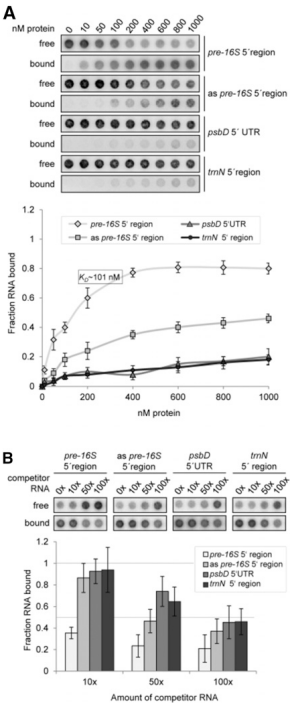


Figure 6. Original: rRAP Binds Preferentially to the 5′ Region of the *16S* Precursor Transcript.

(A) Determination of RNA binding curves. Binding reactions containing 6 pM ^{32}P-labeled RNA of each indicated RNA and increasing molarities of rRAP were filtered through stacked nitrocellulose and nylon membranes using a dot-blot apparatus (top panel). Signal intensities for nitrocellulose bound protein-RNA complexes (bound) as well as nylon membrane–bound free RNAs (free) were quantified by phosphor imaging. The binding curves were determined from three experiments performed as triplicates with the same rRAP preparation (bottom panel). Calculated means are shown with standard deviations indicated by error bars. The equilibrium binding constant (K_d) of rRAP and the *pre-16S* 5′ region probe was determined to be 101 nM as indicated.
(B) Competition experiments. Binding reactions containing rRAP protein, ^{32}P-labeled RNA of the *pre-16S* 5′ region, and the indicated molar excess of competitor RNAs representing the homologous RNA, sequences of the *psbD* 5′ UTR, the *trnN* 5′ noncoding region, or the antisense sequence of the radiolabeled *pre-16S* 5′ region (as *pre-16S* 5′ region), respectively, were treated as described in **(A)**. Signal intensities obtained for each reaction without competitor RNA were set to 1. Three independent experiments were performed as triplicates for each reaction, and calculated means are shown with standard deviations indicated by error bars (bottom panel).

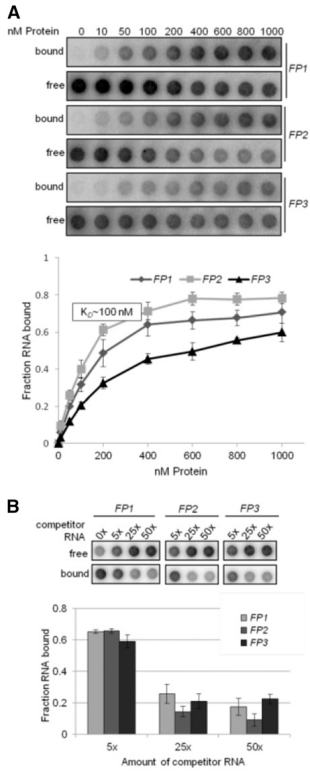


Figure 6. Corrected: RNA Binding Specificity of rRAP to Footprint Regions within the *16S* Precursor Transcript.

(A) Determination of RNA binding curves. Binding reactions containing 6 pM ^{32}P-labeled RNA of each indicated RNA and increasing molarities of rRAP were filtered through stacked nitrocellulose and nylon membranes using a dot-blot apparatus (top panel). Signal intensities for nitrocellulose-bound protein-RNA complexes (bound) as well as nylon membrane–bound free RNAs (free) were quantified by phosphor imaging. The binding curves were determined from three experiments performed as triplicates (bottom panel). Calculated means are shown with standard deviations indicated by error bars. The equilibrium binding constant (K_d) of rRAP and the FP2 probe was determined to be ~100 nM as indicated.
(B) Competition experiments. Binding reactions containing rRAP protein, ^{32}P-labeled FP2 RNA, and the indicated molar excess of competitor RNAs representing unlabeled FP1, FP2, and FP3 RNA oligos, respectively, were treated as described in **(A)**. Signal intensities obtained for each reaction without competitor RNA were set to 1. Three independent experiments were performed as triplicates for each reaction, and calculated means are shown with standard deviations indicated by error bars (bottom panel).

```
                    *          20
A.thaliana    : GAATA---TGAAGCGCATGGA : 18
P.trichocarpa : GAATA---TGAAGCGCATGGA : 18
S.oleracea    : GAATA---TGAAGCGCATGGA : 18
V.vinifera    : GAATA---TGAAGCGCATGGA : 18
Z.mays        : TAATAATGTGAAGCGCATGGA : 21
O.sativa      : TAATAATGTGAAGCGCATGGA : 21
H.vulgare     : TAATAATGTGAAGCGCATGGA : 21
B.distachyon  : TAATAATGTGAAGCGCATGGA : 21
P.patens      : TAATA---TGAAGCAAATGAA : 18
C.reinhardtii : AAATT---AAAAACACTGGGA : 18
```

Figure 8. Original: Conservation of the Putative RAP Binding Site.

Alignment of the *16S* 5′ region corresponding to footprint 1 in *Arabidopsis* (Supplemental Figure 3) with respective segments of the *16S* 5′ region of indicated species. Black shading represents 100% conservation and dark gray and gray 80 and 60%, respectively. For sequence accession numbers, see Methods.

```
                    *          20
A.thaliana    : AAGGAAGCTAATAAGTAATGC : 20
P.trichocarpa : AAGGAAGCTAATAAGTAATGC : 20
S.oleracea    : AAGGAAGCTAATAAGTAATGC : 20
V.vinifera    : AAGGAAGCTAATAAGTAATGC : 20
Z.mays        : AAGGAAGCTAATAAGTAATGC : 20
O.sativa      : AAGGAAGCTAATAAGTAATGC : 20
H.vulgare     : AAGGAAGCTAATAAGTAATGC : 20
B.distachyon  : AAGGAAGCTAATAAGTAATGC : 20
P.patens      : AAGGAAGCTAATAAGTAATTT : 20
C.reinhardtii : AATCCCGACAAAATTAAAC : 20
```

Figure 8. Corrected: Conservation of the Potential RAP Binding Site.

Alignment of the *16S* 5′ region corresponding to footprint 2 in *Arabidopsis* (Supplemental Figure 3) with respective segments of the *16S* 5′ region of indicated species. Black shading represents 100% conservation and dark gray and gray 80 and 60%, respectively. For sequence accession numbers, see Methods.

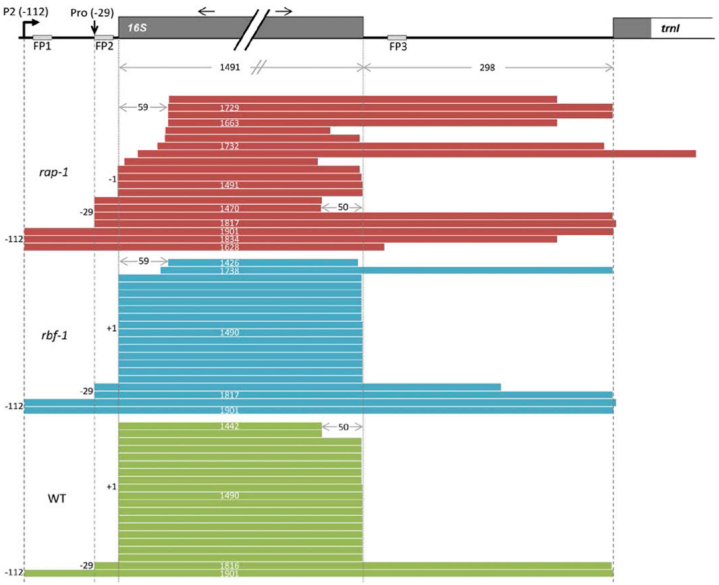


Figure 9. cRT-PCR.

16S rRNA ends were deduced from cRT-PCR clones ($n = 20$). Each bar represents a single clone. A schematic representation of a part of the chloroplast *rrn* operon is shown above the diagram. Dark-gray boxes indicate exons, white boxes introns, and light-gray boxes predicted footprints. The P2 promoter is represented by a bent arrow. The vertical arrow indicates the mapped processing site at −29 with respect to the start of the mature *16S* rRNA (previously annotated as "Pro-31" in Bisanz et al., 2003). Black horizontal arrows indicate the positions and directions of the cRT-PCR primer pair.

METHODS

RNA Preparation and Transcript Analysis

Frozen leaves from 3-week-old plants were ground in liquid nitrogen, and RNA was extracted using TRI Reagent (Sigma-Aldrich) according to the manufacturer's instructions. RNA gel blot analysis of total RNA from *rap-1* and wild-type plants was performed using standard methods. Specific transcripts were detected with digoxigenin-labeled PCR products.

RNA gel blots for detection of small RNAs were basically performed as described by Zhelyazkova et al. (2012). Before hybridization, RNAs were cross-linked to the membrane using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride according to Pall and Hamilton (2008). The oligonucleotides used as probes (FP1, 5′-TCCATGCGCTTCATATTC-3′; FP2, 5′-GCATTACTTATAGCTTCCTT-3′; FP3, 5′-ATACCCAAGAAGCATTAGCTCTCC-3′; miR163, 5′- ATCGAAGTTCCAAGTCCTCTTCAA-3′) were end-labeled with [γ-^{32}P]ATP (Hartmann Analytic) using T4 polynucleotide kinase (New England Biolabs). Unincorporated nucleotides were removed with the QIAquick nucleotide removal kit (Qiagen) according to the manufacturer's instructions. Three DNA oligonucleotides (FP1, 5′-GAATATGAAGCGCATGGA-3′; FP2, 5′-AAGGAAGCTATAAGTAATGC-3′; FP3, 5′-GGAGAGCTAATGCTTCTTGGGTAT-3′) that mimic each sRNA were run on the gel as controls.

Determination of RNA Binding Curves and Competition Experiments

The RNA binding curves and the K_d value for the specific RNA were determined as described by Bohne et al. (2013). Synthetic RNA oligos (Integrated DNA Technologies; FP1, 5′-CGAAUAUGAAGCGCAUGGAUACAA-3′; FP2, 5′-GAAGGAAGCUAUAAGUAAUGCAAC-3′; and FP3, 5′-GGAGAGCUAAUGCUUCUUGGGUAU-3′) were 5′-labeled as described above, and probes were gel purified according to Ostersetzer et al. (2005). Binding reactions were performed at room temperature for 15 min and contained 20 mM HEPES/KOH, pH 7.8, 5 mM $MgCl_2$, 60 mM KCl, and 6 pM of the indicated ^{32}P-labeled RNA probe. Further steps of the filter binding assays were performed as described for the K_d value determination by Bohne et al. (2013). Results were visualized on a Storm phosphor imager and quantified using ImageQuantTL (GE Healthcare).

For competition experiments, reactions containing rRAP (600 nM) and the ^{32}P-labeled synthetic RNA oligo for FP2 (6 pM) premixed with increasing amounts of cold competitor RNA were incubated in binding buffer (20 mM HEPES/KOH, pH 7.8, 5 mM $MgCl_2$, 60 mM KCl, and 0.5 mg/mL heparin) at room temperature for 15 min. Subsequent steps were performed as described for the binding curves.

cRT-PCR

The cRT-PCR method was basically performed as described previously (Zimmer et al., 2012; Hotto et al., 2015). Two and a half micrograms of circularized wild-type, *rap-1*, and *rbf1-1* RNAs were reverse transcribed using SuperScript III with a gene-specific oligo (16S 5′ cRT-PCR F1, 5′-CACCCGTCCGCCACTGGAAACACCA-3′). Twenty percent of the RT reaction was used for amplification with the same oligo as before and an oligo binding close to the 3′ end of the *16S* rRNA (16S 3′ cRT-PCR R1, 5′-CTTAACCGCAAGGAGGGGGGTGCCGAA-3′) using a Taq polymerase. Purified PCR products (NucleoSpin Gel and PCR clean-up; Macherey-Nagel) were cloned with the CloneJET PCR cloning kit (Thermo Fisher Scientific) and sequenced with custom primers.

REFERENCES

Bisanz, C., Bégot, L., Carol, P., Perez, P., Bligny, M., Pesey, H., Gallois, J.-L., Lerbs-Mache, S., and Mache, R. (2003). The Arabidopsis nuclear DAL gene encodes a chloroplast protein which is required for the maturation of the plastid ribosomal RNAs and is essential for chloroplast differentiation. Plant Mol. Biol. **51:** 651–663.

Bohne, A.-V., Schwarz, C., Schottkowski, M., Lidschreiber, M., Piotrowski, M., Zerges, W., and Nickelsen, J. (2013). Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. PLoS Biol. **11:** e1001482.

Fristedt, R., Scharff, L.B., Clarke, C.A., Wang, Q., Lin, C., Merchant, S.S., and Bock, R. (2014). RBF1, a plant homolog of the bacterial ribosome-binding factor RbfA, acts in processing of the chloroplast *16S* ribosomal RNA. Plant Physiol. **164:** 201–215.

Ha, M., Lu, J., Tian, L., Ramachandran, V., Kasschau, K.D., Chapman, E.J., Carrington, J.C., Chen, X., Wang, X.-J., and Chen, Z.J. (2009). Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. Proc. Natl. Acad. Sci. USA **106:** 17835–17840.

Hotto, A.M., Castandet, B., Gilet, L., Higdon, A., Condon, C., and Stern, D.B. (2015). Arabidopsis chloroplast mini-ribonuclease III participates in rRNA maturation and intron recycling. Plant Cell **27:** 724–740.

Ostersetzer, O., Cooke, A.M., Watkins, K.P., and Barkan, A. (2005). CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. Plant Cell **17:** 241–255.

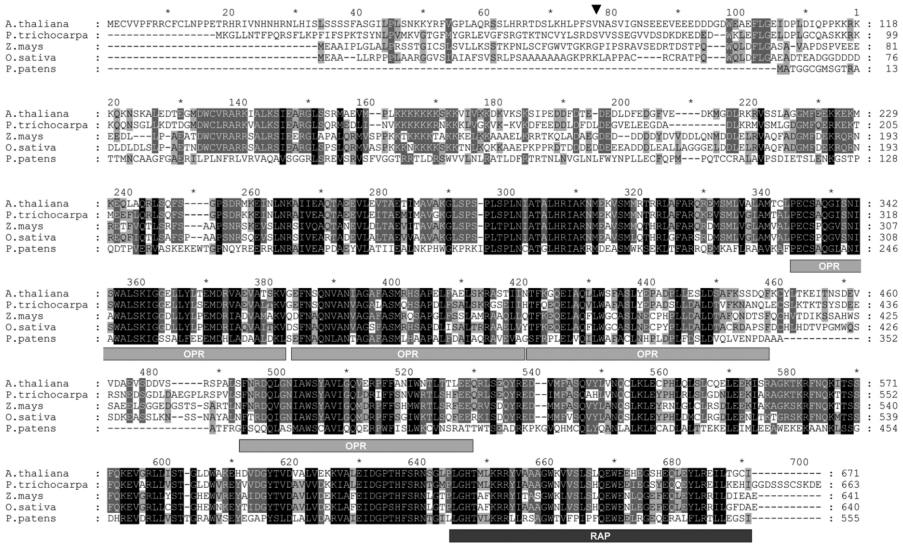
Pall, G.S., and Hamilton, A.J. (2008). Improved northern blot method for enhanced detection of small RNA. Nat. Protoc. **3:** 1077–1084.

Zhelyazkova, P., Hammani, K., Rojas, M., Voelker, R., Vargas-Suárez, M., Börner, T., and Barkan, A. (2012). Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. Nucleic Acids Res. **40:** 3092–3105.

Zimmer, S.L., McEvoy, S.M., Menon, S., and Read, L.K. (2012). Additive and transcript-specific effects of KPAP1 and TbRND activities on 3′ non-encoded tail characteristics and mRNA stability in *Trypanosoma brucei*. PLoS One **7:** e37639.

Editor's note: the corrected figure and accompanying text were reviewed by members of *The Plant Cell* editorial board. Both the original and corrected figures are shown for ease of comparison.

A



B

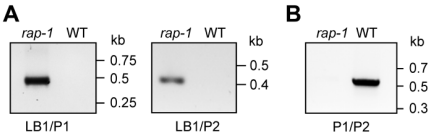
	TargetP[1]	Predotar[2]	WoLF PSORT[3]
A. thaliana	C	(C)	(C)
P. trichocarpa	M	(M)	C
Z. mays	C	C	C
O. sativa	C	-	C
P. patens	C	(C)	C

[1]Emanuelsson et al. 2000, Nielsen et al. 1997, [2]Small et al. 2004; [3]Paul et al. 2007

Supplemental Figure 1. Sequence Alignment and Targeting Predictions for RAP and its Orthologs in Higher Plants and Moss.

(A) Multiple sequence alignment of RAP (*A. thaliana*) and orthologous proteins from *P. patens* (*Physcomitrella patens*), *P. trichocarpa* (*Populus trichocarpa*), *Z. mays* (*Zea mays*) and *O. sativa* (*Oryza sativa cv. Japonica*). Sequence accession numbers are in Methods. The alignment was generated using ClustalW (Thompson et al. 2002) and displayed with Genedoc (http://www.psc.edu/biomed/genedoc; Nicholas et al. 1997). Shading as in Figure 8. OPR repeats and the RAP domain (pfam08373) are indicated. The cleavage site in the chloroplast transit peptide predicted by TargetP (Nielsen et al. 1997; Emanuelsson et al. 2000) is designated for RAP by the filled triangle above the sequence.

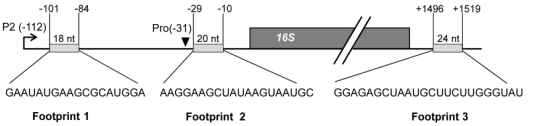
(B) Targeting predictions for proteins shown in **(A)**. C: chloroplast, M: mitochondria. Parentheses indicate a possible localization in the respective organelle.



Supplemental Figure 2. PCR Analysis of *rap-1* Mutants.

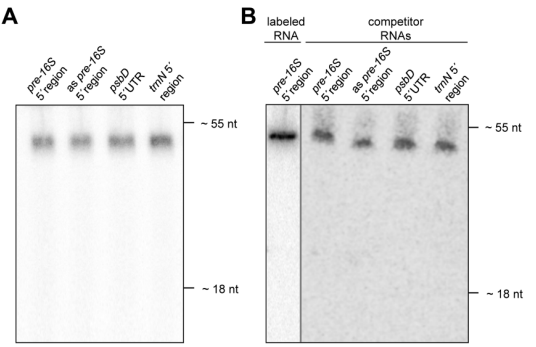
(A) Determination of T-DNA insertion site in *rap-1*. PCR analysis was performed with primers LB1/P1 (left panel) and LB1/ P2 (right panel) which amplify a sequence from both sites of the T-DNA insertion allele in *RAP* and the genomic flanking sequence. PCR products obtained were separated on an agarose gel alongside a size marker (marker lane not displayed) and subjected to sequence analysis to determine the exact site of T-DNA insertion. The positions of the gene- and T-DNA-specific primers used are depicted in Figure 2A.

(B) Identification of mutants homozygous for the insertion in *RAP*. PCR analysis was performed with primers P1 and P2 which bind to the *RAP* gene up- and downstream of the T-DNA insertion site in *rap-1*. PCR products were separated on an agarose gel alongside a size marker (marker lane not displayed). For positions of primers see Figure 2A.



Supplemental Figure 3. Distribution of Footprints within the *16S* rRNA Precursor.

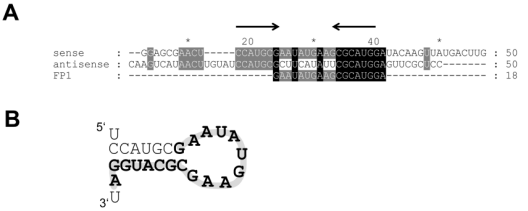
Positions of the PEP promoter (P2, -112) and the precursor processing site (Pro, -31) are indicated with respect to the 5' end of the mature transcript. Footprints identified by Ruwe and Schmitz-Linneweber (2012) are depicted as light grey boxes with given length and nucleotide sequences.



Supplemental Figure 4. Integrity of Probes used for RNA Binding Assays in Figure 6.

(A) Integrity of probes used for binding curves (Figure 6A). 3 fmol of labeled RNA were loaded on a denaturing 10% polyacrylamide gel. The dye migration of xylene cyanol (~55 nt) and bromophenol blue (~18 nt) is indicated on the right. Slight differences in signal intensities of the competitor RNAs correlate with numbers of radiolabeled U-residues in each transcript.

(B) Integrity of probes used for competition experiments (Figure 6B). 3 fmol of labeled RNA or 30 fmol of respective competitor RNAs (labeled with 1/1000 of ^{32}P UTP) were loaded on the same denaturing 10% polyacrylamide gel. As indicated by a grey line, the left part of the gel showing the competitor RNAs was exposed ten times longer than the right part and the contrast was adjusted to visualize the signals. The dye migration of xylene cyanol (~55 nt) and bromophenol blue (~18 nt) is indicated on the right. Slight differences in signal intensities of the competitor RNAs correlate with numbers of radiolabeled U-residues in each transcript.



Supplemental Figure 5. Formation of a Potential Stem Loop Structure at the *16S* rRNA Precursor 5' End.

(A) Sequence alignment of the *16S* precursor sense and antisense probes and the putative RAP binding site (FP1) displayed with Genedoc (Nicholas et al. 1997). The two inverted sequences which have the potential to form the stem loop structure shown in **(B)** are indicated by arrows above the sequences.

(B) Potential stem loop structure formed overlapping the RAP binding site. The binding site (FP1) is highlighted in grey and written in bold letters.

Supplemental References

Emanuelsson, O., Nielsen, H., Brunak, S. and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* **300:** 1005-1016.

Nicholas, K. B., Nicholas, H. B. and Deerfield, D. W. (1997). GeneDoc: analysis and visualization of genetic variation. *EMBNET. NEWS* **4:** 14.

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

Ruwe, H. and Schmitz-Linneweber, C. (2012). Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? *Nucleic Acids Res.* **40:** 3106-3116.

Thompson, J. D., Gibson, T. J. and Higgins, D. G. (2002). Multiple Sequence Alignment Using ClustalW and ClustalX. Current Protocols in Bioinformatics, John Wiley & Sons, Inc.

3.2 Analysis of the plastid RNase-sensitive DLA2 containing complex in Chlamydomonas reinhardtii

Kleinknecht, L., König, AC., Bohne, A.-V., and Nickelsen, J.

This study focuses on the moonlighting enzyme DLA2. While its primary function as a subunit of the cpPDC has been well described, an additional function as a RNA-binding protein involved in chloroplast gene expression under mixotrophic conditions has only been uncovered more recently.

Size exclusion chromatography and 2D-BN-PAGE disclosed that the PDC subunits E1 and E3 are not part of the RNase-sensitive DLA2 complex leading to the conclusion that the RNase-sensitive complex is indeed distinct from the cpPDC. To gain further insights, quantification of the cpPDC subunit protein levels in wild type grown under mixotrophic, photoautotrophic and heterotrophic conditions were performed, revealing that protein levels do not change under these different conditions and therefore the ratio between the subunits stays the same. Thus it can be concluded that no additional DLA2 is provided for its additional RNA binding function, but rather that the same pool of protein needs to be used alternatively. Importantly, it was shown that the E3 binding domain in DLA2 is indeed also the site of RNA binding. Remarkably, the *psbA* mRNA and the E3 protein bind to DLA in a competitive fashion. In addition, it became evident that auto-acetylation of DLA2, induced by incubation with high concentrations of acetyl-CoA, can influence its RNA binding activity. These finding supports the hypothesis that post-translational modifications might regulate the two functions of the DLA2 protein.

All the experiments in this study were performed by me, with the exception of the acetylation assay, which was carried out by AC. König and me conjointly. The manuscript was written by me. A.-V. Bohne and J. Nickelsen designed the research and supervised the project.

Manuscript: Analysis of the plastid RNA-sensitive DLA2 containing complex in *Chlamydomonas reinhardtii*

Laura Kleinknecht, Ann-Christine König, Alexandra-Viola Bohne, Jörg Nickelsen

Introduction

It is becoming more and more evident that proteins especially enzymes can have an alternative function additional to the one which was primary described. It is believed that while these were uni-functional proteins to begin with, throughout time an additional function was acquired in different ways (Jeffery, 2003). In 1988 the first such proteins, crystallins in the human eyes, were described (Piatigorsky et al., 1988). Soon after that first publication even more proteins with two independent functions were discovered and the term moonlighting proteins was introduced, referring to workers working in two jobs whereas the second additional job would often happen at night thereby in moonlight (Jeffery, 1999; Huberts and van der Klei, 2010).

In 2002 the RNA-binding protein RBP63 was described by Ossenbühl et al. (2002). This protein, with an apparent size of 63 kDa, was discovered in a screen for RNA-binding proteins associated with the thylakoid membrane of the unicellular green algae *Chlamydomonas reinhardtii*. RBP63 preferably binds to the 5'UTR of *psbA* mRNA, that encodes the D1 subunit of photosystem II (Ossenbühl et al., 2002). In a following study by Bohne et al. (2013), it became evident that RBP63 is actually the previously described protein DLA2 (Dihydrolipoamidacetyltranserase), the E2 subunit of the chloroplast pyruvate dehydrogenase complex (cpPDC) (Reid et al., 1977; Mooney et al., 1999). This multi-enzyme complex catalyzes the oxidative decarboxylation of pyruvate in the chloroplast stroma thereby providing acetyl-CoA for fatty acid synthesis (Mooney et al., 2002; Lin and Oliver, 2008).

Strikingly DLA2 plays an additional role in the chloroplast of *C. reinhardtii* when they cells are grown in the presence of light and acetate. It was shown by gel filtration analysis that it is part of an RNase-sensitive complex with a size of more than 1 MDa. Further experiments suggest that by binding to the *psbA* mRNA, DLA2 actually tethers this mRNA the so called translational active zones (T-zones) thereby promoting translation of the D1 protein in the thylakoid membrane (Bohne et al., 2013; Bohne and Nickelsen, 2017). This study gives further insight into the RNA binding mode of DLA2, the role of the cpPDC E1 and E3

subunits in the RNase-sensitive complex and suggests how the two functions of DLA2 might be regulated by lysine acetylation.

Results and discussion

E1 and E3 are not part of the DLA2-RNA-complex

To elucidate the role of the other subunits of the cpPDC in the additional function of DLA2 the size exclusion chromatography (SEC) analyses performed previously by Bohne et al (2013) were repeated using a Sephacryl S500 HR column, that is providing a better resolution in the high molecular weight range (Fig 1a). Chloroplasts from a cell-wall deficient strain were isolated and solubilized thylakoids, either treated with RNase or not, before SEC followed by immunoblots using antibodies against all cpPDC subunits was conducted. Once again a clear shift of DLA2 from fraction 5 to 8 to fraction 8 and 9 can be seen. The E1 and E3 subunits accumulate most prominent in fraction 8 to 10 and no change can be seen after the addition of RNase.

To further investigate this we additionally performed a similar experiment using a 2D-BN-PAGE to separate the complexes after solubilizing the thylakoids and RNase treatment (Fig 1b). Once again a similar pattern was observed. E1 and E3 accumulate at around 500 kDa whereas DLA2 shows a different pattern accumulating in a complex of a much higher molecular weight. After addition of RNase there is no change for E1 and E3 whereas a shift can be observed for DLA2. The DLA2 blots were divided into five fractions (Fig 1b) and signals in these fractions were quantified (Fig 1c). The amounts of DLA2 in fraction 1 and 2 were reduced by about 50 % after the addition of RNase.

Taken together these results clearly indicate that the subunits E1 and E3 are not part of the RNA-containing DLA2 complex associated with the thylakoids in the Chlamydomonas chloroplast. In the high molecular weight fractions we could not find the active cpPDC, as it should be localized mainly in the stroma. The previously measured PDC activity in these fractions was probably caused by small amounts of assembled cpPDC in the thylakoid SEC fractions that were not detectable by immunoblot (Bohne et al., 2013). To confirm this hypothesis SEC with stroma will be performed in the future. The composition of the DLA2-RNA-complex still remains elusive. Even though multiple copies of DLA2 attached to the RNA might be able to form a big complex it seems more likely that additional proteins are

part of this complex. Up to now different strategies including Co-IP and pull-downs followed by mass spectrometry have not identified clear candidates.

Accumulation and ratio of cpPDC subunits

The accumulation of the PDC subunits was examined in the available PDC subunit mutants to analyze the dependency on one another (Fig 2). We were able to obtain mutants for the E1β (*pdh2*) and the E3 (*dld2*) subunit, whereas the previously analyzed RNAi line was used for DLA2 (Bohne et al., 2013). Immunoblots were performed using these strains as well as respective wild types (Fig 2a). These blots were quantified by calculating the ratio of PDC subunits over RbcL (Fig 2b).

It can be clearly seen, that the subunits are indeed influenced by the accumulation of their partners. The highest effect on the other subunits can be seen in the *pdh2* mutant. As expected the two isoforms of E1, E1β and E1α are absent in this mutant. Strikingly DLA2 is reduced to about 20 %, whereas there is about 40% of DLD2 accumulating. The DLA2 mutant is not a knock-out mutant but an RNAi-induced knockout line. Therefore the DLA2 protein is not absent but still accumulating to about 15% of wild-type level (Bohne et al., 2013). Both E1 isoforms are reduced to roughly 30-40% whereas E3 still accumulates to about 65%. In consistency with this E1 and E2 are less influenced in the *dld2* mutant compared to their reduction in the other mutants. They accumulate to 60 and 80 %, respectively, even though E3 is completely absent.

This leads to the conclusion that the accumulation of each subunit is depending on the accumulation of the other subunits. Moreover it is interesting that the absence of DLD2 has less effect on the other subunits than the absence of PDH2 or DLA2. Accordingly, DLD2 is less affected by the absence of the other subunits. One can assume that its accumulation is not as dependent on DLA2 and PDH2 as their accumulation is on each other. We can only hypothesize why that is the case. An easy explanation would be to assume a higher stability of the unassembled protein compared to the other subunits. Another hypothesis would be to assume an additional function of the protein, therefore requiring an accumulation of DLD2 independent of the other subunits.

During our work on this topic the closely related question arose whether the ratio of the subunits changes under different conditions. It might be possible that additional DLA2 is produced under mixotrophic conditions thereby providing more DLA2 available for its RNA-binding function. Proteins from cells grown under photo-, mixo- and heterotroph conditions

as well as recombinant proteins as a reference for protein amounts were subjected to immunoblot analyses and quantified (Fig 3). The ratio of E1α:E1β:E2:E3 is roughly 1:2.5:1.8:2.2. No difference in protein accumulation under different conditions and therefore no difference in the ratio was observed. Strikingly, E1α and E1β do not accumulate to the same level as we expected beforehand (Johnston et al., 2000). The observed ratio is different from previously reported ratios of the subunits in PDC complexes, which can be explained by differences between organisms or even more likely by the fact that we observed the ratio of the total protein amounts in the cell whereas other studies identified the ratio within the assembled complex (Johnston et al., 2000). This result clearly indicates that the DLA2 pool is not increasing under mixotrophic conditions and therefore the cpPDC might at least partially disassemble to free DLA2 for its additional function as RNA-binding protein.

The RNA binding site of DLA2

To elucidate which part of the protein is responsible for RNA-binding different recombinant DLA2 deletion mutants were generated. Interestingly, i*n silico* analysis has predicted the RNA-binding site in the same location as the E3 binding site (Fig 4a, Bohne et al., 2013). It contains a number of positively charged amino acids as well as a predicted Rossmann fold, which is typically a NAD+/FAD^+ binding site (Rao and Rossmann, 1973). This motif was reported to be involved in the binding of RNA molecules by enzymes (Nagy and Rigby, 1995; Nagy et al., 2000; Benning, 2009).

Due to the highly repetitive and GC-rich sequence of the DLA2 gene it was impossible to delete only this exact part of the gene. Instead two different deletion mutants were created and cloned both with a His-tag as well was a MBP-tag, to exclude an influence of the tag on the binding activity (Fig 4a). It was shown that the recombinant MBP-tagged protein possessed the same affinity for RNA as the previously published His-tagged version of DLA2 (Bohne et al., 2013, data not shown). For the following experiments the MBP-tagged versions were expressed, but the tag was removed prior to binding studies.

The radiogram of an UV-cross-link experiment in figure 4b shows that both deletion mutants have lost their binding activity of 5' UTR *psbA* mRNA. There is no visible signal in either mutant at the expected size. To gain further insight filter binding assays were performed and quantified to generate RNA binding curves (Fig 4c). Both versions of the deletion protein show a residual RNA binding activity when high molarities of the proteins were present. Nevertheless, the activity it is strongly reduced when compared to native DLA2.

We therefore presume that the RNA binding site of DLA2 indeed overlaps with its E3 binding site. That leads directly to the hypothesis that the E3 subunit and the RNA might bind to DLA2 in a competitive mode. To test this hypothesis UV cross-linking competition experiments were performed (Fig 5). First rDLA2 and the radiolabeled *psbA* mRNA were incubated before rising amounts of rDLD2 were added and the UV-cross-link assay was performed. Increasing amounts of rDLD2 in the sample clearly lead to decrease in RNA binding activity of rDLA2. As a control purified GST was used, which did not affect the RNA binding activity of rDLA2 (Fig S1). In conclusion it seems very likely that the E3 binding site of DLA2 is not only overlapping with the RNA binding site, but moreover that there is a competitive binding of either RNA or DLD2.

Acetylation of DLA2 influences its binding activity

Several studies have shown that besides regulation on RNA level a high number of protein functions is regulated by post-translational modifications (PTMs). PTMs, that occur either at the protein's N- or C-termini or on the amino acid side chains, include most prominently phosphorylation as well as glycosylation, carbonylation and acetylation (Khoury et al., 2011). They are often used as fast molecular switches to change activity or function of proteins (Jing et al., 2013; Castano-Cerezo et al., 2014).

In the case of DLA2 a regulation mechanism through acetylation seemed very likely. Lysine acetylation was described in various organisms including bacteria, yeast, plant and animal cells and requires acetyl-CoA as a substrate (Choudhary et al., 2009; Finkemeier et al., 2011; Weinert et al., 2011; Henriksen et al., 2012; Lundby et al., 2012; Melo-Braga et al., 2012; Choudhary et al., 2014; Mo et al., 2015). The RNA binding activity of DLA2 occurs under only mixotrophic conditions in the presence of light and acetate, which can be converted to acetyl-CoA by two mechanisms (Spalding, 2009). Furthermore, DLA2 carries three lysine acetylation sites including one that lies within the proposed RNA binding site. Interestingly, the acetylation of this specific lysine residue is significantly upregulated under mixo- and heterotrophic conditions when compared to photoautotrophic conditions (König et al., in preparation).

In order to test whether acetylation plays a role in the regulation of DLA2 we performed an *in vitro* auto-acetylation assay with recombinant DLA2 protein. The protein was used untreated, treated with a deacetylase (YHDZ) or incubated in the presence of acetyl-CoA before an immunoblot was performed (Fig 6a). In the sample that was incubated with acetyl-

CoA a clear shift in size can be seen in the ponceau stain when compared to the two other samples (Fig 6a, upper panel). Accordingly this band was also detected using an anti-Ac Lysin antibody for the immunoblot (Fig 6a, lower panel).

The acetylated and deacetylated rDLA2 was also used to perform an UV-cross-linking assay as described before (Fig 6b). The deacetylated protein showed a similar RNA binding activity compared to the untreated rDLA2. Strikingly the RNA binding activity of the acetylated rDLA2 was completely eradicated. This result indicates that the RNA binding activity of DLA2 can be influenced by the acetylation status of DLA2 and that this mechanism could therefore play a role in the regulation of its dual function (Sterner and Berger, 2000). Since the protein was randomly auto-acetylated, it is likely that all available lysine sites were acetylated. In future one has to mutagenize the individual lysine sites and repeat the RNA binding experiment to reveal the role of lysine acetylation in the regulation of DLA2.

Summing up we found that DLA2 is truly part of two different distinct complexes whereas one of these is the active cpPDC including the E1 and the E3 subunits while the other does not contain them but presumably the *psbA* mRNA and additional still unknown factors. Since there is competitive binding of RNA and DLD2 to rDLA2 it seems likely that DLA2 is regulated by a PTM more probable than not by acetylation.

Material and Methods

Algal Strains and Culture Conditions

To allow easy chloroplast isolation we used CC-406, which has a defective cell wall, as the wild-type *C. reinhardtii* strain. The DLA2 RNAi line (*iDLA2-1*) and the respective control that was transformed with an empty vector (WT-Ne) were generated by Bohne et al. (2013). The other cpPDC mutant lines (*pdh2* and *dld2*) were crossed with a cw15 strain to allow for better comparison with the wild type (Dent et al., 2005; Li et al., 2016). Strains were maintained on tris-acetate-phosphate (TAP) plates (containing 0,8% agar) at 23°C and under 30 $\mu E/m^{-2}/s^{-1}$ continuous light (Gorman and Levine, 1965). Liquid cultures were grown at 23 °C in TAP medium containing 1% sorbitol for mixotrophic and heterotrophic conditions and in high-salt minimal (HSM) medium for photoautotrophic growth (Harris, 1989). Light conditions were 30 $\mu E/m^{-2}/s^{-1}$ for mixotrophic growth, 200 $\mu E/m^{-2}/s^{-1}$for phototrophic growth and darkness for heterotrophic growth.

Chloroplast Preparation, SEC and 2D-BN-PAGE

Chloroplast preparation and RNase treatment was performed as described before (Bohne et al., 2013). Gel filtration samples were loaded through an online filter onto a Sephacryl S500 HR column (GE Healthcare), and elution was performed at 4°C with a buffer containing 50 mM KCl, 2.5 mM EDTA, 5 mM ε-aminocaproic acid, 0.1% Triton X-100, and 20 mM tricine-KOH, pH 7.8, at a flow of 0.3 mL/min using an ÄKTApurifier 10 system (GE Healthcare). Aliquots of each elution fraction were subjected to immunoblotting.

2D-BN-PAGE samples were mixed with an equal amount of loading buffer before a 2-BN-PAGE was performed as described previously (Schottkowski et al., 2009). Second dimensions were subjected to immunoblotting.

Expression of recombinant proteins

Recombinant DLA2 with His-tag was expressed as previously described by Bohne et al. (2013). The same sequence was inserted into the plasmid pMAL-c5x (NewEnglandBiolabs). The DLA2 deletion mutants were cloned into the plasmid pMAL-c5x as well (rDLA2ΔE3-1: Δaa136-aa241; rDLA2ΔE3-2: Δaa208-aa261).

Expression was performed in *Escherichia coli* Rosetta cells (Novagen) by induction with 1 mM isopropyl b-D-1-thiogalactopyranoside (IPTG) for 3 h at 30°C. Purification of the recombinant protein was performed according to the New England Biolabs protocol for purification of MBP-tagged recombinant proteins, including the removal of the MBP tag by proteolytic digestion with factor Xa.

The other cpPDC subunits were all cloned into the pGEX-4T-1 plasmid (Amersham Pharmacia Biotech) and purified according to the manufacturers protocol. For PDC2 amino acids 35 to 397 were inserted and overexpression took place at 18°C over night after induction with 1 mM IPTG. For PDH2 amino acids 33 to 370 were inserted and overexpression took place at 27°C over night after induction with 1 mM IPTG. For DLD2 amino acids 241 to 585 were inserted and overexpression took place at 18°C over night after induction with 0.5 mM IPTG.

In Vitro Synthesis of RNA and UV Cross-Linking

In vitro synthesis of RNA and UV cross-linking experiments were basically performed as described by Zerges and Rochaix (1998) and Bohne et al. (2013). The DNA template for the *in vitro* synthesis of the *psbA* 5' UTR RNA probe was generated by PCR using the

following primers: T7psbA5 (*5′-gtaatacgactcactatagggTACCATGCTTTTAATAGAAG-3′***)** and 2054-psbA **(***5′-GATCCATGG TCATATGTTAATTTTTTTAAAG-3′***).**

Binding reactions (20 µl) were performed at RT for 10 min and contained 20 mM HEPES/KOH, pH 7.8, 5 mM $MgCl_2$, 60 mM KCl, and 100 ng of protein unless indicated otherwise. Each reaction contained 100 kcpm of ^{32}P-RNA probe. For competition experiments indicated amounts of competitor protein were added to the reaction. Radiolabeled RNA and DLA2 were mixed prior to the addition of competing proteins.

The K_D was determined as described by Ostersetzer et al. (2005). Increasing amounts of recombinant DLA2 protein and the two rDLA deletion versions were incubated for 15 min at RT with *in vitro* transcribed ^{32}P-labeled *psbA* mRNA (6.7 pM) in 20 µl reactions in the same binding buffer used for UV cross-linking assays. Subsequently, the reactions were filtered through stacked nitrocellulose (Reprobe nitrocellulose plus, 0.45 µm; Applichem) and nylon membranes (Nylon plus, 0.45 µm; C. Roth) using a dot blot apparatus (Minifold SRC96, Schleicher & Schuell). The membranes were washed once with 100 µl of binding buffer, dried, and subjected to phosphorimaging and quantitation with AlphaEase software (Alpha Innotech Corporation).

Acetylation Assay

Recombinant His-DLA2 was left untreated or incubated with either 100 ng Deacetylase (YHDZ) or 10 mM Acetyl-CoA at room temperature for two hours. Hereafter a ponceau S staining (Sigma) followed by immunoblotting with an anti-acetyl-lysine antibody or an UV-cross-link assay was performed as described above.

References

Benning, C. (2009). Mechanisms of lipid transport involved in organelle biogenesis in plant cells. Annual review of cell and developmental biology *25*, 71-91.

Bohne, A.-V., and Nickelsen, J. (2017). Metabolic Control of Chloroplast Gene Expression: An Emerging Theme. Mol Plant *10*, 1-3.

Bohne, A.-V., Schwarz, C., Schottkowski, M., Lidschreiber, M., Piotrowski, M., Zerges, W., and Nickelsen, J. (2013). Reciprocal Regulation of Protein Synthesis and Carbon Metabolism for Thylakoid Membrane Biogenesis. PLoS Biol *11*, e1001482.

Castano-Cerezo, S., Bernal, V., Post, H., Fuhrer, T., Cappadona, S., Sanchez-Diaz, N.C., Sauer, U., Heck, A.J., Altelaar, A.F., and Canovas, M. (2014). Protein acetylation

affects acetate metabolism, motility and acid stress response in Escherichia coli. Mol Syst Biol *10*, 762.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.

Choudhary, C., Weinert, B.T., Nishida, Y., Verdin, E., and Mann, M. (2014). The growing landscape of lysine acetylation links metabolism and cell signalling. Nat Rev Mol Cell Biol *15*, 536-550.

Dent, R.M., Haglund, C.M., Chin, B.L., Kobayashi, M.C., and Niyogi, K.K. (2005). Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of Chlamydomonas reinhardtii. Plant Physiol *137*, 545-556.

Finkemeier, I., Laxa, M., Miguet, L., Howden, A.J., and Sweetlove, L.J. (2011). Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. Plant Physiol *155*, 1779-1790.

Gorman, D.S., and Levine, R.P. (1965). Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of *Chlamydomonas reinhardi*. Proc Natl Acad Sci USA *54*, 1665-1669.

Harris, E.H. (1989). The Chlamydomonas Sourcebook. A Comprehensive Guide to Biology and Laboratory Use. , Vol 246, 1989/12/15 edn (San Diego, CA: Academic Press).

Henriksen, P., Wagner, S.A., Weinert, B.T., Sharma, S., Bacinskaja, G., Rehman, M., Juffer, A.H., Walther, T.C., Lisby, M., and Choudhary, C. (2012). Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in Saccharomyces cerevisiae. Mol Cell Proteomics *11*, 1510-1522.

Huberts, D.H.E.W., and van der Klei, I.J. (2010). Moonlighting proteins: An intriguing mode of multitasking. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research *1803*, 520-525.

Jeffery, C.J. (1999). Moonlighting proteins. Trends in biochemical sciences *24*, 8-11.

Jeffery, C.J. (2003). Moonlighting proteins: old proteins learning new tricks. Trends in genetics *19*, 415-417.

Jing, E., O'Neill, B.T., Rardin, M.J., Kleinridders, A., Ilkeyeva, O.R., Ussar, S., Bain, J.R., Lee, K.Y., Verdin, E.M., Newgard, C.B., *et al.* (2013). Sirt3 regulates metabolic flexibility of skeletal muscle through reversible enzymatic deacetylation. Diabetes *62*, 3404-3417.

Johnston, M.L., Miernyk, J.A., and Randall, D.D. (2000). Import, processing, and assembly of the alpha- and beta-subunits of chloroplast pyruvate dehydrogenase. Planta *211*, 72-76.

Khoury, G.A., Baliban, R.C., and Floudas, C.A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. Scientific Reports *1*, 90.

Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., Yue, R., Robertson, J.M., Lefebvre, P., Fitz-Gibbon, S.T., *et al.* (2016). An indexed, mapped mutant library enables reverse genetics studies of biological processes in Chlamydomonas reinhardtii. Plant Cell *28*, 367-87

Lin, M., and Oliver, D.J. (2008). The Role of Acetyl-Coenzyme A Synthetase in Arabidopsis. Plant Physiology *147*, 1822.

Lundby, A., Lage, K., Weinert, B.T., Bekker-Jensen, D.B., Secher, A., Skovgaard, T., Kelstrup, C.D., Dmytriyev, A., Choudhary, C., Lundby, C., *et al.* (2012). Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. Cell Rep *2*, 419-431.

Melo-Braga, M.N., Verano-Braga, T., Leon, I.R., Antonacci, D., Nogueira, F.C., Thelen, J.J., Larsen, M.R., and Palmisano, G. (2012). Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (Vitis vinifera) mesocarp and exocarp owing to Lobesia botrana infection. Mol Cell Proteomics *11*, 945-956.

Mo, R., Yang, M., Chen, Z., Cheng, Z., Yi, X., Li, C., He, C., Xiong, Q., Chen, H., Wang, Q., *et al.* (2015). Acetylome analysis reveals the involvement of lysine acetylation in photosynthesis and carbon metabolism in the model cyanobacterium Synechocystis sp. PCC 6803. J Proteome Res *14*, 1275-1286.

Mooney, B.P., Miernyk, J.A., and Randall, D.D. (1999). Cloning and Characterization of the DihydrolipoamideS-Acetyltransferase Subunit of the Plastid Pyruvate Dehydrogenase Complex (E2) from Arabidopsis. Plant Physiology *120*, 443-452.

Mooney, B.P., Miernyk, J.A., and Randall, D.D. (2002). THE COMPLEX FATE OF α-KETOACIDS. Annual Review of Plant Biology *53*, 357-375.

Nagy, E., Henics, T., Eckert, M., Miseta, A., Lightowlers, R.N., and Kellermayer, M. (2000). Identification of the NAD(+)-binding fold of glyceraldehyde-3-phosphate dehydrogenase as a novel RNA-binding domain. Biochem Biophys Res Commun *275*, 253-260.

Nagy, E., and Rigby, W.F. (1995). Glyceraldehyde-3-phosphate dehydrogenase selectively binds AU-rich RNA in the NAD(+)-binding region (Rossmann fold). J Biol Chem *270*, 2755-2763.

Ossenbühl, F., Hartmann, K., and Nickelsen, J. (2002). A chloroplast RNA binding protein from stromal thylakoid membranes specifically binds to the 5′ untranslated region of the psbA mRNA. FEBS J *269*, 3912-3919.

Ostersetzer, O., Cooke, A.M., Watkins, K.P., and Barkan, A. (2005). CRS1, a chloroplast group II intron splicing factor, promotes intron folding through specific interactions with two intron domains. Plant Cell *17*, 241-255.

Piatigorsky, J., O'Brien, W.E., Norman, B.L., Kalumuck, K., Wistow, G.J., Borras, T., Nickerson, J.M., and Wawrousek, E.F. (1988). Gene sharing by delta-crystallin and argininosuccinate lyase. Proc Natl Acad Sci USA *85*, 3479-3483.

Rao, S.T., and Rossmann, M.G. (1973). Comparison of super-secondary structures in proteins. J Mol Biol *76*, 241-256.

Reid, E.E., Thompson, P., Lyttle, C.R., and Dennis, D.T. (1977). Pyruvate Dehydrogenase Complex from Higher Plant Mitochondria and Proplastids. Plant Phy *59*, 842-848.

Schottkowski, M., Gkalympoudis, S., Tzekova, N., Stelljes, C., Schünemann, D., Ankele, E., and Nickelsen, J. (2009). Interaction of the Periplasmic PratA Factor and the PsbA (D1) Protein during Biogenesis of Photosystem II in Synechocystis sp. PCC 6803. J Biol Chem *284*, 1813-1819.

Spalding, M.H. (2009). The CO2-Concentrating Mechanism and Carbon Assimilation. In: D.B. Stern, and G.B. Witman, eds. The Chlamydomonas Sourcebook (Second Edition). London, Academic Press, 257-301.

Sterner, D.E., and Berger, S.L. (2000). Acetylation of Histones and Transcription-Related Factors. Microbiol Mol Biol Rev *64*, 435-459.

Weinert, B.T., Wagner, S.A., Horn, H., Henriksen, P., Liu, W.R., Olsen, J.V., Jensen, L.J., and Choudhary, C. (2011). Proteome-wide mapping of the Drosophila acetylome demonstrates a high degree of conservation of lysine acetylation. Sci Signal *4*, ra48.

Zerges, W., and Rochaix, J.-D. (1998). Low Density Membranes Are Associated with RNA-binding Proteins and Thylakoids in the Chloroplast of Chlamydomonas reinhardtii. J Cell Biol *140*, 101-110.

FIGURE 1

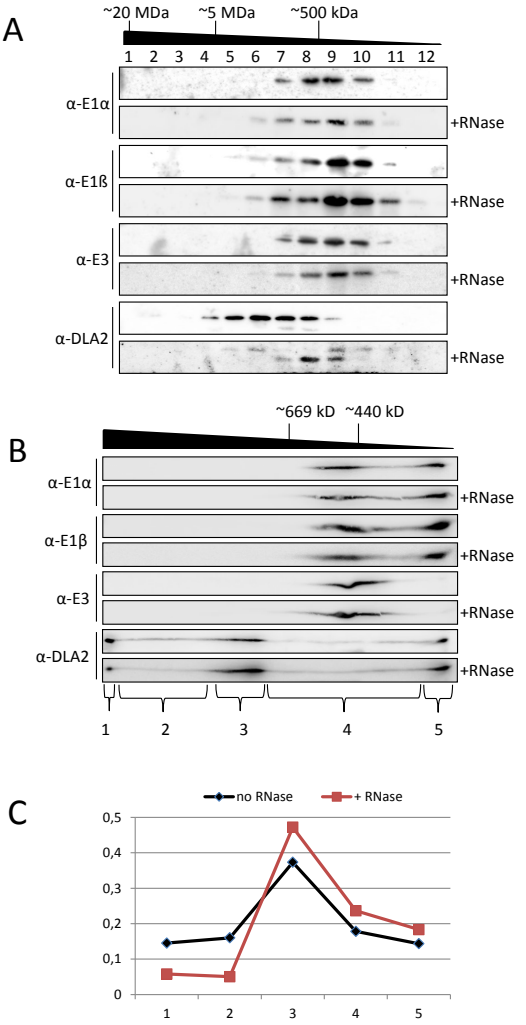


Figure 1: RNase-sensitive complex.
A) Gel filtration analysis. Thylakoids from the wild-type strain CC-406 were solubilized and either treated with RNase or loaded directly on the column without prior RNase treatment. Fractions were collected and 1/10th of each fraction was loaded on SDS-gels. Subsequently immunoblots were performed using antibodies against all PDC subunits. Molecular masses shown at the top were estimated by parallel analysis of high molecular mass calibration markers.
B) 2D-BN-PAGE. Thylakoids from the wild-type strain CC-406 were solubilized and either treated with RNase or incubated on ice for the same time without the addition of RNase before 2D-BN-PAGE was performed. Second dimensions were subjected to immunoblot analysis.
C) Quantification of DLA2. The DLA2 blots from the 2D-BN-PAGE were divided into different fractions (B) and quantified using the program ImageQuantTL. The sum of the signal in all fractions was set to 100%.

FIGURE 2

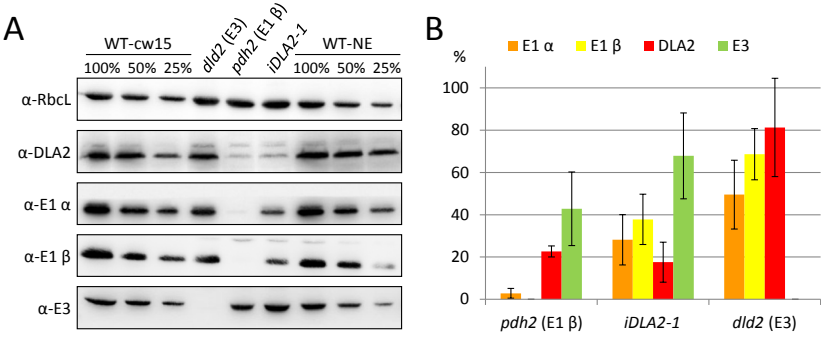


Figure 2: Quantification of PDC subunits in mutant backgrounds.
A) Protein samples of wild type, the different PDC mutant strains and WT-NE strain as a control for the iDLA2-2 strain were separated by SDS-PAGE and subjected to immunoblot analysis.
B) Quantification of immunoblot signals of three independent experiments. Quantification was performed by calculating the ratio of PDC subunits over RbcL. Values obtained for respective wild type were set to 100%.

FIGURE 3

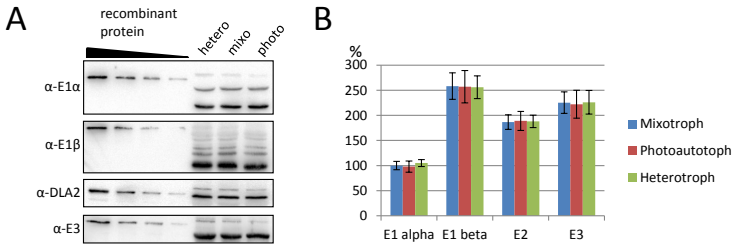


Figure 3: Quantification of PDC subunits under different growth conditions.
A) Protein fractions (15 µg) of wild-type grown under heterotrophic, mixotrophic or photoautotrophic conditions were separated by SDS-PAGE and subjected to immunoblot analysis. Specific amounts (50, 25, 15, 10 ng) of the respective recombinant proteins were loaded on the gels for means of quantification.
B) Quantification of immunoblot signals of three independent experiments. Quantification was performed by comparing the amounts in moles of protein with the respective defined amount of recombinant proteins. The value obtained for E1α under heterotrophic growth conditions was set to 100%.

FIGURE 4

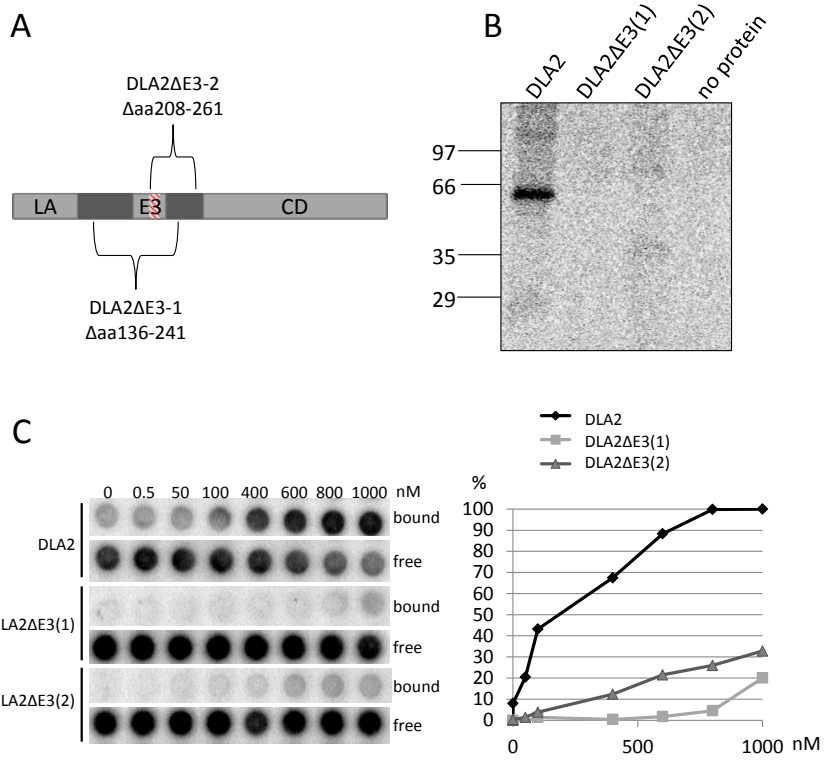


Figure 4: RNA-binding domain of DLA2.

A) Scheme of the mature DLA2 protein. Identified domains are marked by light grey boxes (LA = lipoyl attachement site, E3 = E3 binding site/predicted RNA binding site, CD = catalytic domain). The predicted Rossmann fold in the E3 binding site is symbolized by red stripes. Deleted parts in DLA2ΔE3-1 and DLA2ΔE3-2 are marked by brackets.

B) UV-cross-link experiment with different versions of DLA2 containing deletions of the putative RNA binding domain and the *psbA* 5' UTR mRNA. Deleted parts are depicted in (A). 100 ng of each protein and 100 kcpm of radiolabeled *psbA* 5' UTR mRNA were used.

C) Filter binding assays with different versions of DLA2 containing deletions of the putative RNA binding domain were performed to determine RNA binding curves. Binding reactions containing 6.7 pM ^{32}P-labeled *psbA* 5' UTR RNA and indicated molarities of the different DLA2 versions were filtered through stacked nitrocellulose and nylon membranes (left panel) using a dot blot apparatus according to Ostersetzer et al. (2005). Signal intensities of nitrocellulose-bound protein–RNA complexes (bound) as well as nylon membrane-bound free RNAs (free) were measured for quantification. The amount of bound RNA in the 800 nM DLA2 probe was set to 100 %.

FIGURE 5

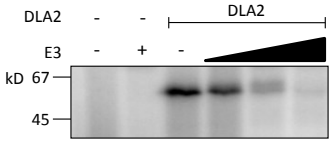


Figure 5: Competitive binding of *psbA* mRNA and rE3 to rDLA2.
We incubated 100 ng of DLA2 with 100 kcpm of radiolabeled *psbA* 5' UTR mRNA and rising amounts of E3 (1x, 5x and 10x molar excess over DLA2) before a standard UV-cross-link experiment was performed.

FIGURE 6

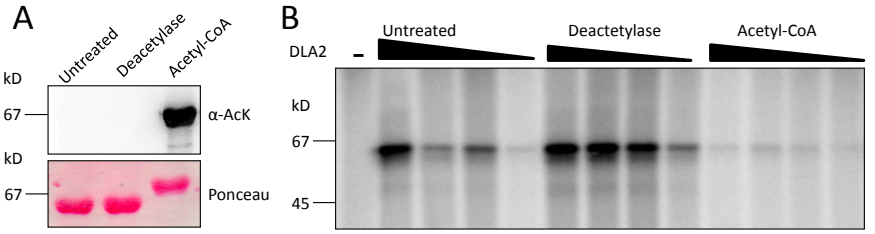
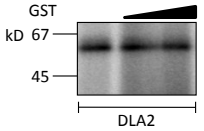


Figure 6: Acetylation influences RNA binding capacity of rDLA2.
A) Untreated, deacetylated and acetylated rDLA2 was run on a SDS-PAGE followed by an immunoblot. Prior to incubation with the anti-acetyl-lysine antibody (α-AcK) the membrane was stained with ponceau (lower panel).
B) UV-cross-link experiment with decreasing amounts of untreated, deacetylated and acetylated rDLA2 protein (1000 ng, 100 ng, 10 ng and 1 ng) and 100 kcpm of ^{32}-P labeled *psbA* 5' UTR mRNA.

Supplemental
FIGURE 1



Supplemental Figure 1: Control UV-Crosslink experiment
We incubated 100 ng of DLA2 with 100 kcpm of radiolabeled *psbA* 5' UTR mRNA and rising amounts of GST (5x and 10x molar excess over DLA2) before a standard UV-cross-link experiment was performed.

3.3 Unanticipated T7 RNA polymerase activity using annealed oligonucleotides as transcription template

Kleinknecht, L., Nickelsen, J., and Bohne, A.-V. (2017). Endocytobiosis and Cell Research 28, 33-37.

The bacteriophage T7 RNA polymerase (RNAP) is regularly used in molecular biology to efficiently synthesize specific RNA sequences. For RNA synthesis a DNA template, including a specific double stranded short promotor element, is required. During ongoing work we noticed an artificial activity in the absence of a correct promoter element. Non-complementary primers showed a self-annealing capacity that led to the synthesis of RNA in the absence of a perfectly double-stranded promotor element. Moreover, even some perfectly self-complementary primers showed a high tendency to self-anneal. RNase T1 digestion of the generated transcripts was performed, showing indeed that transcription took place using only one of the self-annealed primers instead of the correct template. In conclusion, we uncovered an unanticipated T7 RNAP activity and feel that experiments involving this method should be conducted with caution. Ideally, a possible self-annealing potential of primers should be tested, before using this method.

All experiments in this study were performed by me. The manuscript was written by A.-V. Bohne and me and revised by J. Nickelsen.

Endocytobiosis and Cell Research (2017) **VOL 28** | Issue 1 | 33–37 | International Society of Endocytobiology
zs.thulb.uni-jena.de/content/main/journals/ecb/info.xml

Endocytobiosis and Cell Research

Unanticipated T7 RNA polymerase activity using annealed oligonucleotides as transcription template

Laura Kleinknecht, Jörg Nickelsen and Alexandra-Viola Bohne*

Ludwig-Maximilians University Munich, Molecular Plant Sciences, 82152 Planegg-Martinsried, Germany; *correspondence to: Alexandra-Viola Bohne; Tel.: +49 89 218074774; E-mail: Alexandra.Bohne@lmu.de; ORCID: 0000–0001–9638–0762

The bacteriophage T7 RNA polymerase (RNAP) is an important tool used in molecular biology to efficiently synthesize specific RNA sequences for diverse *in vitro* applications. Its activity depends on a DNA template that must contain a specific double-stranded short promoter element. During ongoing work, we noticed an unanticipated activity of the T7 RNAP, which was able to generate RNAs even in the absence of a known T7 promoter motif. Here, we describe the observed artificial activity of the T7 RNAP and attempt to uncover the underlying mechanism.

Endocytobiosis and Cell Research (2017) **28**(1):33–37
Category: Communication
Keywords: T7 RNA polymerase, *in vitro* transcription, RNA synthesis, T7 promoter

Accepted: June 07, 2017
Published online: June 09, 2017

Introduction

The bacteriophage T7 DNA-dependent RNAP efficiently synthesizes RNA molecules using a DNA template strand. It is a monomeric enzyme that has been cloned and can be expressed in bacteria (Arnaud et al. 1997; Davanloo et al. 1984; Tabor and Richardson 1985). It does not require any auxiliary factors for its activity, therefore making it extremely useful for *in vitro* synthesis of RNA (Chamberlin and Ryan 1982). T7 RNAP promoters share a 23 base pair consensus sequence which is positioned from −17 to +6 relative to the transcription initiation site at +1 (Dunn et al. 1983). The T7 RNAP recognizes specific nucleotides of this double-stranded promoter, contacting both the template and the non-template strand (Ikeda and Richardson 1986; Oakley and Coleman 1977; Sousa et al. 1993). The enzyme builds an unstable initiation complex producing short RNA products of 8–12 bps after which the polymerase undergoes a conformational change and thereby starts the highly efficient elongation phase (Cheetham et al. 1999; Steitz 2009). The termination occurs either by a terminator sequence or by reaching the 5' end of a linear template (Landick 1997).

It was already shown more than 30 years ago that the T7 RNAP can produce high amounts of RNA from synthetic DNA templates *in vitro* and requires only a short double stranded promoter region, whereas the downstream region may be single stranded (Milligan et al. 1987; Oakley and Coleman 1977). As of today, there are at least four possible template options for *in vitro* transcription. When the method was established, genes were cloned into plasmids containing a T7 promoter element (Studier and Moffatt 1986). For transcription the plasmid was linearized to ensure the termination of the RNA. This method was largely replaced by less time-consuming methods that do not require the cloning of genes namely by using cDNAs, PCR products or annealed synthetic DNA oligonucleotides as DNA templates for *in vitro* transcription reactions. The best choice in order to synthesize short stretches of RNA are usually annealed DNA oligonucleotides (for examples see Ruwe and Schmitz-Linneweber 2011; Williams-Carrier et al. 2008).

Here, we report on a so far unknown artificial activity of the T7 RNAP, which appears to be able to generate RNAs even in the absence of an obvious promoter element.

Materials and Methods

Annealing of oligonucleotides

1 µg of primers (Table 1) was incubated individually or together in 10 µl hybridization buffer (150 mM KCl, 10 mM Tris-HCl pH 8.3, 1 mM EDTA) at 95°C for 3 min and either transferred directly on ice (Non-annealed) or annealed by cooling them down to 25°C over the course of 90 min (Annealed). To analyze the annealing efficiency, the oligonucleotides were then separated on native 15% polyacrylamide gels and stained with ethidium bromide.

Table 1: Sequences of oligonucleotides used for the generation of RNAs. The T7 RNA polymerase promoter sequence is indicated by capitalized letters. Note that primer 1 is identical to primer 5 except for three additional nucleotides at the 5' end of primer 5.

Primer	Sequence 5' -> 3'
1	**TAATACGACTCACTATAGGG**tcattccaagtcgtggcttgtatccatgcgcttcatattc
2	attatgctgagtgatatcccagtaaggttcagtattgaacataggtacgcgaagtataag
3	**TAATACGACTCACTATAGGG**cagatgcttcttccttcgatattcattacgttgatactta
4	attatgctgagtgatatcccgtctacgaagaaggaagctataagtaatgcaactatgaat
5	atg**TAATACGACTCACTATAGGG**tcattccaagtcataacttgtatccatgcgcttcatattc
6	gaatatgaagcgcatggatacaagttatgacttggaatga**CCCTATAGTGAGTCGTATTa**cat
7	atg**TAATACGACTCACTATAGGG**attcatagttgcattacttatagcttccttcttcgtagac
8	gtctacgaagaaggaagctataagtaatgcaactatgaat**CCCTATAGTGAGTCGTATTA**cat

In vitro transcription

In vitro transcription was carried out with annealed oligonucleotides using standard T7 RNA polymerase (Promega) according to the manufacturer's protocol. Loading dye (90% (v/v) deionized formamide, 20 mM Tris-HCl pH 7.5, 20 mM EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol) was added to the samples before fractionation on 15% denaturing polyacrylamide gels. Gels were analyzed by phosphor imaging.

RNase T1 digestion

Generated RNAs at a size of ~ 40 nt were excised from polyacrylamide gels and extracted at 4°C over night by the addition of 200 µl elution buffer (0.5 M NH_4OAc, 0.25% SDS, 1 mM EDTA). The gel-purified RNAs (~ 500 cpm) were digested with 20 U RNase T1 (Thermo Scientific) at 37°C for 45 min. Subsequently, RNAs were run on a 17.5% polyacrylamide sequencing gel to compare fragment patterns. Gels were dried before being analyzed by phosphor imaging.

Results and Discussion

During our ongoing work on the molecular analysis of chloroplast RNA metabolism, we used the above mentioned method of primer annealing to synthesize radiolabeled RNAs by *in vitro* transcription. Interestingly, we realized that even an annealing of non-complementary primers – which cannot form the described double stranded T7 promoter – led repetitively to the synthesis of RNA by the T7 RNAP. More precisely, we used annealed primers which represent only the non-template strand of the T7 promoter sequence (compare Figure 1d) but not the template strand (see Table 1 for primer sequences). Next we set out to uncover the underlying mechanism. As to current knowledge the prerequisite of T7 transcription is the formation of a double stranded promoter sequence, we intended to obtain indications for potentially formed secondary structures by these non-complementary primers. First, we analyzed the primers' self-annealing capacity. Therefore, two non-complementary primer pairs (primers 1 and 2 or 3 and 4, respectively), where only primers 1 and 3 exhibited the T7 promoter sequence, were subjected individually or together to the above described annealing procedure. To promote the annealing reaction primers were boiled and gradually cooled down (Annealed) or as a control immediately incubated on ice after boiling to minimize the formation of non-covalent bonds and prevent hybridization (Non-annealed). Subsequently, the migration behavior of primers was compared by native gel electrophoresis. By examining primers 1 and 2, it became obvious that the annealed primer 1 migrates slower than the non-annealed primer 1 which indicates a self-hybridization capacity (Figure 1a, left panel). In contrary, no self-hybridization was observed for primer 2. When primers 1 and 2 were added to the annealing reaction, two bands at the size of self-annealed primer 1 and non-annealed primer 2 were detected. This suggests that both oligonucleotides did not hybridize with each other.

The analysis of the second pair of primers implied an even higher tendency of the non-template strand containing primer 3 to self-anneal. Even immediate incubation on ice after boiling did not change its running behavior on native gels as compared to the annealed primer (Figure 1a, right panel). In contrast, primer 4 did not show self-hybridization. Once more, bands of the sizes of individual primers 3 and 4 were detected in the lane where both primers were annealed suggesting again that they did not hybridize with each other.

When annealed primers 1 and 2 or 3 and 4, respectively, were provided as templates for *in vitro* transcription by the T7 RNAP, we observed an RNA synthesis yielding transcripts with the expected size of ~ 40 nt even though no known T7 promoter could be generated (Figure 1b). Strikingly, the transcript patterns of annealed primer pairs resembled those obtained from transcription of only primer 1 or 3. Conversely, when only the annealed primer 2 or 4, which did not show a self-annealing capacity, was provided as templates for the T7 RNAP no transcripts of expected sizes were detected.

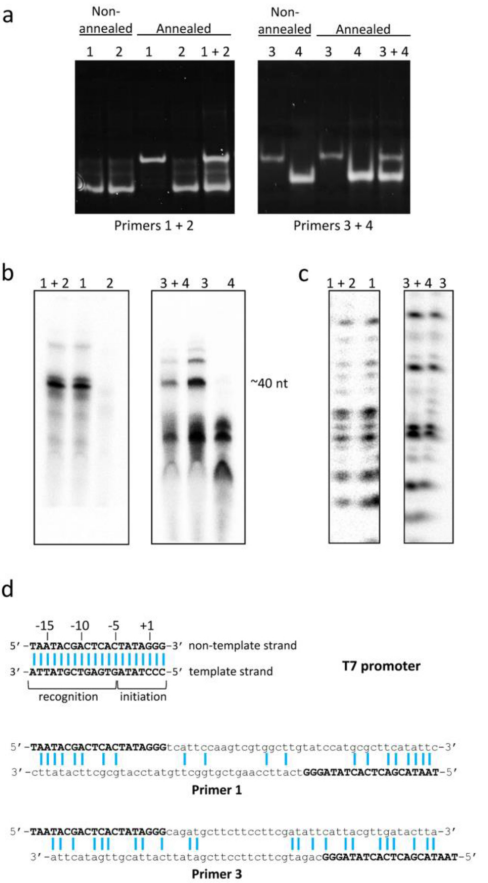


Figure 1: (a) Self-annealing capacity of oligonucleotides. Primers 1 and 2 as well as 3 and 4 (Table 1) were incubated individually or together at 95°C and either transferred directly on ice (Non-annealed) or cooled down gradually (Annealed). Subsequently, they were run on native 15% polyacrylamide gels and stained with ethidium bromide after the run; **(b) Generation of RNAs.** *In vitro* transcription was carried out with annealed oligonucleotides shown in (a) as indicated. Samples were fractionated on 15% denaturing polyacrylamide gels and analyzed by phosphor imaging. Expected sizes of transcripts were estimated from the running front of xylene cyanol (~ 40 nt); **(c) RNase T1 digestion.** Generated RNAs migrating at ~ 40 nt (b) were gel purified, digested with RNase T1, and separated on a 17.5% polyacrylamide sequencing gel to compare fragment patterns. Gels were analyzed by phosphor imaging; **(d) Prediction of possible self-annealing capacity of primers.** Oligonucleotides 1 and 3 were analyzed for possible self-annealing by using the self-dimer prediction tool of OligoAnalyzer 3.1 on the Integrated DNA Technologies webpage (http://eu.idtdna.com/calc/analyzer) with focus on generation of a double stranded T7 promoter sequence (indicated by capitalized letters). For better comparison the T7 promoter sequence is given on top. Positions are labeled with respect to the transcription initiation site at +1. Regions important for recognition by the RNAP and initiation of transcription are indicated.

To obtain indications on the nature of synthesized RNAs, transcripts at ~ 40 nt generated from only primer 1 or 3 or annealed primers were excised from the gel and incubated with RNase T1 to compare their digestion patterns. In both cases the pattern generated with only primer 1 or 3 was exactly the same as when having annealed both primers (Figure 1c). Therefore, it is very likely that the unexpected RNA products that were synthesized by the T7 RNAP originate from transcription initiation at self-annealed primers 1 or 3.

The T7 RNAP has been reported to require the double stranded promoter element for recognition and transcription initiation (Figure 1d, upper panel). Hence, it is likely that the investigated primers 1 and 3 which gave rise to RNA transcripts form at least a partially double stranded T7 promoter element. These double stranded regions might originate from primer secondary structures produced by either intermolecular or intramolecular interactions, where the primer is homologous to itself. However, as a slower migration behavior of these primers observed in native gels

(Figure 1a) is not expected from intramolecular base-pairing, intermolecular interactions were predicted *in silico* for primers 1 and 3. As shown in Figure 1d (middle and lower panel) both primers might form self-dimers, where approximately half of the nucleotides upstream of the transcription initiation site at +1 could form complementary base pairs. Notably, this appears to be sufficient for recognition by the T7 RNAP and is to our knowledge not described so far.

Remarkably, the use of primers being perfectly complementary to each other did not necessarily solve the perceived difficulties (see Table 1 for primer sequences). When carrying out the same experiments as described above with complementary primers 5 and 6 or 7 and 8, respectively, a strong self-hybridization which could not even be resolved by boiling and immediate cooling on ice was observed for all four primers tested (Figure 2a).

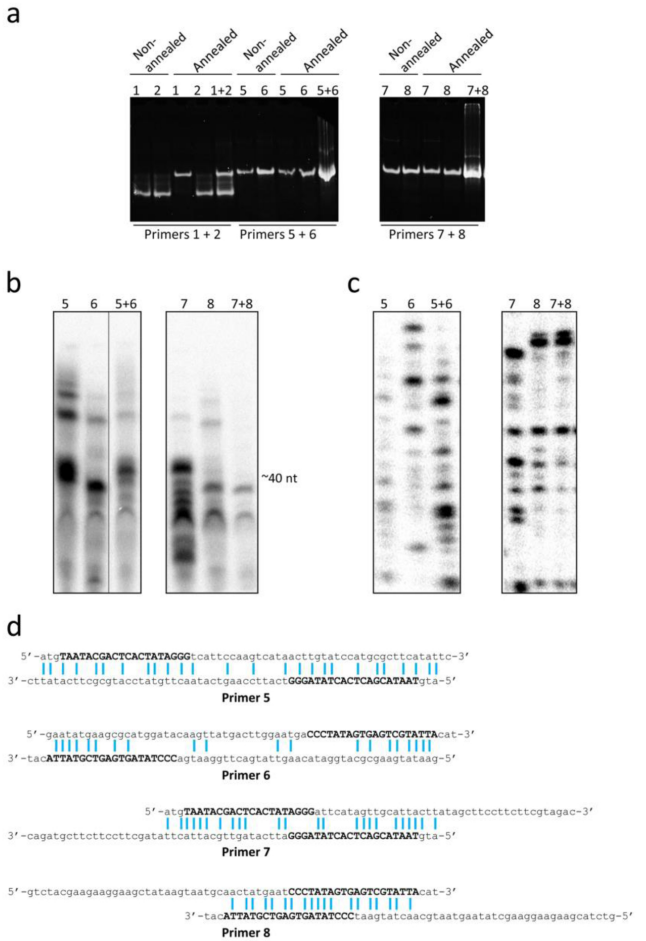


Figure 2: (a) Self-annealing capacity of oligonucleotides. Oligonucleotides 5 and 6 as well as 7 and 8 (Table 1) were annealed and analyzed as described in Figure 1a. Oligonucleotides 1 and 2 were loaded for size comparison; **(b) Generation of RNAs.** *In vitro* transcription was carried out as described in Figure 1b with annealed oligonucleotides shown in a) as indicated; **(c) RNase T1 digestion.** Generated RNAs at ~40 nt were analyzed as described in Figure 1c; **(d) Prediction of possible self-annealing capacity of primers.** Oligonucleotides 5, 6, 7 and 8 were analyzed for possible self-annealing as described in Figure 1d. The T7 promoter sequence is indicated by capitalized letters.

For all annealed oligonucleotides, whether annealed individually or together with its complementary sequence, RNA of ~40 nt was produced by T7 RNAP transcription (Figure 2b). As from this experiment no conclusions can be drawn as to whether the RNAs that were transcribed using both oligonucleotides are indeed transcribed from paired complementary oligonucleotides or individual self-annealed oligonucleotides, we again performed RNase T1 digestions

with gel-purified RNAs (Figure 2c). By comparing the digestion patterns, it became apparent that even for RNAs synthesized from complementary oligonucleotides which have undergone the annealing procedure RNase T1 patterns resembled those synthesized from either individual primer 5 or primer 8. It can therefore be concluded that the tendency for self-annealing of the analyzed primers was even higher than the tendency to anneal with the complementary primer and no defined RNAs could be synthesized. The prediction of the possible self-annealing capacity of primers 5–8 shown in Figure 2d again opens the possibility for the formation of an at least partially double stranded T7 promoter. However, whether these hybridization products really occur needs further investigation.

It was previously shown by the introduction of point mutations that not all nucleotides within the T7 promoter region are equally important for the binding of the T7 RNAP and transcription initiation (Cheetham and Steitz 2000; Ikeda et al. 1992; Rong et al. 1998; Temme et al. 2012). For example, the exchange of single nucleotides at positons –11 to –7 can abolish binding of the RNAP to the promoter, while mutations at positions –12 and –6 revealed no effect (Ikeda et al. 1992; Rong et al. 1998). On the other hand, mutations in the –3 to +2 region had no effect on promoter binding, but rather on its activity (Temme et al. 2012). However, all these mutations consider the exchange of one or more nucleotides on both the template and the non-template strand. To our knowledge, the impact of introduced mismatches between template and non-template strand on the binding and activity of the T7 RNAP has not been analyzed yet. Based on our data, it seems very likely that an only partially double-stranded promoter element is sufficient to promote binding of the T7 RNAP and lead to transcription initiation. In conclusion, we feel that experiments involving primer annealing approaches for RNA synthesis by the T7 RNAP – in particular when primers exhibit high self-annealing potentials – should be taken with caution.

Acknowledgments

Financial support from the Deutsche Forschungsgemeinschaft to JN (Ni390/7-1) and from Studienstiftung des Deutschen Volkes to LK is acknowledged.

References

Arnaud N, Cheynet V, Oriol G, Mandrand B, Mallet F. (1997) Construction and expression of a modular gene encoding bacteriophage T7 RNA polymerase *Gene* 199:149–156. doi:10.1016/S0378-1119(97)00362–4.

Chamberlin M, Ryan T. (1982) Bacteriophage DNA-dependent RNA polymerases. In: Paul DB (ed) The Enzymes, vol 15. Academic Press, Cambridge, pp. 87–108. doi:10.1016/S1874–6047(08)60276–0.

Cheetham GM, Jeruzalmi D, Steitz TA. (1999) Structural basis for initiation of transcription from an RNA polymerase-promoter complex *Nature* 399:80–83. doi:10.1038/19999.

Cheetham GM, Steitz TA. (2000) Insights into transcription: structure and function of single-subunit DNA-dependent RNA polymerases. *Curr Opin Struct Biol*. 10:117–123.

Davanloo P, Rosenberg AH, Dunn JJ, Studier FW. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase *Proc Natl Acad Sci USA* 81:2035–2039.

Dunn JJ, Studier FW, Gottesman M. (1983) Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *J Mol Biol*. 166:477–535. doi:10.1016/S0022-2836(83)80282–4.

Ikeda RA, Richardson CC. (1986) Interactions of the RNA polymerase of bacteriophage T7 with its promoter during binding and initiation of transcription *Proc Natl Acad Sci USA* 83:361–3618.

Ikeda RA, Warshamana GS, Chang LL. (1992) *In vivo* and *in vitro* activities of point mutants of the bacteriophage T7 RNA polymerase promoter *Biochemistry* 31:9073–9080.

Landick R. (1997) RNA polymerase slides home: pause and termination site recognition *Cell* 88:741–744 doi:10.1016/S0092–8674(00)81919–4.

Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res*. 15:8783–8798.

Oakley JL, Coleman JE. (1977) Structure of a promoter for T7 RNA polymerase. *Proc Natl Acad Sci USA* 74:4266–4270.

Rong M, He B, McAllister WT, Durbin RK. (1998) Promoter specificity determinants of T7 RNA polymerase. *Proc Natl Acad Sci USA* 95:515–519.

Ruwe H, Schmitz-Linneweber C. (2011) Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? *Nucleic Acids Res*. 40:3106–3116. doi:10.1093/nar/gkr1138.

Sousa R, Chung YJ, Rose JP, Wang BC. (1993) Crystal structure of bacteriophage T7 RNA polymerase at 3.3 A resolution. *Nature* 364:593-599. doi:10.1038/364593a0.

Steitz TA. (2009) The structural changes of T7 RNA polymerase from transcription initiation to elongation. *Curr Opin Struct Biol*. 19:683. doi:10.1016/j.sbi.2009.09.001.

Studier FW, Moffatt BA. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*. 189:113-130. doi:10.1016/0022–2836(86)90385–2.

Tabor S, Richardson CC. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc Natl Acad Sci USA* 82:1074–078.

Temme K, Hill R, Segall-Shapiro TH, Moser F, Voigt CA. (2012) Modular control of multiple pathways using engineered orthogonal T7 polymerases. *Nucleic Acids Res*. 40:8773–8781. doi:10.1093/nar/gks597.

Williams-Carrier R, Kroeger T, Barkan A. (2008) Sequence-specific binding of a chloroplast pentatricopeptide repeat protein to its native group II intron ligand. *RNA* 14:1930–1941. doi:10.1261/rna.1077708.

4 Discussion

4.1 RAP, the sole OPR protein in Arabidopsis, plays an important role in chloroplast biogenesis

Helical repeat proteins play a major role as RNA binding proteins in chloroplast gene expression. The *A. thaliana* protein RAP belongs to the family of OPR proteins, of which several described proteins in algae are actually involved in gene expression (Auchincloss et al., 2002; Balczun et al., 2005; Boulouis et al., 2015; Eberhard et al., 2011; Merendino et al., 2006; Murakami et al., 2005; Perron et al., 2004; Rahire et al., 2012; Wang et al., 2015). It was therefore assumed that RAP, the only OPR protein in Arabidopsis, might fulfill a similar role.

4.1.1 RAP is involved in the maturation of the *16S* rRNA

The protein is named after the RAP domain (RNA-binding domain abundant in Apicomplexans) located at its C-terminus, which is a putative RNA-binding domain found in various eukaryotic proteins, especially abundant in apicomplexans (Lee and Hong, 2004). The knockout mutant *rap-1* exhibits a retarded growth phenotype, a defect in photosynthetic activity as well as an impaired protein translation in the chloroplast (Fig 2B-D, Kleinknecht et al., 2014). Most remarkably, a severe reduction of the plastid encoded *16S* rRNA was uncovered. Detailed analysis of the plastid rRNA operon showed no changes in levels or pattern of *23S*, *4.5S* and *5S* rRNA. In strong contrast, *16S* rRNA precursor transcripts accumulated, while the mature transcript was missing entirely (Fig 3, Kleinknecht et al., 2014).

As described in detail in section 1.4.2, the *rrn* operon is transcribed from both a NEP and a PEP promotor, named PC and P2 respectively, of which P2 is predominantly used in green tissue of *A. thaliana* (Sriraman et al., 1998; Swiatecka-Hagenbruch et al., 2007). Furthermore, a mapped precursor processing site (PRO) is located at nucleotide -31 in respect to the start of the mature transcript (Lerbs-Mache, 2000). In wild-type leaves, mature *16S* rRNA as wells as transcripts originating from the transcription initiation site P2, the precursor processing site PRO and the mature 5' end can be observed (Bisanz et al., 2003). Even though all three *16S* 5' ends appear to be present in the *rap-1* mutant, the mature RNA is severely reduced while the amount of the P2 and PRO precursors as well as many processed/degraded transcripts downstream of the P2 promotor are enriched (Fig 4B, Kleinknecht et al., 2014).

This was confirmed by circular RT-PCR, which additionally disclosed that no correctly processed mature *16S* rRNA accumulated in *rap-1*, but that all seemingly mature transcripts were processed at -1 instead of the +1 nucleotide (Fig 9, Kleinknecht et al., 2014). Most of the chloroplast RNAs analyzed in *rap-1* were also less abundant than the respective transcripts in the wild type (Fig 2E, Kleinknecht et al., 2014). This phenotype was observed in several rRNA mutants and it was previously documented that some transcripts require ribosomal loading and/or translation for their stability (Barkan, 1993; Bisanz et al., 2003; Yamamoto et al., 2000). These findings can be summed up in so far that *rap-1* exhibits significant defects in processing of the *16S* rRNA, which in turn lead to its observed phenotype.

RNA deep-sequencing data sets were analyzed extensively in a study to identify short RNAs in the chloroplast genome (Ruwe and Schmitz-Linneweber, 2011). 50 abundant small chloroplast RNAs (sRNA) were identified, most of which are located in non-coding regions, often shortly upstream of start codons. These sRNAs are proposed to be binding sites of PPR proteins, PPR-like proteins and more general RNA-binding proteins. It was shown so far for maize PPR10 and CRP1 as well as for Arabidopsis HCF152 and HCF107 proteins, that identified sRNAs resemble their RNA binding sites (Hammani et al., 2012; Ruwe and Schmitz-Linneweber, 2011; Zhelyazkova et al., 2011). Short stretches of RNA are protected against nucleolytic cleavage by these proteins binding to them; sRNAs thus called footprints (FPs) accumulate. The accumulation of three footprints, located in close proximity to the mature *16S* rRNA, in dependence of RAP was analyzed (Fig 4A and C, Kleinknecht et al., 2014). Interestingly, footprint 2 accumulates in the wild-type RNA but not in *rap-1*, suggesting this region as a possible binding site of RAP or RAP-dependent binding of another protein at this site.

These observations are in agreement not only data obtained for *E. coli*, but also with the phenotypes of several plant rRNA maturation mutants. As rRNA maturation and ribosome assembly are closely linked processes, mutants with defective rRNA processing often have defects in the translation of proteins (Bisanz et al., 2003; Bollenbach et al., 2007; Schmitz-Linneweber et al., 2006; Williams and Barkan, 2003). This can be explained by the fact that defects in rRNA maturation might lead to reduced polysomal loading, which in turn will lead to the observed decreased translation rate (Barkan, 1993; Beligni and Mayfield, 2008; Bellaoui et al., 2003; Sharwood et al., 2011). We propose that the accumulation of precursors in *rap-1* leads to an impeded translational activity by preventing the formation of active ribosomes as previously suggested by Bisanz et al. (2003) for the *dal* mutant, which exhibits a

similar accumulation of *16S* rRNA precursors. Moreover, it is documented that in *E. coli* complete processing of the *16S* rRNA requires at first an interaction with the 30S ribosome subunit and hereafter even the assembly of complete polysomes (Shajani et al., 2011). In addition, *in vitro* reconstitution assays uncovered that the bacterial 30S subunit loaded with premature *16S* rRNA are inactive, thereby suggesting that correct processing of the *16S* rRNA is required for protein translation (Wireman and Sypherd, 1974). Nevertheless, more recent data revealed that extensions at the *16S* rRNA 5' end might have an influence rather on translational fidelity due to misfolding of the rRNA than on the actual assembly of the ribosome (Gutgsell and Jain, 2012; Roy-Chaudhuri et al., 2010).

Taken together the data discussed above strongly suggest that RAP mediates its function by binding just downstream of the precursor processing site upstream of the 5′ end of the mature *16S* rRNA. In agreement with this hypothesis *in vitro* RNA binding experiments revealed that RAP has an intrinsic RNA binding capacity (Fig 5, Kleinknecht et al., 2014). It needs to be determined whether the RNA binding capacity of RAP is effectuated by its OPR domain, its RAP domain, or both of them as all three scenarios seem plausible. It was demonstrated by Rahire et al. (2012) that the translation initiation factor Tab1 displays its RNA binding activity solely through its OPR domain. On the other hand it was proposed by Lee and Hong (2004), that the RNA-binding protein Raa3, which is involved in splicing of the *psaA* Operon in *C. reinhardtii*, binds to its target mediated by its RAP domain.

With this information in mind, we can speculate how exactly RAP influences the maturation process of the *16S* rRNA. RAP itself could resemble an endonuclease that directly binds to its target and processes the 5' leader region of the *16S* rRNA. This hypothesis seems unlikely for two reasons: (I) Processing in the *rap-1* mutant is very ineffective but still takes place. (II) An *in vitro* assay performed in order to uncover an intrinsic RNase activity of RAP, did not reveal any nucleolytic activity (data not shown). A far more likely option, is that RAP interacts with one or more endonucleases guiding them to the correct processing site, since it is assumed that plastid RNases (e.g. RNase J or RNase R) possess little to no sequence specificity of their own (Germain et al., 2013; Stoppel and Meurer, 2011). Interestingly, a possible concerted function of RAP and RNase J in processing of the *16S* rRNA is supported by the phenotype of RNase J-deficient tobacco and Arabidopsis plants, which display a decreased accumulation of mature *16S* rRNA as well as an over-accumulation of 5' extended precursor *16S* rRNAs (Sharwood et al., 2011). Additionally, Luro et al. (2013) observed an interactive role of RNase J and helical repeat proteins in tobacco, as they postulate that the

enzyme does trim chloroplast 5' mRNA ends to their mature form defined by bound PPR proteins. If RAP indeed interacts with RNase J - which still needs to be determined - it would resemble the protein RHON1, which is described to interact with RNase E and confer sequence specificity to it in order to correctly trim the *5'* end of the 23S rRNA in Arabidopsis chloroplasts (Stoppel et al., 2012). Last but not least, a third possible working model that needs to be taken into account is that RAP could change the secondary structure of the *16S* rRNA precursor by binding to it and thereby facilitating the accessibility of sequence-specific RNases and/or other involved factors.

Moreover, it has been suggested that chloroplast transcription and ribosome assembly are closely linked (Germain et al., 2013; Majeran et al., 2012). Recent findings, like the nucleoid localization of the protein PAC, which is involved in *23S* rRNA processing, support the hypothesis that nucleoids are the site of rRNA processing within the chloroplast (reviewed in Bohne, 2014; Jeon et al., 2014; Meurer et al., 2017). Consistently, the maize ortholog of RAP has been identified in the nucleoid proteome (Majeran et al., 2012), and this localization was confirmed cytologically (Fig 7, Kleinknecht et al., 2014).

4.1.2 RAP as negative regulator of plant defense

In a study conducted by Katiyar-Agarwal et al. (2007), a different effect in the *rap-1* mutant was observed. The mutant is less sensitive to infection with the parasitic microbe *Pseudomonas syringae*. In wild-type plants *AtlsiRNA-1*, a short RNA that belongs to a novel class of short interfering RNAs, is expressed upon infiltration with *P. syringae* and specifically leads to degradation of *RAP* mRNA. Therefore, RAP was proposed to act as a negative regulator in disease resistance (Katiyar-Agarwal et al., 2007; Katiyar-Agarwal and Jin, 2010). Its role in *16S* rRNA processing might be either an additional function or explain the effects observed in their study. By down-regulating processing of the *16S* rRNA and therefore diminishing translation of chloroplast proteins in infected cells, spread of the bacteria might be embanked by limiting their access to nutrients and lead to faster death of the infected cells. Further analyses - like Northern blots to analyze the processing of *16S* rRNA in wild-type plants after *P. syringae* infection – will be necessary to elucidate RAP's molecular function in plant defense.

4.1.3 How conserved is the function of the single OPR protein in higher plants?

Interestingly, only a single OPR protein was identified in Streptophyta, like *A. thaliana*, whereas more than 100 OPR proteins were identified in the Chlorophyta *C. reinhardtii* (O. Vallon, A. Bohne, L. Cerutti, J. D. Rochaix, unpublished data). The number of another family of helical repeat proteins, the PPR proteins, is reversed. The family of PPRs consists of more than 450 members in *A. thaliana*, but only of 11 proteins in *C. reinhardtii* (Schmitz-Linneweber and Small, 2008). Presuming that a certain number of RNA binding proteins is required to ensure correct and specific processing of many organellar transcripts, this leads to the assumption that PPR proteins in higher plants take over the function of OPR proteins in algae. Whether OPR proteins were present in a high number before the transition to land plants occurred and subsequently substituted by PPR proteins is unknown so far. It is also possible, that the high number of OPR proteins is due to an expansion of the family in Chlorophytes that took place after the separation of Viridaeplantae in Streptophyta and Chlorophyta (O. Vallon, A. Bohne, L. Cerutti, J. D. Rochaix, unpublished data). Further analyses in future will hopefully show, which of the above mentioned theories holds true.

A multiple sequence alignment of the OPR proteins from several Streptophyta showed a high sequence conservation indicating the same evolutionary origin, whereas in the Chlorophyta *C. reinhardtii* no direct ortholog could be found (Fig S1, Kleinknecht et al., 2014). Most interestingly, Zhang et al. (2016) recently analyzed a rice mutant *al1* (albino leaf1), which is orthologous to the Arabidopsis *RAP* gene. The mutant exhibited an albino phenotype in leaves and resulted in seedling lethality. A complementation of the *rap-1* mutant with the *Al1* gene led to a wild-type like phenotype indicating that AL1 is a functional ortholog of RAP. Mature *16S* rRNA accumulated to a lesser degree in the *al1* mutant than in wild-type plants whereas the accumulation of precursors was unfortunately not determined. In agreement with the situation in *rap-1*, the accumulation of several chloroplast proteins and transcripts is reduced. Furthermore, they do suggest an additional function of AL1 in homeostasis of ribosomal proteins.

These findings clearly support the hypothesis that the OPR proteins in Streptophyta are highly conserved not only in sequence but also in function. It needs to be determined in future studies whether they do not only play a similar function in *16S* rRNA processing as RAP, but also whether and when in evolution additional functions were acquired.

4.2 The multifunctional moonlighting enzyme DLA2

Even though helical repeat proteins constitute supposedly the majority of RNA binding proteins in the organelles of plants and algae, proteins that do not belong to this superfamily are frequently identified to play important roles as well. Some of these belong to the interesting group of moonlighting proteins, which acquired an additional function next to their - often enzymatic - main function (see section 1.5 and Huberts and van der Klei, 2010). The protein DLA2, the E2 subunit of the chloroplast pyruvate dehydrogenase complex in *C. reinhardtii*, belongs to this group of dual-functioning proteins (Bohne et al., 2013; Ossenbühl et al., 2002). As described in section 1.5.1 in detail, the cpPDC catalyzes the oxidative decarboxylation of pyruvate in the chloroplast stroma thereby providing acetyl-CoA for fatty acid synthesis in the chloroplast (Lin and Oliver, 2008; Mooney et al., 2002). Bohne et al. (2013) were able to show that under mixotrophic conditions DLA2 additionally forms part of an RNP complex with the *psbA* mRNA. Furthermore, they are proposing based on their results that it plays a role in localizing this mRNA to the T-zone for its translation and the synthesis of the PSII protein D1.

4.2.1 Competition between the two functions of DLA2

The conducted SEC and 2D-BN-PAGE disclosed that the other two cpPDC subunits, E1 and E3, are not part of the RNase-sensitive DLA2 complex leading to the conclusion that the RNP complex is completely distinct from the cpPDC (Fig 1, Kleinknecht et al., in preparation). The composition of the DLA2-RNA-complex remains elusive so far. Even though multiple copies of DLA2 attached to the RNA might be able to form a high molecular weight complex, it seems more likely that additional proteins are part of this complex. Therefore, the question, whether the cpPDC is continuously assembled and disassembled to free DLA2 for its secondary function or whether additional DLA2 accumulates during mixotrophic conditions, arose. The analysis of mutants of all three cpPDC subunits under mixotrophic conditions revealed that the accumulation of the subunits is dependent on the presence and/or amount of the other subunit, whereas E3 was not influencing the other two subunits as much as they influence each other and vice versa. This of course leads directly to speculation why the accumulation of E3 is more independent. The easiest explanation is to postulate a higher stability of the unassembled protein compared to the other subunits. A more complex and daring hypothesis is to assume an additional function of the protein, therefore requiring an accumulation of E3 independent from the other subunits. Up to now, there is no

additional proof to support this hypothesis and future studies will hopefully answer this question.

To gain further insights, quantification of the cpPDC subunit protein levels in wild type grown under mixotrophic, photoautotrophic and heterotrophic conditions was performed, revealing that protein levels do not change under these different conditions and therefore the ratio between the subunits stays the same. Interestingly, the two E1 isoforms, E1α and E1β, do not accumulate to the expected 1:1 ratio but rather in a 1:2.5 ratio (Fig 3, Kleinknecht et al., in preparation). This observed ratio differs from previously reported ratios, which might be due to the fact that the ratio is different between organisms or even more likely that here the ratio of the total protein amounts in the cell was observed, whereas other studies identified the ratio within the assembled complex (Johnston et al., 2000). Nevertheless, the interesting conclusion was drawn, that there is no change in subunit ratio under different growth conditions. This further confirms the idea that no additional DLA2 is accumulating under mixotrophic conditions, but rather that the cpPDC might at least partially disassemble to free DLA2 for its additional function as RNA-binding protein. In agreement with this, it was postulated before that the cpPDC is less defined as compared to mitochondrial PDC complexes and might disassemble slightly easier (Camp and Randall, 1985; Zhou et al., 2001)

In conformity with these findings, Bohne et al. (2013) postulated that the RNA binding site of DLA2 overlaps with its E3 binding domain. *In silico* analyses have shown that it contains a number of positively charged amino acids suitable for RNA binding as well as a predicted Rossmann fold, which is typically a dinucleotide binding site (Rao and Rossmann, 1973). Interestingly, it was reported that this short motif is involved in the RNA binding activity of several enzymes, e.g. the glyceraldehyde-3-phosphate dehydrogenase (Benning, 2009; Nagy et al., 2000; Nagy and Rigby, 1995). This hypothesis of an overlapping E3 and RNA binding site in DLA2 was not only confirmed, but in addition a competitive binding mode of the *psbA* mRNA and E3 to this site was revealed *in vitro* (Fig 4 and 5, Kleinknecht et al., in preparation). In summary, the described results reveal that E1 and E3 are not part of the DLA2-RNA complex and even more importantly that there is there is a competitive binding mode of RNA and E3 to DLA2, raising the question in which way these two functions of DLA2 are regulated.

4.2.2 Metabolic control of *psbA* gene expression

Changes in gene expression are often an adaptation to environmental conditions and the metabolic status of the cell. There is profound evidence that bacterial as well as eukaryotic

gene expression is modulated by carbon metabolism (Barańska et al., 2013; Wellen and Thompson, 2012). One specific example are mammalian glycolytic enzymes involved in the carbon metabolism in the cytoplasm, that are regulators of transcription, translation or can affect mRNA stability (reviewed in Kim and Dang, 2005). Many of those are rate limiting enzymes and therefore particularly suitable to play a role in coordination of gene expression as an adaptation to a changing metabolic status.

The first such enzyme described in plants is a sulfite reductase that – next to its enzymatic function – possesses a DNA-binding capacity and is known to regulate transcription by condensing the chloroplast nucleoids (Sekine et al., 2007; Sekine et al., 2002). The second one is the large subunit of the well-known ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) which catalyzes the carboxylation of ribulose-1,5-bisphosphate – the first major step in carbon fixation via the Calvin cycle (see section 1.2.1). Cohen et al. (2006) were able to show that it acquires an RNA-binding activity under oxidizing conditions and might therefore be involved in an autoregulatory feedback translational repression. Moreover, it is involved in the formation of stress granules in the chloroplast, most likely in order to degrade oxidized RNAs (Uniacke and Zerges, 2008; Zhan et al., 2015). The third and final one is the here described DLA2 protein.

It was described above that the gene expression is adjusted in accordance with the demands of the cell. Metabolites can serve as low molecular weight signaling molecules or as substrates for post-translational modifications (PTMs). Several studies have shown that besides regulation on RNA level a high number of protein functions are regulated by PTMs, that occur either at the protein's N- or C-terminus or on the amino acid side chains. They include most prominently phosphorylation as well as other modifications like glycosylation, carbonylation and acetylation (Khoury et al., 2011). These alterations can unmask the additional function of metabolic enzymes (Jeffery, 2003). Interestingly, Bohne et al (2013) demonstrated that the formation of the DLA2-RNA-complex relies on mixotrophic growth conditions. In this conditions, the present acetate can be converted to acetyl-CoA by the acetate synthetase (ACS) and/or by the acetyl-kinase/phosphate acetyltransferase (ACK/PAT) system (Spalding, 2009). This in turn can be used for fatty acid synthesis in the chloroplast and can lead to a product inhibition of the cpPDC and/or the acetylation of its subunits (Tovar-Méndez et al., 2003; Wellen and Thompson, 2012; Xing and Poirier, 2012). Therefore, a regulation mechanism of DLA2 through acetylation seems very likely. Lysine acetylation is a ubiquitous PTM and was described in various organisms including bacteria,

yeast, plant and animal cells (Choudhary et al., 2009; Choudhary et al., 2014; Finkemeier et al., 2011; Henriksen et al., 2012; Lundby et al., 2012; Melo-Braga et al., 2012; Mo et al., 2015; Weinert et al., 2011).

In fact, DLA2 carries three lysine acetylation sites including one that is significantly upregulated under mixo- and heterotrophic conditions when compared to photoautotrophic conditions (Section 6.1, König et al., in preparation). Most remarkably, this regulated residue K_{197} lies within the E3 and proposed RNA binding domain of DLA2. Moreover, it was shown that the RNA binding activity of DLA2 can be influenced by the acetylation status of DLA2, further supporting the idea of regulation through acetylation (Section 3.1, Fig 6, Kleinknecht et al., in preparation). In future, one has to mutagenize the individual lysine sites and repeat the RNA binding experiment to reveal the role of lysine acetylation in the regulation of DLA2. An additional supporting this hypothesis is that the translation of D1 is induced by the addition of actetate to the growth medium (Michaels and Herrin, 1990). In conclusion, metabolic control of the *psbA* gene expression through the acetylation of DLA2 seems a likely possibility.

4.2.3 Is the multifunctionality of DLA2 evolutionary conserved?

Bohne et al. (2013) suggested that DLA2's role in gene expression might be a very ancient feature of the dihydrolipoamide acetyltransferases, which likely occurred first before the separation of mitochondrial and chloroplastic homologues. It was demonstrated that recombinant proteins from *C. reinhardtii*, human, yeast and the cyanobacterium *Synechocystis* sp. PCC 6803 all have an intrinsic RNA binding activity *in vitro*. Obviously, *in vitro* experiments do not prove a role of these E2 proteins in gene expression, but there are several other hints that support this hypothesis.

The analysis of high molecular weight complexes from *A. thaliana* by SEC revealed that LTA2, the cpPDC E2 subunit, occurred not only in a fraction of about 1-2 MDa in size together with the other cpPDC subunits, but in a second high molecular weight complex of a size of more than 5 MDa (Olinares et al., 2010). Interestingly, several ribosomal subunits and RNA-binding proteins were found in the same fraction. One of these RNA-binding proteins, CSP41 was analyzed by several groups in more detail. It interacts with a number of mRNAs of photosynthetic proteins as well as with the *16S* and *23S* rRNAs and is postulated to stabilize and/or process these RNAs (Beligni and Mayfield, 2008; Bollenbach et al., 2009; Chevalier et al., 2015; Qi et al., 2012). Remarkably, LTA2 was found to be an interacting partner of CSP41 in a co-immunoprecipitation experiment conducted by Qi et al. (2012),

further supporting the existence of a second LTA2-containing complex. Furthermore, an additional function is indicated by the phenotypes of T-DNA insertion lines of LTA2 and the E3 subunit. Strikingly, the LTA2 line is lethal, while the E3 mutant line is viable (Chen et al., 2010; Lin et al., 2003). This implies that Arabidopsis plants are viable without the cpPDC complex, and raises the question which additional function of LTA2 leads to the lethality of its knockout mutant.

Moreover, recent data obtained by 2D-BN-PAGE analyses of *Synechocystis* thylakoids, strongly support the assumption of an evolutionary conservation of DLA2's role in gene expression. It was shown that the E2 subunit (SynE2) does form an RNP complex *in vivo* under the same conditions as DLA2 in *Chlamydomonas* (Fig S1, Appendix). It will be most interesting to analyze SynE2's role in more detail and to elucidate its target RNA as well as its regulatory function. In its entirety, a conserved role of the PDC E2 subunit in gene expression seems very likely and in future the characterization SynE2 as well as that of other E2 subunits promises further insight into the development of DLA2's moonlighting function.

Taken together, the results of this thesis enhanced the understanding of several aspects that fulfill important functions in the post-transcriptional regulation of chloroplast gene expression in *A. thaliana* and *C. reinhardtii*.

5 References

Allen, J.F., de Paula, W.B., Puthiyaveetil, S., and Nield, J. (2011). A structural phylogenetic map for chloroplast photosynthesis. Trends Plant Sci *16*, 645-655.

Allison, L.A. (2000). The role of sigma factors in plastid transcription. Biochimie *82*, 537-548.

Armbrust, E.V., Berges, J.A., Bowler, C., Green, B.R., Martinez, D., Putnam, N.H., Zhou, S., Allen, A.E., Apt, K.E., Bechner, M., *et al.* (2004). The genome of the diatom *Thalassiosira pseudonana*: ecology, evolution, and metabolism. Science *306*, 79-86.

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

Balczun, C., Bunse, A., Hahn, D., Bennoun, P., Nickelsen, J., and Kück, U. (2005). Two adjacent nuclear genes are required for functional complementation of a chloroplast trans-splicing mutant from *Chlamydomonas reinhardtii*. Plant J *43*, 636-648.

Barańska, S., Glinkowska, M., Herman-Antosiewicz, A., Maciąg-Dorszyńska, M., Nowicki, D., Szalewska-Pałasz, A., Węgrzyn, A., and Węgrzyn, G. (2013). Replicating DNA by cell factories: roles of central carbon metabolism and transcription in the control of DNA replication in microbes, and implications for understanding this process in human cells. Microb Cell Fact *12*, 55.

Barkan, A. (1993). Nuclear mutants of maize with defects in chloroplast polysome assembly have altered chloroplast RNA metabolism. Plant Cell *5*, 389-402.

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

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

Barnes, D., and Mayfield, S.P. (2003). Redox control of posttranscriptional processes in the chloroplast. Antioxid Redox Signal *5*, 89-94.

Bassham, J.A., Benson, A.A., and Calvin, M. (1950). The path of carbon in photosynthesis. J Biol Chem *185*, 781-787.

Beligni, M.V., and Mayfield, S.P. (2008). *Arabidopsis thaliana* mutants reveal a role for CSP41a and CSP41b, two ribosome-associated endonucleases, in chloroplast ribosomal RNA metabolism. Plant Mol Biol *67*, 389-401.

Bellaoui, M., Keddie, J.S., and Gruissem, W. (2003). DCL is a plant-specific protein required for plastid ribosomal RNA processing and embryo development. Plant Mol Biol *53*, 531-543.

Benning, C. (2009). Mechanisms of lipid transport involved in organelle biogenesis in plant cells. Annu Rev Cell Dev Biol *25*, 71-91.

Bisanz, C., Bégot, L., Carol, P., Perez, P., Bligny, M., Pesey, H., Gallois, J.-L., Lerbs-Mache, S., and Mache, R. (2003). The Arabidopsis nuclear *DAL* gene encodes a chloroplast protein which is required for the maturation of the plastid ribosomal RNAs and is essential for chloroplast differentiation. Plant Mol Biol *51*, 651-663.

Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. BioEssays *21*, 932-939.

Bock, R., and Timmis, J.N. (2008). Reconstructing evolution: gene transfer from plastids to the nucleus. BioEssays *30*, 556-566.

Bohne, A.-V., and Nickelsen, J. (2017). Metabolic control of chloroplast gene expression: an emerging theme. Mol Plant *10*, 1-3.

Bohne, A.-V., Schwarz, C., Schottkowski, M., Lidschreiber, M., Piotrowski, M., Zerges, W., and Nickelsen, J. (2013). Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. PLoS Biol *11*, e1001482.

Bohne, A.V. (2014). The nucleoid as a site of rRNA processing and ribosome assembly. Front Plant Sci *5*, 257.

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

Bollenbach, T.J., Schuster, G., Portnoy, V., and Stern, D.B. (2007). Processing, degradation, and polyadenylation of chloroplast transcripts. In Cell and Molecular Biology of Plastids, R. Bock, ed. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 175-211.

Bollenbach, T.J., Schuster, G., and Stern, D.B. (2004). Cooperation of endo- and exoribonucleases in chloroplast mRNA turnover. Prog Nucleic Acid Res Mol Biol *78*, 305-337.

Bollenbach, T.J., Sharwood, R.E., Gutierrez, R., Lerbs-Mache, S., and Stern, D.B. (2009). The RNA-binding proteins CSP41a and CSP41b may regulate transcription and translation of chloroplast-encoded RNAs in Arabidopsis. Plant Mol Biol *69*, 541-552.

Börner, T., Aleynikova, A.Y., Zubo, Y.O., and Kusnetsov, V.V. (2015). Chloroplast RNA polymerases: Role in chloroplast biogenesis. Biochimica Biophys Acta *1847*, 761-769.

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

Boulouis, A., Drapier, D., Razafimanantsoa, H., Wostrikoff, K., Tourasse, N.J., Pascal, K., Girard-Bascou, J., Vallon, O., Wollman, F.A., and Choquet, Y. (2015). Spontaneous dominant mutations in chlamydomonas highlight ongoing evolution by gene diversification. Plant Cell *27*, 984-1001.

Bresinsky, A., Körner, C., Kadereit, J.W., Neuhaus, G., and Sonnewald, U. (2008). Strasburger – Lehrbuch der Botanik. (Heidelberg: Spektrum Akademischer Verlag).

Buchanan, B.B. (2016). The carbon (formerly dark) reactions of photosynthesis. Photosyn Res *128*, 215-217.

Camp, P.J., and Randall, D.D. (1985). Purification and characterization of the pea chloroplast pyruvate dehydrogenase complex : a source of acetyl-Coa and NADH for fatty acid biosynthesis. Plant Physiol *77*, 571-577.

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

Chateigner-Boutin, A.L., des Francs-Small, C.C., Delannoy, E., Kahlau, S., Tanz, S.K., de Longevialle, A.F., Fujii, S., and Small, I. (2011). OTP70 is a pentatricopeptide repeat protein of the E subgroup involved in splicing of the plastid transcript *rpoC1*. Plant J *65*, 532-542.

Chen, W., Chi, Y., Taylor, N.L., Lambers, H., and Finnegan, P.M. (2010). Disruption of *ptLPD1* or *ptLPD2*, genes that encode isoforms of the plastidial lipoamide dehydrogenase, confers arsenate hypersensitivity in Arabidopsis. Plant Physiol *153*, 1385-1397.

Chevalier, F., Ghulam, M.M., Rondet, D., Pfannschmidt, T., Merendino, L., and Lerbs-Mache, S. (2015). Characterization of the *psbH* precursor RNAs reveals a precise endoribonuclease cleavage site in the *psbT/psbH* intergenic region that is dependent on *psbN* gene expression. Plant Mol Biol *88*, 357-367.

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

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.

Choudhary, C., Weinert, B.T., Nishida, Y., Verdin, E., and Mann, M. (2014). The growing landscape of lysine acetylation links metabolism and cell signalling. Nat Rev Mol Cell Biol *15*, 536-550.

Cohen, I., Sapir, Y., and Shapira, M. (2006). A conserved mechanism controls translation of Rubisco large subunit in different photosynthetic organisms. Plant Physiol *141*, 1089-1097.

D'Andrea, L.D., and Regan, L. (2003). TPR proteins: the versatile helix. Trends Biochem Sci *28*, 655-662.

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

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

Danon, A., and Mayfield, S.P. (1994b). Light-regulated translation of chloroplast messenger RNAs through redox potential. Science *266*, 1717-1719.

Dekker, J.P., and Boekema, E.J. (2005). Supramolecular organization of thylakoid membrane proteins in green plants. Biochim Biophys Acta *1706*, 12-39.

del Campo, E.M. (2009). Post-transcriptional control of chloroplast gene expression. Gene Regu Syst Biol *3*, 31-47.

Derelle, E., Ferraz, C., Rombauts, S., Rouze, P., Worden, A.Z., Robbens, S., Partensky, F., Degroeve, S., Echeynie, S., Cooke, R., *et al.* (2006). Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. Proc Natl Acad Sci USA *103*, 11647-11652.

Eberhard, S., Loiselay, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., Choquet, Y., and Wollman, F.-A. (2011). Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of *atpA* transcripts in *Chlamydomonas reinhardtii* chloroplasts. Plant J *67*, 1055-1066.

Favory, J.J., Kobayshi, M., Tanaka, K., Peltier, G., Kreis, M., Valay, J.G., and Lerbs-Mache, S. (2005). Specific function of a plastid sigma factor for *ndhF* gene transcription. Nucleic Acids Res *33*, 5991-5999.

Finkemeier, I., Laxa, M., Miguet, L., Howden, A.J., and Sweetlove, L.J. (2011). Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. Plant Physiol *155*, 1779-1790.

Foth, B., M Stimmler, L., Handman, E., Crabb, B., N Hodder, A., and McFadden, G. (2005). The malaria parasite *Plasmodium falciparum* has only one pyruvate dehydrogenase complex, which is located in the apicoplast. Mol Microbiol *55*, 39-53.

Gancedo, C., and Flores, C.L. (2008). Moonlighting proteins in yeasts. Microbiol Mol Biol Rev *72*, 197-210

Germain, A., Hotto, A.M., Barkan, A., and Stern, D.B. (2013). RNA processing and decay in plastids. RNA *4*, 295-316.

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

Gray, J.C., Sullivan, J.A., Wang, J.H., Jerome, C.A., and MacLean, D. (2003). Coordination of plastid and nuclear gene expression. Biol Sci *358*, 135-144.

Guest, J.R., Lewis, H.M., Graham, L.D., Packman, L.C., and Perham, R.N. (1985). Genetic reconstruction and functional analysis of the repeating lipoyl domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. J Mol Biol *185*, 743-754.

Gully, B.S., Cowieson, N., Stanley, W.A., Shearston, K., Small, I.D., Barkan, A., and Bond, C.S. (2015). The solution structure of the pentatricopeptide repeat protein PPR10 upon binding *atpH* RNA. Nucleic Acids Res *43*, 1918-1926.

Gunstone, F.D., Harwood, J.L., and Padley, F.B. (1994). The Lipid Handbook, Second Edition (Taylor & Francis).

Gutgsell, N.S., and Jain, C. (2012). Gateway role for rRNA precursors in ribosome assembly. J Bacteriol *194*, 6875-6882.

Hammani, K., Cook, W.B., and Barkan, A. (2012). RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. Proc Natl Acad Sci USA *109*, 5651-6.

Harris, E.H. (1989). The Chlamydomonas Sourcebook. A Comprehensive Guide to Biology and Laboratory Use. , Vol 246, 1989/12/15 edn (San Diego, CA: Academic Press).

Harris, E.H., Boynton, J.E., and Gillham, N.W. (1994). Chloroplast ribosomes and protein synthesis. Microbiol Rev *58*, 700-754.

Harwood, J.L. (2009). Fatty acid biosynthesis. In Plant Lipids: Biology, Utilisation and Manipulation, D.J. Murphy, ed. (Oxford: Blackwell Publishing), pp. 27-66.

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

Henriksen, P., Wagner, S.A., Weinert, B.T., Sharma, S., Bacinskaja, G., Rehman, M., Juffer, A.H., Walther, T.C., Lisby, M., and Choudhary, C. (2012). Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in *Saccharomyces cerevisiae*. Mol Cell Proteom *11*, 1510-1522.

Herrin, D.L., and Nickelsen, J. (2004). Chloroplast RNA processing and stability. Photosyn Res *82*, 301-314.

Hess, W.R., Prombona, A., Fieder, B., Subramanian, A.R., and Börner, T. (1993). Chloroplast *rps15* and the *rpoB/C1/C2* gene cluster are strongly transcribed in ribosome-deficient plastids: evidence for a functioning non-chloroplast-encoded RNA polymerase. Embo J *12*, 563-571.

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

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

Hohmann-Marriott, M.F., and Blankenship, R.E. (2011). Evolution of photosynthesis. Annu Rev Plant Biol *62*, 515-548.

Huberts, D.H.E.W., and van der Klei, I.J. (2010). Moonlighting proteins: An intriguing mode of multitasking. Biochim Biophys Acta *1803*, 520-525.

Jalal, A., Schwarz, C., Schmitz-Linneweber, C., Vallon, O., Nickelsen, J., and Bohne, A.V. (2015). A small multifunctional pentatricopeptide repeat protein in the chloroplast of *Chlamydomonas reinhardtii*. Mol Plant *8*, 412-426.

Jarvis, P., and Soll, J. (2001). Toc, Tic, and chloroplast protein import. Biochim Biophys Acta *1541*, 64-79.

Jeffery, C.J. (1999). Moonlighting proteins. Trends Biochem Sci *24*, 8-11.

Jeffery, C.J. (2003). Moonlighting proteins: old proteins learning new tricks. Trends Gen *19*, 415-417.

Jeffery, C.J. (2018). Protein moonlighting: what is it, and why is it important? Philos Trans R Soc Lond B Biol *19*, 373.

Jeon, Y., Ahn, C.S., Jung, H.J., Kang, H., Park, G.T., Choi, Y., Hwang, J., and Pai, H.S. (2014). DER containing two consecutive GTP-binding domains plays an essential role in chloroplast ribosomal RNA processing and ribosome biogenesis in higher plants. J Exp Bot *65*, 117-130.

Johnston, M.L., Miernyk, J.A., and Randall, D.D. (2000). Import, processing, and assembly of the alpha- and beta-subunits of chloroplast pyruvate dehydrogenase. Planta *211*, 72-76.

Katiyar-Agarwal, S., Gao, S., Vivian-Smith, A., and Jin, H. (2007). A novel class of bacteria-induced small RNAs in Arabidopsis. Genes Dev *21*, 3123-3134.

Katiyar-Agarwal, S., and Jin, H. (2010). Role of small RNAs in host-microbe interactions. Annu Rev Phytopathol *48*, 225-246.

Khoury, G.A., Baliban, R.C., and Floudas, C.A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. Scientific Rep *1*, 90.

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

Kim, J.W., and Dang, C.V. (2005). Multifaceted roles of glycolytic enzymes. Trends Biochem Sci *30*, 142-150.

Kleinknecht, L., Wang, F., Stübe, R., Philippar, K., Nickelsen, J., and Bohne, A.-V. (2014). RAP, the sole Octotricopeptide Repeat Protein in Arabidopsis, is required for chloroplast *16S* rRNA maturation. Plant Cell *26*, 777-787.

Komatsu, T., Kawaide, H., Saito, C., Yamagami, A., Shimada, S., Nakazawa, M., Matsui, M., Nakano, A., Tsujimoto, M., Natsume, M., *et al.* (2010). The chloroplast protein BPG2 functions in brassinosteroid-mediated post-transcriptional accumulation of chloroplast rRNA. Plant J *61*, 409-422.

Koornneef, M., and Meinke, D. (2010). The development of Arabidopsis as a model plant. Plant J *61*, 909-921.

Lee, I., and Hong, W. (2004). RAP-a putative RNA-binding domain. Trends Biochem Sci *29*, 567-570.

Lerbs-Mache, S. (2000). Regulation of rDNA transcription in plastids of higher plants. Biochimie *82*, 525-535.

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

Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., Yue, R., Robertson, J.M., Lefebvre, P., Fitz-Gibbon, S.T., *et al.* (2016). An indexed, mapped mutant library enables reverse genetics studies of biological processes in *Chlamydomonas reinhardtii*. Plant Cell *28*, 367-87

Liere, K., Weihe, A., and Börner, T. (2011). The transcription machineries of plant mitochondria and chloroplasts: Composition, function, and regulation. Plant Physiol *168*, 1345-1360.

Lin, M., Behal, R., and Oliver, D.J. (2003). Disruption of plE2, the gene for the E2 subunit of the plastid pyruvate dehydrogenase complex, in Arabidopsis causes an early embryo lethal phenotype. Plant Mol Biol *52*, 865-872.

Lin, M., and Oliver, D.J. (2008). The role of acetyl-coenzyme a synthetase in Arabidopsis. Plant Physiol *147*, 1822.

Lundby, A., Lage, K., Weinert, B.T., Bekker-Jensen, D.B., Secher, A., Skovgaard, T., Kelstrup, C.D., Dmytriyev, A., Choudhary, C., Lundby, C., *et al.* (2012). Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. Cell Rep *2*, 419-431.

Luro, S., Germain, A., Sharwood, R.E., and Stern, D.B. (2013). RNase J participates in a pentatricopeptide repeat protein-mediated 5′ end maturation of chloroplast mRNAs. Nucleic Acids Res *41*, 9141-51.

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

Majeran, W., Friso, G., Asakura, Y., Qu, X., Huang, M., Ponnala, L., Watkins, K.P., Barkan, A., and van Wijk, K.J. (2012). Nucleoid-enriched proteomes in developing

plastids and chloroplasts from maize leaves: a new conceptual framework for nucleoid functions. Plant Physiol *158*, 156-189.

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

Manthey, G.M., and McEwen, J.E. (1995). The product of the nuclear gene PET309 is required for translation of mature mRNA and stability or production of intron-containing RNAs derived from the mitochondrial COX1 locus of *Saccharomyces cerevisiae.* EMBO J *14*, 4031-4043.

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

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

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

McFadden, G.I., and van Dooren, G.G. (2004). Evolution: red algal genome affirms a common origin of all plastids. Current Biol *14*, R514-516.

Melo-Braga, M.N., Verano-Braga, T., Leon, I.R., Antonacci, D., Nogueira, F.C., Thelen, J.J., Larsen, M.R., and Palmisano, G. (2012). Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (*Vitis vinifera*) mesocarp and exocarp owing to *Lobesia botrana* infection. Mol Cell Proteomics *11*, 945-956.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Marechal-Drouard, L., *et al.* (2007). The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science *318*, 245-250.

Merendino, L., Perron, K., Rahire, M., Howald, I., Rochaix, J.D., and Goldschmidt-Clermont, M. (2006). A novel multifunctional factor involved in trans-splicing of chloroplast introns in Chlamydomonas. Nucleic Acids Res *34*, 262-274.

Meurer, J., Schmid, L.M., Stoppel, R., Leister, D., Brachmann, A., and Manavski, N. (2017). PALE CRESS binds to plastid RNAs and facilitates the biogenesis of the 50S ribosomal subunit. Plant J *92*, 400-413.

Michaels, A., and Herrin, D.L. (1990). Translational regulation of chloroplast gene expression during the light-dark cell cycle of Chlamydomonas: evidence for control by ATP/energy supply. Biochem Biophy Res Commun *170*, 1082-1088.

Mo, R., Yang, M., Chen, Z., Cheng, Z., Yi, X., Li, C., He, C., Xiong, Q., Chen, H., Wang, Q., *et al.* (2015). Acetylome analysis reveals the involvement of lysine acetylation in

photosynthesis and carbon metabolism in the model cyanobacterium *Synechocystis sp.* PCC 6803. J Proteome Res *14*, 1275-1286.

Mooney, B.P., Miernyk, J.A., and Randall, D.D. (1999). Cloning and characterization of the dihydrolipoamides-acetyltransferase subunit of the plastid pyruvate dehydrogenase complex (E2) from Arabidopsis. Plant Physiol *120*, 443-452.

Mooney, B.P., Miernyk, J.A., and Randall, D.D. (2002). The complex fate of α-ketoacids. Annu Rev Plant Biol *53*, 357-375.

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

Nagy, E., Henics, T., Eckert, M., Miseta, A., Lightowlers, R.N., and Kellermayer, M. (2000). Identification of the NAD(+)-binding fold of glyceraldehyde-3-phosphate dehydrogenase as a novel RNA-binding domain. Biochem Biophys Res Commun *275*, 253-260.

Nagy, E., and Rigby, W.F. (1995). Glyceraldehyde-3-phosphate dehydrogenase selectively binds AU-rich RNA in the NAD($^{+}$)-binding region (Rossmann fold). J Biol Chem *270*, 2755-2763.

Nelson, N. (2011). Photosystems and global effects of oxygenic photosynthesis. Biochim Biophys Acta *1807*, 856-863.

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

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

Nickelsen, J., and Kück, U. (2000). The unicellular green alga Chlamydomonas reinhardtii as an experimental system to study chloroplast RNA metabolism. Naturwissenschaften *87*, 97-107.

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. Science *289*, 920-930.

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

Olinares, P.D.B., Ponnala, L., and van Wijk, K.J. (2010). Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size sxclusion chromatography, mass spectrometry, and hierarchical clustering. Mol Cell Prot *9*, 1594-1615.

Ossenbühl, F., Hartmann, K., and Nickelsen, J. (2002). A chloroplast RNA binding protein from stromal thylakoid membranes specifically binds to the 5′ untranslated region of the *psbA* mRNA. FEBS J *269*, 3912-3919.

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

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

Perron, K., Goldschmidt-Clermont, M., and Rochaix, J.D. (2004). A multiprotein complex involved in chloroplast group II intron splicing. RNA *10*, 704-711.

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

Pfannschmidt, T., Brautigam, K., Wagner, R., Dietzel, L., Schroter, Y., Steiner, S., and Nykytenko, A. (2009). Potential regulation of gene expression in photosynthetic cells by redox and energy state: approaches towards better understanding. Ann Bot *103*, 599-607.

Piatigorsky, J. (1998). Gene sharing in lens and cornea: facts and implications. Prog Retin Eye *17*, 145-174.

Piatigorsky, J., O'Brien, W.E., Norman, B.L., Kalumuck, K., Wistow, G.J., Borras, T., Nickerson, J.M., and Wawrousek, E.F. (1988). Gene sharing by delta-crystallin and argininosuccinate lyase. Proc Natl Acad Sci USA *85*, 3479-3483.

Plader, W., and Sugiura, M. (2003). The Shine-Dalgarno-like sequence is a negative regulatory element for translation of tobacco chloroplast *rps2* mRNA: An additional mechanism for translational control in chloroplasts, Plant J *34* 377-82.

Plant, A., and Gray, J. (1988). Introns in chloroplast protein-coding genes of land plants. Photosynth Res *16*, 23-39.

Prikryl, J., Rojas, M., Schuster, G., and Barkan, A. (2011). Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. Proc Natl Acad Sci USA*108*, 415-420.

Qi, Y., Armbruster, U., Schmitz-Linneweber, C., Delannoy, E., de Longevialle, A.F., Ruhle, T., Small, I., Jahns, P., and Leister, D. (2012). Arabidopsis CSP41 proteins form multimeric complexes that bind and stabilize distinct plastid transcripts. J Exp Bot *63*, 1251-1270.

Rahire, M., Laroche, F., Cerutti, L., and Rochaix, J.-D. (2012). Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. Plant J *72*, 652-661.

Rao, S.T., and Rossmann, M.G. (1973). Comparison of super-secondary structures in proteins. J Mol Biol *76*, 241-256.

Reid, E.E., Thompson, P., Lyttle, C.R., and Dennis, D.T. (1977). Pyruvate dehydrogenase complex from higher plant mitochondria and proplastids. Plant Physiol *59*, 842-848.

Reyes-Prieto, A., Weber, A.P., and Bhattacharya, D. (2007). The origin and establishment of the plastid in algae and plants. Annu Review Gen *41*, 147-168.

Roy-Chaudhuri, B., Kirthi, N., and Culver, G.M. (2010). Appropriate maturation and folding of *16S* rRNA during 30S subunit biogenesis are critical for translational fidelity. Proc Natl Acad Sci USA *107*, 4567-4572.

Ruwe, H., and Schmitz-Linneweber, C. (2011). Short non-coding RNA fragments accumulating in chloroplasts: footprints of RNA binding proteins? Nucleic Acids Res *40*, 3106-3116.

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

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

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

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

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

Schottkowski, M., Peters, M., Zhan, Y., Rifai, O., Zhang, Y., and Zerges, W. (2012). Biogenic membranes of the chloroplast in *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA *109*, 19286-19291.

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

Schwenkert, S., Soll, J., and Bölter, B. (2011). Protein import into chloroplasts—How chaperones feature into the game. Biochim Biophys Acta *1808*, 901-911.

Sekine, K., Fujiwara, M., Nakayama, M., Takao, T., Hase, T., and Sato, N. (2007). DNA binding and partial nucleoid localization of the chloroplast stromal enzyme ferredoxin-sulfite reductase. FEBS J *274*, 2054-2069.

Sekine, K., Hase, T., and Sato, N. (2002). Reversible DNA compaction by sulfite reductase regulates transcriptional activity of chloroplast nucleoids. J Biol Chem *277*, 24399-24404.

Sengupta, S., Ghosh, S., and Nagaraja, V. (2008). Moonlighting function of glutamate racemase from *Mycobacterium tuberculosis*: racemization and DNA gyrase inhibition are two independent activities of the enzyme. Microbiol *154*, 2796-2803.

Shajani, Z., Sykes, M.T., and Williamson, J.R. (2011). Assembly of bacterial ribosomes. Annu Rev Biochem *80*, 501-526.

Sharwood, R.E., Hotto, A.M., Bollenbach, T.J., and Stern, D.B. (2011). Overaccumulation of the chloroplast antisense RNA *AS5* is correlated with decreased abundance of *5S* rRNA in vivo and inefficient *5S* rRNA maturation in vitro. RNA *17*, 230-243.

Shikanai, T. (2006). RNA editing in plant organelles: machinery, physiological function and evolution. Cell Mol Life Sci *63*, 698-708.

Small, I.D., and Peeters, N. (2000). The PPR motif - a TPR-related motif prevalent in plant organellar proteins. Trends Biochem Sci *25*, 46-47.

Smith , A.C., and Purton, S. (2002). The transcriptional apparatus of algal plastids, Eur J Phycol *37*, 301-311

Spalding, M.H. (2009). The CO2-Concentrating Mechanism and Carbon Assimilation. In The Chlamydomonas Sourcebook (Second Edition), D.B. Stern, and G.B. Witman, eds. (London: Academic Press), pp. 257-301.

Sriraman, P., Silhavy, D., and Maliga, P. (1998). Transcription from heterologous rRNA operon promoters in chloroplasts reveals requirement for specific activating factors. Plant Physiol *117*, 1495-1499.

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

Stoppel, R., Manavski, N., Schein, A., Schuster, G., Teubner, M., Schmitz-Linneweber, C., and Meurer, J. (2012). RHON1 is a novel ribonucleic acid-binding protein that supports RNase E function in the Arabidopsis chloroplast. Nucleic Acids Res *40*, 8593-606.

Stoppel, R., and Meurer, J. (2011). The cutting crew – ribonucleases are key players in the control of plastid gene expression. J Exp Bot *63*, 1663-73.

Suga, M., Akita, F., Hirata, K., Ueno, G., Murakami, H., Nakajima, Y., Shimizu, T., Yamashita, K., Yamamoto, M., Ago, H., *et al.* (2015). Native structure of photosystem II at 1.95 A resolution viewed by femtosecond X-ray pulses. Nature *517*, 99-103.

Sutoh, K., Kato, H., and Minamikawa, T. (1999). Identification and possible roles of three types of endopeptidase from germinated wheat seeds. J Biochem *126*, 700-707.

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

Tillich, M., Beick, S., and Schmitz-Linneweber, C. (2010). Chloroplast RNA-binding proteins: repair and regulation of chloroplast transcripts. RNA biol *7*, 172-178.

Tillich, M., Lehwark, P., Morton, B.R., and Maier, U.G. (2006). The evolution of chloroplast RNA editing. Mol Biol Evol *23*, 1912-1921.

Tovar-Méndez, A., Miernyk, J.A., and Randall, D.D. (2003). Regulation of pyruvate dehydrogenase complex activity in plant cells. Eur J Biochem *270*, 1043-1049.

Uniacke, J., and Zerges, W. (2008). Stress induces the assembly of RNA granules in the chloroplast of *Chlamydomonas reinhardtii*. J Cell Biol *182*, 641-646.

Uniacke, J., and Zerges, W. (2009). Chloroplast protein targeting involves localized translation in Chlamydomonas. Proc Natl Acad Sci USA *106*, 1439-1444.

Walter, M., Kilian, J., and Kudla, J. (2002). PNPase activity determines the efficiency of mRNA 3'-end processing, the degradation of tRNA and the extent of polyadenylation in chloroplasts. EMBO J *21*, 6905-6914.

Wang, F., Johnson, X., Cavaiuolo, M., Bohne, A.V., Nickelsen, J., and Vallon, O. (2015). Two Chlamydomonas OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome *b6f*. Plant J *82*, 861-873.

Weinert, B.T., Wagner, S.A., Horn, H., Henriksen, P., Liu, W.R., Olsen, J.V., Jensen, L.J., and Choudhary, C. (2011). Proteome-wide mapping of the Drosophila acetylome demonstrates a high degree of conservation of lysine acetylation. Sci Signal *4*, ra48.

Wellen, K.E., and Thompson, C.B. (2012). A two-way street: reciprocal regulation of metabolism and signaling. Nat Rev Mol Cell Biol *13*, 270-276.

Williams, P.M., and Barkan, A. (2003). A chloroplast-localized PPR protein required for plastid ribosome accumulation. Plant J *36*, 675-686.

Wireman, J.W., and Sypherd, P.S. (1974). Properties of 30S ribosomal particles reconstituted from precursor *16S* ribonucleic acid. Biochem *13*, 1215-1221.

Wistow, G.J., and Piatigorsky, J. (1988). Lens crystallins: the evolution and expression of proteins for a highly specialized tissue. Ann Rev Biochem *57*, 479-504.

Xing, S., and Poirier, Y. (2012). The protein acetylome and the regulation of metabolism. Trends Plant Sci *17*, 423-430.

Xiong, J., and Bauer, C.E. (2002). Complex evolution of photosynthesis. Annu Rev Plant Biol *53*, 503-521.

Yamamoto, Y.Y., Puente, P., and Deng, X.-W. (2000). An Arabidopsis cotyledon-specific albino locus: a possible role in *16S* rRNA maturation. Plant and Cell Physiology *41*, 68-76.

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

Yin, P., Li, Q., Yan, C., Liu, Y., Liu, J., Yu, F., Wang, Z., Long, J., He, J., Wang, H.-W., *et al.* (2013). Structural basis for the modular recognition of single-stranded RNA by PPR proteins. Nature *504*, 168-171.

Yohn, C.B., Cohen, A., Rosch, C., Kuchka, M.R., and Mayfield, S.P. (1998). Translation of the chloroplast *psbA* mRNA requires the nuclear-encoded poly(A)-binding protein, RB47. J Cell Biol *142*, 435-442.

Zghidi, W., Merendino, L., Cottet, A., Mache, R., and Lerbs-Mache, S. (2007). Nucleus-encoded plastid sigma factor SIG3 transcribes specifically the *psbN* gene in plastids. Nucleic Acids Res *35*, 455-464.

Zhan, Y., Dhaliwal, J.S., Adjibade, P., Uniacke, J., Mazroui, R., and Zerges, W. (2015). Localized control of oxidized RNA. J Cell Sci *128*, 4210.

Zhang, Z., Tan, J., Shi, Z., Xie, Q., Xing, Y., Liu, C., Chen, Q., Zhu, H., Wang, J., Zhang, J., *et al.* (2016). Albino leaf1 that encodes the sole octotricopeptide repeat protein is responsible for chloroplast development. Plant Physiol *171*, 1182-1191.

Zhelyazkova, P., Hammani, K., Rojas, M., Voelker, R., Vargas-Suárez, M., Börner, T., and Barkan, A. (2011). Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. Nucleic Acids Res *40*, 3092-105.

Zhou, Z.H., McCarthy, D.B., O'Connor, C.M., Reed, L.J., and Stoops, J.K. (2001). The remarkable structural and functional organization of the eukaryotic pyruvate dehydrogenase complexes. Proc Natl Acad Sci USA *98*, 14802-14807.

6 Appendix

The following section contains a manuscript in preparation and a supplemental figure supporting findings of Section 3.2.

6.1 Dynamic regulation of the proteome and lysine acetylome in Chlamydomonas reinhardtii responding to light and acetate

König, AC., Füßl, M,. Hartl, M., **Kleinknecht, L.**, Bohne, A.-V., Harzen, A., Kramer, K., Nickelsen, J. and Finkemeier, I.

The research in this study focuses on proteome and acetylome dynamics in the green algae *C. reinhardtii*. One major way of controlling protein functions and metabolism is by the use of post-translational modifications. They can act as fast and often reversible molecular switches. Here a mass spectrometry based analysis was performed to study the lysine acetylation and proteome dynamics in Chlamydomonas under different growth conditions. Liquid cultures were grown under mixotrophic (light and acetate), photoautotrophic (light only) or heterotrophic (dark and acetate) conditions before the MS analyses were carried out. Overall, 4.065 protein groups were identified, which carried a total of 254 acetylation sites. The integration of acetate seems to be mainly performed by the acetyl-CoA synthase 3 that is localized to the peroxisomes. Remarkably, lysine acetylation of nearly all the enzymes involved in the glyoxylate cycle were dynamically regulated within the different growth conditions. Additionally, it was shown that lysine acetylation has a functional relevance for the citrate synthase activity, which upon deacetylation after a treatment with a recombinant deactelyase from *E. coli*, revealed decreased activity. Overall, this study proves that lysine acetylation in *C. reinhardtii* is a dynamic process that is able to regulate the activity of metabolic enzymes.

My contribution to this study was to perform immunoblot analyses in collaboration with AC. König. Moreover, I revised the manuscript together with other co-authors.

Manuscript: Dynamic regulation of the proteome and lysine acetylome in *Chlamydomonas reinhardtii* responding to light and acetate

Ann-Christine König[a,b,#], Magdalena Füßl[a,c,#], Markus Hartl[a], Laura Kleinknecht[b], Alexandra Bohne[b], Anne Harzen[a], Katharina Kramer[a], Jörg Nickelsen[b], Iris Finkemeier[a,c,*]

[a] Plant Proteomics, Max Planck Institute for Plant Breeding Research, Carl von Linné Weg 10, DE-50829 Cologne, Germany

[b] Molecular Plant Sciences, Biocenter Martinsried, Ludwig-Maximilians University, Grosshaderner Strasse 2-4, DE-82152 Munich, Germany.

[c] Plant Physiology, Institute of Plant Biology and Biotechnology, University of Muenster, Schlossplatz 7, DE-48149 Muenster, Germany

[#] these authors contributed equally to this work

*To whom correspondence should be addressed:

Prof. Dr. Iris Finkemeier, Plant Physiology, Institute of Plant Biology and Biotechnology, University of Muenster, Schlossplatz 7, 48149 Muenster, Germany

Tel.: +49(0)251-8323805; Email: iris.finkemeier@uni-muenster.de

Author contributions: A.-C.K, M.F., M.H., L.K., A.H., A.B., K.K. performed research; A.-C.K., M.H., A.B., J.N., I.F. designed research; A.-C.K., M.F., M.H., and K.K. analyzed data; A.-C.K., J.N. and I.F. wrote the paper.

Running title: Proteome and acetylome dynamics in Chlamydomonas

Abbreviations:

AAA-metalloprotease (FTSH1)

Acetyl-CoA acyltransferase 1 (ATO1)

Acetyl-CoA synthetase 2 (ACS2)

Adenosylhomocysteinase (SAH1)

Alcohol/aldehyde dehydrogenase 1 (ADH1)

Aldehyde dehydrogenase 2 (ALDH2)

Apocytochrome f (PetA)

Aspartate aminotransferase (AST1)

ATP-synthase subunit alpha (AtpA)

ATP-synthase subunit b' (AtpX)

ATP-synthase subunit beta (AtpB)

ATP-synthase subunit delta (AtpD)

ATP-synthase subunit epsilon (AtpE)

Basic extraction buffer (BEB)

Basic extraction buffer containing deacetylase inhibitors (DI)

Calcium/calmodulin dependent protein kinase kinase 1 (CDPKK1)

Calcium/Calmodulin-dependent protein kinase kinase 1 (CDPKK1, A8IBS4)

Carbon dioxide (CO2)

Chlamydomonas reinhardtii (*C. reinhardtii*)

Chloroplastic pyruvate dehydrogenase complex (cpPDC)

Citrate synthase 2 (CIS2)

Dihydrolipoamide acetyltransferase (DLA2)

Fatty acid (FA)

Filter-aided sample preparation (FASP)

Glyceraldehyde 3-phosphate dehydrogenase dominant|minor splicing variant (GAP1a|GAP1b)

Glyceraldehyde-3-phosphate dehydrogenase (GAP3)

Glyoxylate cycle (GC)

Heat shock protein 70B (HSP70)

High-salt minimal medium (HSM)

Histone H3 (HTR14)

Hydroxymethylpyrimidine phosphate synthase (THICb/THICa)

Isocitrate lyase 1 (ICL1)

Label-free quantification (LFQ)

Light-harvesting protein of photosystem I (LHCA1)

Malate dehydrogenase 2 (MDH2)

Malate synthase 1 (MS1)

Mass spectrometry (MS)

Mg protoporphyrin IX S-adenosyl methionine O-methyl transferase (CHLM)

NAD-dependent protein deacetylase (YHDZ)

Oxygen-evolving enhancer protein 1 of photosystem II (PSBO)

Pentose phosphate pathway (OPPP)

Phosphoenolpyruvate (PEP)

Phosphoenolpyruvate carboxykinase (PCK1a/PCK1b)

Photosystem I P700 chlorophyll a apoprotein A2 (PsaB)

Photosystem I reaction center subunit II (PsaD)

Post-translational modifications (PTMs)

Pyruvate (PYR)

Pyruvate carboxykinase (PCK1a/PCK1b)

Pyruvate phosphate dikinase (PPD1)

Pyruvate: ferredoxin oxidoreductase (PFR1)

Ribosomal protein S11 (Rps11)

Ribosomal protein S4 (Rsp4)

Ribulose-1,5- bisphosphate carboxylase large chain (RBCL)

Solute carrier family 25 (SLC25)

Transketolase (TRK1)

Tricarboxylic acid cycle (TCA)

Tris/acetate/phosphate (TAP)

Tris/acetate/phosphate medium containing 1% sorbitol (TAPS)

UDP-Glucose:protein transglucosylase (EZY11)

Summary

The soil-dwelling green algae *Chlamydomonas reinhardtii* is one of the most studied microorganisms concerning the production of renewable alternative energy. To understand its metabolic regulation upon variable environmental conditions is a major task in research of this century. Post-translational modifications have a key role in acting as molecular switches for the control of protein functions and metabolism. However, the post-translational control of enzyme functions in Chlamydomonas, except for phosphorylation, is largely unexplored. Acetylation of the ε-amino acid of a lysine residue is a dynamic and major modification coupled to the energy metabolism of a cell as it is dependent on acetyl-CoA levels. Therefore, we performed a mass spectrometry based analysis to study lysine acetylation and proteome dynamics in *C. reinhardtii* under varying growth conditions. Liquid cultures of Chlamydomonas were transferred from mixotrophic (light and acetate as carbon source) to heterotrophic (dark and acetate), or photoautotrophic (light only) growth conditions for 30 h before harvest. In total, 4,065 protein groups were identified with a protein FDR <1 %, which carried 254 lysine acetylation sites. The proteome and acetylome changes between the different growth conditions were quantified using dimethyl-labelling. The presence of lysine acetylation on nearly all enzymes involved in the glyoxylate cycle being dynamically regulated within the different growth conditions was one of the major results of this study. Acetate seems to be predominantly integrated by the acetyl-coA synthase 3 which is localized to the peroxisomes. A functional relevance of lysine acetylation could be shown for the citrate synthase activity which exhibits less activity after treatment with a recombinant deacetylase from *Escherichia coli*. Our results show that lysine acetylation is a dynamic process with the potential of genetic engineering possibilities for future alternative energy supply.

Introduction

One of the main challenges of this century is the rising demand of energy in combination with the need to prevent damages to the environment. Therefore, environmentally sustainable fuels and energy have a huge potential as alternatives to the available fossil fuels. Especially algae possess abilities in generating energy products, such as bio-oil, methane, methanol, and hydrogen (Jones and Mayfield, 2012; Velasquez-Orta et al., 2009); hence they have been the target of genetic manipulation over the last 70 years. The

soil-dwelling green algae *Chlamydomonas reinhardtii* (*C. reinhardtii*) is the most studied microorganism in the production of renewable alternative energy (Leite et al., 2013; Merchant et al., 2012). Using chloroplastic and nuclear engineering as well as homologous recombination and high-throughput screening, an improvement of characteristics for biofuel production has been achieved over the last years (Scranton et al., 2015). Still there is a huge gap of knowledge in the basic research of underlying biochemistry, cell biology, and genetics of *C. reinhardtii*. Hence, to improve the idea of using *C. reinhardtii* as an alternative energy source for biofuels the main task is to understand these basic principles of metabolic functions and pathways.

C. reinhardtii is able to grow photoautotrophically with just light as an energy source, as well as heterotrophically with acetate as carbon source. In addition, *C. reinhardtii* is often grown in a mixture between both conditions (acetate and light), which is named mixotrophic growth (Hoober, 1989). Due to different metabolic pathways that have to be activated in dependence on the growth conditions, such as photosynthesis in the light, several metabolic adaptations have to occur in *C. reinhardtii* to prevent any futile cycling in metabolism. Such acclimation to growth conditions can occur at several levels which include transcriptional and translational control of mRNA as well as protein stability and product inhibition or allosteric effects on enzyme activities (Erickson et al., 2015; Hoober, 1989; Ledford et al., 2007). Another layer of regulation of protein functions is mediated by post-translational modifications (PTMs). PTMs are often reversible and can act as fast molecular switches which enable proteins to change their activities, functions or even localizations in the cell (Castano-Cerezo et al., 2014; Jing et al., 2013). Recent advances in mass spectrometry (MS)-based proteomics allowed the identification of thousands of different PTMs (Choudhary and Mann, 2010; Jensen, 2006; Macek et al., 2009). Within these PTMs the lysine residue is the most modified amino acid with more than 200 different types of modifications (Hershko and Ciechanover, 1998; Martin and Zhang, 2005; Weinert et al., 2013; Zhang et al., 2009). One of these more prominent lysine modifications is lysine acetylation which occurs on the ε-amino group of lysine residues. Lysine acetylation has been studied in much detail on histone proteins where it is linked to transcriptional regulation by changing the interaction of transcription factors with chromatin (Sterner and Berger, 2000). The first non-nuclear lysine acetylation protein discovered was the alpha-tubulin of Chlamydomonas axonemal microtubules more than 20 years ago (L'Hernault and Rosenbaum, 1983, 1985). From then on, the discovery of lysine acetylation on non-histone proteins gradually increased with

improvements in the enrichment techniques as well as with the sensitivity and scanning-speed and resolution of mass spectrometers. Lysine acetylation is nowadays known to occur in various organisms and subcellular localizations in bacteria, yeast, plant and animal cells (Choudhary et al., 2009; Finkemeier et al., 2011; Henriksen et al., 2012; Lundby et al., 2012; Melo-Braga et al., 2012; Mo et al., 2015; Weinert et al., 2011). Beside the total number of lysine acetylation within different organism the differential analysis of lysine acetylation between different genotypes, mutants as well as growth conditions is of upcoming interest in particular to learn more about the functional relevance of lysine acetylation (Yan and Chen, 2005). In this work, a stable isotope dimethyl labeling technique was used to compare the dynamics of lysine acetylation and the proteome between the different growth conditions of *C. reinhardtii* (Boersema et al., 2009). Since lysine acetylation is dependent on the presence of acetyl-CoA, the regulation of the proteome by lysine acetylation represents a tight connection between the metabolic state and the resulting metabolic pathway regulation (Choudhary et al., 2014; Galdieri et al., 2014; Glasser et al., 2014). Hence, this study gives a novel insight into the functional relevance of lysine acetylation for *C. reinhardtii* transferred for 30h from mixotrophic growth conditions to light, light and acetate and only acetate as a carbon source, respectively.

Materials and Methods

Algal Strain and Culture Conditions

For *C. reinhardtii*, we used the cell wall deficient mt- strain CC-3491 (Chlamydomonas Resource Center) which showed a reduced abundance of acetylated tubulin due to a high proportion of non-flagellated cells. The strain was maintained on 0.8% agar-solidified Tris/acetate/phosphate (TAP) medium (Harris et al. 1989) at 25°C under constant light (30 µmol/m^2/s). Liquid cultures were incubated under agitation at 25°C. For analysis, a pre-culture was grown in 2 L of TAP medium containing 1% sorbitol (TAPS) to a density of ~5·10^6 cells/mL. Cells were harvested by centrifugation (5 min, RT, 1000 x g), washed once with 200 mL high-salt minimal medium (HSM) ((Hoober, 1989)) and resuspended in 120 mL HSM (~8 x 10^7 cells/mL). This suspension was used to inoculate the following cultures: For heterotrophic growth, 3 x 10 mL resuspended culture was used to inoculate 3 x 300 mL TAPS, for mixotrophic growth, 3 x 8 mL was used to inoculate 3 x 300 mL TAPS, and for growth under photoautotrophic conditions, 3 x 15 mL were used to inoculate 3 x 500 mL

HSM. For mixotrophic and photoautotrophic growth cells were then incubated in the light (100 µE/m^2/s) or kept in complete darkness for heterotrophic growth.

Immunoenrichment and Western blot analysis

For native protein isolation, cells were grown under indicated light and time conditions as described above. Cells were harvested by centrifugation and lysed under pipetting in 2ml basic extraction buffer (BEB) containing 50 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol, 2mM EDTA, 0.5% Triton X-100, as well as 5 mM dithiothreitol and protease inhibitor cocktail (cOmplete Tablets, Roche). Additionally, to avoid deacetylation of proteins, 2µg mL^{-1} apicidin and 1mM nicotinamide were added to the BEB(DI). For lysine acetylation immunoenrichment, 200µg protein extract was incubated with 20µl acetyl lysine antibody immobilized to agarose beads (ImmuneChem Pharmaceuticals) for 3h at 4°C. Immunoprecipitates were washed three times with extraction buffer and eluted by boiling in gel loading buffer for 5 min. For Western-blot analysis, proteins were separated by SDS-PAGE, transferred to a nitrocellulose membrane, and probed using acetyl lysine antibody in a 1:1,000 dilution (ImmuneChem Pharmaceuticals). Secondary anti-horseradish peroxidase antibody was used in a 1:10,000 dilution.

Protein extraction, Filter-aided sample preparation (FASP) and trypsin digestion

Protein pellets from *C. reinhardtii* were extracted in 10ml heated SDT-lysis buffer containing 4% (w/v) SDS, 100 mM Tris/HCl pH 7.6 and 100 mM DTT for 10min at 95°C with occasional mixing followed by 15min of sonication. Protein extracts were cleared by centrifugation and the protein amount was determined using the 660 nm Pierce protein assay with compatibility reagent (Thermo Scientific) as described before in detail (Hartl et al., 2015).

To remove excess SDS and to prepare the sample for tryptic digestion, a FASP was used as explained in (Wisniewski et al., 2009). Briefly, 10mg of the protein extract were diluted with 8 M urea in 100 mM Tris/HCl (pH 8) until a SDS concentration of <0.5 % was reached, loaded on a filter (Amicon Ultra-15 centrifugal filter units, Millipore) and washed three times until SDS was completely removed. The extract was alkylated using 50 mM iodoacetamide for 30min in the dark and excess reagent was washed through the filter and replaced by 50 mM NH_4HCO_3. The reduced and alkylated proteins were digested using

trypsin in an enzyme to protein ratio of 1:100. Peptides obtained by FASP were quantified at 280 nm.

Enrichment of lysine acetylated peptides

Fractionated peptides were pooled together to obtain a total number of six fractions. Dried peptides were dissolved in 1 ml TBS buffer (50 mM Tris/HCl pH 7.6, 150 mM NaCl) and the peptide concentrations were determined. From each fraction a sample for the analysis of the total proteome was taken. 50 µl acetyl lysine antibody beads (ImmuneChem Pharmaceuticals) were used per mg peptide. Eluted peptides were desalted as described in König (2014).

MS data acquisition

Dried peptides were redissolved in 2% ACN, 0.1% TFA for analysis. Total proteome samples were adjusted to a final concentration of 0.2 µg/µl. Samples were analyzed using an EASY-nLC 1000 (Thermo Fisher) coupled to a Q Exactive Plus mass spectrometer (Thermo Fisher). Peptides were separated on 16 cm frit-less silica emitters (New Objective, 0.75 µm inner diameter), packed in-house with reversed-phase ReproSil-Pur C18 AQ 3 µm resin (Dr. Maisch). Peptides (1 µg for total proteome samples, half of the eluate for acetylated peptides) were loaded on the column and eluted for 120 min using a segmented linear gradient of 0% to 95% solvent B (solvent A 5% ACN, 0.5% FA; solvent B 100% ACN, 0.5% FA) at a flow-rate of 250 nL/min. Mass spectra were acquired in data-dependent acquisition mode with a Top15 method. MS spectra were acquired in the Orbitrap analyzer with a mass range of 300–1750 *m/z* at a resolution of 70,000 FWHM and a target value of $3x10^6$ ions. Precursors were selected with an isolation window of 1.3 m/z. HCD fragmentation was performed at a normalized collision energy of 25. MS/MS spectra were acquired with a target value of 10^5 ions and an intensity threshold of $7.3x10^5$ (acetylated peptides) or $1.1x10^5$ (total proteome samples) at a resolution of 17,500 FWHM and a fixed first mass of *m/z* 100. Peptides with a charge of +1, greater than 6, or with unassigned charge state were excluded from fragmentation for MS2, dynamic exclusion for 30 s prevented repeated selection of precursors.

MS data analysis

Raw data were processed using MaxQuant software (version 1.5.2.8, http://www.maxquant.org/) (Cox and Mann, 2008)with label-free quantification (LFQ) and iBAQ enabled (Cox et al., 2014). MS/MS spectra were searched by the Andromeda search engine (integrated in MaxQuant 1.5.2.8) against the UniProt Chamydomonas reference proteome (www.uniprot.org; database retrieved Aug 2012; 14,337 entries). Sequences of 248 common contaminant proteins and decoy sequences were automatically added during the search. Data of total proteome and acetylated lysine-enriched samples were separated into two parameter groups to permit combined analysis. The match between runs option was enabled, but fractions of total proteome samples were defined as 1 to 7 and those of enriched samples as 11 to 17 to prevent transferal of identifications between both sample groups. Dimethylation of lysines and peptide N-termini were set as light (H_4C_2), medium (D_4C_2), and heavy ($-H_2+D_6{}^{13}C_2$) labels; the re-quantify option was enabled. Trypsin specificity was required and a maximum of two or four missed cleavages allowed for total proteome and enriched samples, respectively. Minimal peptide length was set to seven amino acids. Carbamidomethylation of cysteine residues was set as fixed, oxidation of methionine and protein N-terminal acetylation as variable modifications. Acetylation of lysines was added as a variable modification for the anibody-enriched samples. Allowed mass deviation was 4.5 ppm for peptides and 20 ppm for fragments. The minimum score and delta score for modified peptides were filtered for a score higher than 35 and 6, respectively. Peptide-spectrum-matches and proteins were retained if they were below a false discovery rate of 1%. Subsequent quantitative statistical analyses were performed in Perseus (version 1.5.2.6, http://www.maxquant.org/; Cox and Mann 2012).

Citrate synthase activity

Protein extraction was carried out in BWB followed by desalting with PD-10 columns (GE Healthcare). The protein extract was split and either treated with DI or with a recombinant NAD-dependent protein deacetylase (YHDZ) for 2h at room temperature (Seidel et al., 2016). Subsequently the samples were alkylated with 20 mM iodoacetamide for 30min in the dark. The citrate synthase activity was measured spectrophotometrically as described previously (Anoop et al., 2003; Schmidtmann et al., 2014). The assay was based on the absorbance of DTNB after reaction with CoA at 412nm. 100µg of the prepared protein extract were incubated with 0.5 mM acetyl-CoA (AppliChem) in 1 mM DTNB (in 100 mM Tris–HCl pH 8.0) and the reaction was started after addition of 10 mM oxaloacetate.

Results

Lysine acetylation in Chlamydomonas reinhardtii in dependence on light and acetate

In order to analyze the impact of the growth condition on the lysine acetylation status of *C. reinhardtii*, a pre-culture of CC-3491 cells was first grown under mixotrophic conditions and a medium light intensity (30 µmol/m^2/s). The cells were then washed with acetate-free media and used to inoculate the main cultures which were grown either under hetero-, photoauto- or mixotrophic conditions for 30 h (Fig. 1A). The light intensity of the photoautotrophic and mixotrophic growth conditions was increased to 100 µmol/m^2/s to support photosynthesis. After 30 h of growth under the respective condition, the pelleted cells were harvested frozen and stored for further analysis. We selected the 30 h time point, since there was a maximum difference in the acetylation status of a putative histone protein (below 29 kDa) visible on the Western-blot using anti-lysine acetylation between the light and dark grown cultures (Fig. 1B). To detect whether Chlamydomonas possesses more lysine acetylated proteins, an enrichment of soluble lysine acetylated proteins was performed before Western blot analysis by using anti-lysine acetylation agarose beads (Fig. 1C). After this enrichment step, many more lysine acetylated proteins were detected on the Western-blot, such as a prominent protein band between 45 kDa and 66 kDa, which most likely refers to the ribulose-1,5- bisphosphate carboxylase large chain (RBCL). It was previously shown that the RBCL is highly acetylated in Arabidopsis (Finkemeier et al., 2011). No histone proteins could be detected by this approach, probably because the immunoprecipitation buffer did not solubilize the nuclear envelope during protein extraction.

Mass spectrometry-based profiling of the Chlamydomonas proteome and acetylome

The variety of lysine acetylated proteins detected on the Western blot prompted us to profile the lysine acetylome of *C. reinhardtii* in more detail, and to quantify total proteome and acetylome changes under the different growth conditions in a large-scale MS-based shot-gun proteomic approach combined with a stable isoptopic dimethyl-labelling strategy. An overview of the mass-spectrometry-based work flow is shown in Fig. 2 A. Samples from the three different growth conditions were analysed in three biological replicates. Total proteins from each replicate were extracted, digested with trypsin, and free-amino groups of peptides were labeled with light, medium, and heavy stable isotope dimethyl-forms, respectively (Boersema et al., 2009; Konig et al., 2014). To prevent any labeling bias, a label-swap of the

light, medium, and heavy dimethyl-forms was performed on the third replicate. Equal amounts of the labeled peptides from the three growth conditions were combined and subjected to ZIC-HILIC fractionation to reduce the sample complexity. After fractionation the samples were enriched for lysine acetylated peptides by immunoaffinity purification. All fractions were analysed on a Q-Exactive Plus (Thermo Scientific) mass spectrometer and the resulting raw files were processed using MaxQuant. Protein and lysine acetylation peptide ratios for every peptide detected in at least two biological replicates were subjected to a LIMMA statistical analysis using R (Ritchie et al., 2015; Smyth et al., 2005).

First, the proteome changes between the different growth conditions were compared (Fig. 2B-D; Suppl. table 1-3). In total, 4065 protein groups were identified in at least one replicate with a protein FDR of <1%. Overall, the different growth conditions affected only a subset of proteins, which have specific physiological functions important under the respective condition. Under heterotrophic growth conditions there were two enzymes, the pyruvate: ferredoxin oxidoreductase (PFR1) and the phosphoenolpyruvate carboxykinase (PCK1a/PCK1b), which were both highly increased in abundance (2.5 to 16-fold) compared to mixotrophic and photoautotrophic growth conditions. PFR1 is part of the fermentative pathways involved in dark anaerobic hydrogen (H_2) production (Noth et al., 2013), while the pyruvate carboxykinase (PCK1a/PCK1b) is part of the gluconeogenesis pathway (Subramanian et al., 2014). Under mixotrophic conditions, there was only one enzyme, the hydroxymethylpyrimidine phosphate synthase (THICb/THICa), which was particularly more abundant than compared to heterotrophic (8-fold) and photoautotrophic (3-fold) growth conditions, respectively. THIC is involved in thiamine synthesis, which is a cofactor of several central metabolic enzymes (Moulin et al., 2013). As expected, under photoautothrophic and mixotrophic conditions, mainly proteins related to carbon dioxide (CO_2) fixation and important for a proper function of the Calvin-Benson cycle were increased in abundance compared to heterotrophic growth conditions (Winck et al., 2013). This included several low CO_2-inducible proteins as well as RuBisCO activase and light harvesting proteins, for example. In addition to the light-regulated proteins, we were able to identify acetate-dependent proteins, since both mixotrophic and heterotrophic conditions contained acetate in the growth medium. One example for an acetate-regulated protein is the SLC25-like protein that belongs to the peroxisomal solute carrier family. This carrier is supposedly responsible for the transport of nucleotides into the peroxisome like CoA, FAD, and NAD^+, which are required for the assimilation of acetate (Agrimi et al., 2012).

Identification of lysine acetylation sites and Functional Annotation of Acetylated Proteins in C. reinhardtii

In total, 254 lysine acetylation sites (FDR < 1 %, score > 30, delta score > 6) were detected on 149 protein groups of Chlamydomonas. To gain some insight into the biological functions of the lysine-acetylated proteins, a functional annotation classification using the UniProtKB database (http://www.uniprot.org/) was performed (Fig. 3A). Interestingly, 20 % of the 149 protein groups are involved in protein biosynthesis and most of the proteins belong to cytosolic as well as chloroplastic ribosomes (Suppl Tab. 3). Proteins that play a role in photosynthesis including the Calvin cycle reaction represent 17 % of the lysine acetylated proteins. Five of these proteins belong to the chloroplastic ATP-synthase including the ATP-synthase subunit b' (AtpX), the ATP-synthase subunit alpha and beta (AtpA and AtpB), as well as the epsilon and delta subunit (AtpE and AtpD). Interestingly, the same subunits were also discovered as lysine acetylated in *Synechocystis sp.* PC 6803 (Mo et al., 2015). In several other organisms, the mitochondrial homolog of the ATP-synthase in the respiratory chain was discovered to possess several lysine acetylated proteins (Henriksen et al., 2012; Konig et al., 2014; Lundby et al., 2012; Rardin et al., 2013; Weinert et al., 2011). Hence, a functional relevance of this modification for the ATP Synthase protein therefore seems to be likely (Hosp et al., 2017). For example, Liko and co-workers recently discovered that the highly conserved acetylation sites within the cytochrome c oxidase protein interact with membrane lipids and therefore could be responsible for the fine tuning of protein-lipid interactions (Liko et al. 2016). About 9 % of the lysine acetylated proteins of Chlamydomonas belong to the functional category of carbon metabolism which includes glycolysis, tricarboxylic acid cycle (TCA) cycle, starch synthesis, pentose phosphate pathway (OPPP), and the glyoxylate cycle (GC). Interestingly, almost every enzyme of the GC (acetyl-CoA synthetase 2 (ACS2), isocitrate lyase 1 (ICL1), malate dehydrogenase 2 (MDH2), malate synthase 1 (MS1), and citrate synthase 2 (CIS2)) contains acetylated lysine residues (Suppl Tab. 3). Only the aconitase was not detected as lysine acetylated, but also no unmodified peptides were detected for this protein in our study. Interestingly almost all enzymes of the GC pathway were found to be differentially lysine acetylated under the different growth conditions, which will be discussed in the following section. Another 7 % of the lysine acetylated proteins belong to the category of histones and DNA metabolism, among those were the four histone proteins H2A, H2B, H3 and H4, as well as DNA-directed RNA-polymerase and subunit 1 of the DNA-replication factor C complex. Interestingly, 5 % of the lysine acetylated proteins fall into the

category of proteins involved in fatty acid (FA) and acetyl-CoA metabolism. These enzymes are in direct contact with acetyl-CoA, which is used as substrate for lysine acetylation. Five different lysine acetylation sites (K35, K205, K210, K218, K663) were identified on the Chlamydomonas acetyl-CoA synthetase 3 (ACS3, A8JFR9). K609 of the acetyl-CoA synthetase in *Salmonella enterica* was previously shown to be actively regulated by lysine acetylation which inhibits the activity of this enzyme (Starai et al., 2002). However, K609 of Salmonella which is conserved in ACS3 of Chlamydomonas was not identified as lysine acetylated in our study. The functional group of kinases comprised 5 % of the lysine acetylated proteins. This group consists of diverse group of kinases involved in several metabolic processes and signalling, such as the ATM/ATR-like kinase as part of the histone acetyltransferase complex (A8I8Y6) and a Calcium/Calmodulin-dependent protein kinase kinase 1 (CDPKK1, A8IBS4). Furthermore, 13 % of the lysine acetylated proteins comprise a group of predicted proteins, and 23 % of the proteins are grouped in a category miscellaneous, which contains several interesting regulatory proteins and enzymes such as a superoxide dismutase, heat shock proteins and several proteases for example (Suppl Tab. 3). To investigate whether specific sequence motifs can be found around the lysine acetylation sites, the IceLogo tool was used to generate a sequence logo using the 10 amino acids on each site of the lysine residue (Colaert et al., 2009). Whereas no specific pattern could be detected, an accumulation of glycine and lysine residues was observed for all three growth conditions.

Differential lysine acetylation between the different growth conditions.

To investigate the changes of lysine acetylation between heterotrophic, photoautotrophic and mixotrophic conditions we used the fact that peptides of each growth condition were labelled with a different dimethyl label (Fig. 2A). Therefore lysine acetylation of heterotrophic/photoautotrophic, heterotrophic/mixotrophic and mixotrophic/photoautotrophic was compared and visualized by volcano plots (Fig.4A-C). lysine acetylated peptides which have a $-\log_{10}$ p-value more or equal to 1.3 and a $\log_2$FC by more or equal to 1, were considered to be significant regulated between the different growth conditions. Positive $\log_2$FC numbers indicate an increased acetylation for (A) heterotrophic compared to photoautotrophic; (B) Hetero compared to mixotrophic and (C) mixotrophic compared to photoautotrophic conditions, whereas negative $\log_2$FC show lower levels of acetylation for the mentioned conditions.

By having a closer look at the lysine acetylated sites which were upregulated under heterotrophic conditions compared to photoautotrophic conditions, 32 lysine acetylated sites were identified which belong to 19 unique proteins. Several of these lysine acetylated sites belong to proteins which are localized to the peroxisome and are part of the GC pathway as mentioned before. The ACS3 provides one entry to the GC pathways by producing acetyl-CoA by converting acetate and was recently localized to the peroxisome (Lauersen et al., 2016). ACS3 has two lysine acetylation sites which were increased under heterotrophic conditions compared to photoautotrophic, ACS3K205 (4.04$\log_2$FC) and ACS3K218 (3.67 $\log_2$FC) (Fig. 2A). Acetyl-CoA is then converted to citrate by the CIS2 that carries three up-regulated lysine acetylation sites which are CIS2K100 (3.85$\log_2$FC); CIS2K341 (4.36 $\log_2$FC) and CIS2K447 (3.76$\log_2$FC). The next enzyme in the GC, the ACO which produces isocitrate out of citrate, was not detected in our experiment. Isocitrate is then converted to succinate by the ICL1 which has one lysine acetylation, ICL1K195 upregulated by 2.0 $\log_2$FC. Subsequently, succinate is converted to malate by the MS1 which has one lysine acetylation MS1K422 increased by 4.07 $\log_2$FC. The MDH metabolises malate to oxaloacetate and carries two lysine acetylation sites at position MDHK176 and MDHK342 that were not significantly regulated between heterotrophic and photoautotrophic conditions. Beside the known enzymes of the GC pathway several predicted proteins with lysine acetylation sites were identified with a potential function in ß-oxidation as well as peroxisomal in FA synthesis. *A8JBL6* (Pp1) has a predicted 3-hydroxyacyl-CoA dehydrogenase activity and is supposed to be involved fatty acid metabolic process (Pp1K305, 4.96$\log_2$FC; Pp1K553, 3.44$\log_2$FC; Pp1K565, 3.87$\log_2$FC). Additionally two more peroxisome related proteins with lysine acetylated sites were discovered. *A8IYU9* (Pp3) a peroxisome fission related protein (Pp2K55, 3.83$\log_2$FC; Pp2K116, 4.62$\log_2$FC) and *A8IMQ3* (Pp2) a peroxisomal multifunctional enzyme type 2 protein (Pp3K142, 4.77 $\log_2$FC). Interestingly the acetyl-CoA acyltransferase 1 (ATO1) is as well related to peroxisomes as it carries an incomplete peroxisomal targeting sequence and it was related to ß-oxidation (Atteia et al., 2009; Goodenough et al., 2014). ATO1 has three highly increased lysine acetylation sites under heterotrophic conditions, ATO1K230, 3.05$\log_2$FC; ATO1K232, 3.79$\log_2$FC; ATO1K274, 3.08$\log_2$FC and ATO1K457, 4.25$\log_2$FC. Therefore out of the 19 unique proteins with a significant increase in lysine acetylation under heterotrophic conditions 8 are related to different peroxisomal functions. The other proteins belong to diverse localizations and categorical functions like the RbcL which is chloroplast localized and carries a lysine

acetylation site at K175 2.54log_2FC. As mentioned before the influence of lysine acetylation on the RUBISCO was already known and the increase in lysine acetylation during heterotrophic conditions is in concert with the reported decrease in activity (Finkemeier et al., 2011; Gao et al., 2016). Also one of the already mentioned ATP-synthase subunit, AtpB showed increased lysine acetylation under heterotrophic conditions at K215 with a log_2FC of 1.25. The dihydrolipoamide acetyltransferase (DLA2) is a subunit of the chloroplastic pyruvate dehydrogenase complex (cpPDC) and is involved in the conversion from pyruvate to acetyl-coA as well as in the translation of the *psbA* mRNA from photosystem II (Bohne et al., 2013). DLA2 carries in total three lysine acetylation sites including one that is significantly upregulated under heterotrophic conditions (DLA2K161, 1.56 log_2FC). Interestingly, parts of the glycolytic pathway of *C.reinhardtii* take place in the chloroplast (Johnson and Alric, 2013). The Glyceraldehyde-3-phosphate dehydrogenase is part of the glycolysis and one lysine is upregulated under heterotrophic conditions (GAP3K162, 1.42log_2FC). The last chloroplast protein with an increased lysine acetylation under heterotrophic conditions was the alcohol/aldehyde dehydrogenase 1 (ADH1) which converts acetyl-coA to ethanol (ADH1K206, 2.31log_2FC) and is involved in the fermentative pathways (Catalanotti et al., 2013; Terashima et al., 2010). Within the cytosol one different lysine acetylated protein was identified with elevated lysine acetylation sites the Ribosomal protein L7 (RPL7K57, 1.24log_2FC). Likewise in mitochondria one protein showed changed acetylation, the aspartate aminotransferase (AST1K66, 3.99log_2FC) which is responsible for the conversion of aspartate and α-ketogluterate to glutamate and oxaloacetate (Goodenough et al., 2014). The remaining proteins could not be mapped to any localization.)The glyceraldehyde 3-phosphate dehydrogenase has two different splicing variants (GAP1a|GAP1b) that cannot be distinguished by MS. Therefore the dominant splicing variant GAP1a has a lysine acetylation at position K257 which correlates with the lysine acetylation of minor splicing variant GAP1b at position K204 which is the same peptide and therefore these two lysine acetylation sites have the same log_2FC which is 2.18. Identical behaves GAP1aK199 and GAP1bK146 with a log2FC of 1.98. The malonyl-CoA:acyl-carrier-protein transacylase (2*) and Pp4 have each one increased lysine acetylation site, 2*K313 1.23log_2FC and Pp4 1.47K304 log_2FC. Beside changes in lysine acetylation also the changes in protein levels between heterotrophic and photoautotrophic conditions were distinct and visualized in Fig. 4D in which the log_2FC of the heterotrophic/photoautotrophic protein ration is plotted against the log_2FC of lysine acetylation heterotrophic/photoautotrophic ratio. Interestingly mainly proteins which are

related to peroxisomes were upregulated on protein level under heterotrophic conditions including the ACS3 (2.94log_2FC), CIS2 (2.5log_2FC), Pp2 (2.23log_2FC), MS1 (2.17log_2FC), ICL1 (2.14log_2FC), Pp1 (1.48log_2FC) and ATO1 (1.42log_2FC). This could be due to the fact that the GC is necessary during darkness because during longer periods of darkness plants start to degrade their endomembrane system and depend on fatty acid β-oxidation (Kunze et al., 2006). Additionally it is known that the number of peroxisomes increases with acetate in the media (Hayashi et al., 2015). Another three proteins were significant more upregulated which were so far not related to peroxisomes, the AST1 (2.11log_2FC), ADH1 (2.08log_2FC) and GAP1a|GAP1b (1.95log_2FC). By having a look at the lysine acetylation sites which were down regulated under heterotrophic conditions and therefore upregulated under photoautotrophic conditions, seven unique proteins were discovered. Five of these seven where associated with chloroplast function, Apocytochrome f (PetAK219, -2.58log_2FC), a 3-oxoacyl-ACP like protein (1*K280,-1.88log_2FC), Mg protoporphyrin IX S-adenosyl methionine O-methyl transferase (CHLMK271, -1.81log_2FC), Photosystem I reaction center subunit II (PsaDK114, -1.67log_2FC) and Photosystem I P700 chlorophyll a apoprotein A2 (PsaBK8, -1.63log_2FC; PsaBK4 -1.19log_2FC). Furthermore a Calcium/calmodulin dependent protein kinase kinase 1 (CDPKK1-1.24log_2FC) with unknown localization was identified. Interestingly GAP3 which also carries an upregulated lysine acetylation site also exhibits a lysine acetylation site on position K205 which is -1.3log_2FC downregulated. On total protein level just three proteins have a significant decrease, PsaD (-1.51log_2FC), PsaB (-1.5log_2FC) and CHLM (-1.1log_2FC) (Fig. 2D).

In the next step heterotrophic versus mixotrophic conditions were compared on lysine acetylation level (Fig. 4B). In this case it was interesting that lysine acetylation changes were in general not as distinct between the two conditions when compared with heterotrophic against photoautotrophic. This could be a hint that the major lysine acetylation changes are provoked by acetate. The acetate dependent changes cannot be seen between heterotrophic and mixotrophic since both growth conditions contain acetate in the media therefore the visible lysine acetylation changes depend solely on dark versus light conditions. This can explain that the -log_{10}p-values were not as defined, as well. Under heterotrophic conditions in total 22 lysine acetylation sites were significantly upregulated compared to mixotrophic which belong to 16 unique protein groups. Also in this comparison 8 out of 16 proteins belong to proteins involved in peroxisomal metabolism. Again several enzymes of the GC were lysine acetylated, the ACS3 (ACS3K205, 1.34 log_2FC and K218, 1.31log_2FC), CIS2 (CIS2K100,

1.7log$_2$FC; K341, 3.69log$_2$FC, CISK447, 1.43log$_2$FC) ICL1 (ICL1K195, 1.93log2FC) and MS1 (MS1K422, 2.82log$_2$FC). The FA related protein Pp1 showed likewise an upregulation of lysine acetylation sites but clearly less increased when compared with heterotrophic versus photoautotrophic conditions (Pp1K305, 2.31log$_2$FC; Pp1K553, 1.68log$_2$FC; Pp1K565, 1.24log$_2$FC). The same is true for Pp3 which had just one lysine acetylation site which was significantly regulated (Pp3K116, 1.37log$_2$FC), ATO1 (ATO1K232, 1.21log$_2$FC) as well as for Pp2 (Pp2K142, 3.29 log$_2$FC). Concerning proteins localized to the chloroplast just three proteins showed increased lysine acetylation. These were again the RbcL K175 2.54log$_2$FC, as well as a nucleoside diphosphate kinase (3*K49, 1.19log$_2$FC) and GAP3 (GAP3K162, 1.1log$_2$FC). The mitochondrial AST1 showed as well an upregulation in lysine acetylation (AST1K66, 1.88log2FC). The last three proteins could not be assigned to any localization, UDP-Glucose:protein transglucosylase (EZY11K312, 1.92og$_2$FC), Pp5 (Pp5K32, 1.81og$_2$FC) and again the GAP1a/GAP1b splicing variants (GAP1aK257|GAP1bK204, 1.2log$_2$FC). On total proteome level (Fig. 4E) similar aspects were seen for heterotrophic/mixotrophic conditions as for heterotrophic/photoautotrophic conditions (Fig. 4D) but not as strongly pronounced. Again especially proteins of the peroxisomal metabolism show an upregulation on protein level, ACS3 (1.52log$_2$FC), CIS2 (1.46log$_2$FC), Pp3 (1.08log$_2$FC), MS1 (1.5log$_2$FC), ICL1 (1.78log$_2$FC) and ATO1 (1.58log$_2$FC). Additionally AST1 (1.73log$_2$FC) and GAP1a|GAP1b (1.87log$_2$FC) were significantly increased. Whereas the lysine acetylation sites of the upregulated proteins between heterotrophic and mixotrophic conditions were quite similar to the once in comparison between heterotrophic and photoautotrophic conditions, just three identical lysine acetylation sites were found for the downregulated ones which were CHLM (K271, -2.3log$_2$FC) and CDKK1 (K271, -1.73log$_2$FC) and PsaD (K114, -1.2log$_2$FC). Further chloroplast localized lysine acetylation sites were detected, the Light-harvesting protein of photosystem I (LHCA1K114, -1.2log$_2$FC), AtpE (AtpEK111, -1.0log$_2$FC), AtpB (AtpBK215, -1.95log$_2$FC), Transketolase (TRK1K72, -1.18log$_2$FC) and two sites of the Oxygen-evolving enhancer protein 1 of photosystem II (PSBOK121, -1.14log$_2$FC; PSBOK231, -1.2log$_2$FC). One lysine acetylation site of an mitochondrial localized adenosylhomocysteinase (SAH1K236, , -1.29log$_2$FC) was identified to be downregulated under heterotrophic conditions (Atteia et al., 2009). On protein level LHCA1 (-1.57log$_2$FC), CHLM (-1.21log$_2$FC) and PsaD (-1.167log$_2$FC) were significant downregulated as well (Fig. 4E).

The last two growth conditions which were compared were mixotrophic against photoautotrophic. Under mixotrophic conditions 29 lysine acetylation sites were upregulated compared to photoautotrophic conditions which belong to 23 unique proteins. Starting again with the proteins that were involved in peroxisomal metabolism eight out 23 proteins belong to this subcellular localization. As before the ACS3 (ACS3K205, 2.56 log_2FC and K218, 2.56log_2FC), CIS2 (CIS2K100, 2.1log_2FC; CISK447, 2.36log_2FC) and MS1 (MS1K422, 1.77log_2FC) were discovered and upregulated under mixotrophic conditions. Interestingly the ICL1 did not show any changes in lysine acetylation between these conditions (ICL1K195, -0.084log2FC). Pp1 showed again three lysine acetylation sites being upregulated (Pp1K305, 2.56log_2FC; Pp1K553, 1.73log_2FC; Pp1K565, 2.86log_2FC) as well as Pp3 showed one upregulated lysine acetylation site (Pp3K116, 2.47log_2FC). ATO1 carried two increased lysine acetylation sites (ATO1K232, 2.65log_2FC; ATO1K274, 1.97log_2FC) as well as the fatty acid regulated protein Pp2 (Pp2K142, 2.02 log_2FC). Three proteins with increased lysine acetylation site under mixotrophic conditions are localized to the chloroplast, the RbcL (RbcLK175, 1.02log_2FC), AtpB (AtpBK215, 2.91log_2FC), a membrane AAA-metalloprotease (FTSH1K, 1.39 log_2FC), DLA2 (DLA2K161, 1.91log_2FC) and a ribosomal protein S11 (Rps11K43, 1.15 log_2FC). Furthermore five proteins were found to be localized in the cytosol, a heat shock protein 70B (HSP70K278, 2.33log_2FC), RPL7 (RPL7K57, 1.24log_2FC), GAP1a|GAP1b (GAP1aK257|GAP1bK204, 1.12log_2FC; GAP1aK199|GAP1bK146, 1.07log_2FC), ADH1 (ADH1K206, 1.45log_2FC). Two mitochondrial localized proteins, the AST1 (AST1K66, 2.24log_2FC) and an Adenosylhomocysteinase (SAH1K236, 2.53 log_2FC) were upregulated under mixotrophic conditions compared to photoautotrophic. Likewise three proteins with unknown localization showed increased acetylation, 2* (2*K313, 1.09log_2FC), a Ring3 protein (4*K230, 1.76log_2FC) and a predicted protein, A8I4R4 (Pp8) with an N-acyltransferases activity (Pp8K143, 1.07log_2FC). By comparing mixotrophic to photoautotrophic on total proteome level just the ACS3 is upregulated significantly under heterotrophic conditions (1.47log_2FC). Significant downregulated lysine acetylation sites for mixotrophic conditions compared to photoautotrophic were found on three chloroplastic proteins, a CR051 protein (1*K280, -2.6log_2FC), a ribosomal protein S4 (Rsp4K252, -1.24log_2FC) and 3* (3*K49, -1.12log_2FC). One histone protein was discovered, Histone H3 (HTR14) which was significantly down regulated under mixotrophic conditions (HTR14K24, -1.32 log_2FC). All other proteins with decreased lysine acetylation sites were of unknown localization, an ATM/ATR-like kinase (6*K3333, -1.73log_2FC), A8JBL8 (Pp6K320, -

1.21log_2FC and A8HMW6 (Pp7K780, -1.19log_2FC). No significant changes on total proteome level were observed (Fig. 4F).

Activity changes of citrate synthase after deacetylase treatment

To validate our data for functional relevance, the peroxisomal CIS2 was picked in order to perform activity assays. CIS2 was discovered to be dynamically regulated within the different growth conditions on several lysine acetylation sites (Fig. 4A-C). So far nothing was known about the influence of lysine acetylation on CS activity. Until now, a dimer dependent redox regulation of the mitochondrial CS has been shown in Arabidopsis thaliana (Schmidtmann et al., 2014). lysine acetylation As the CIS2 represents a key enzyme of the GC and is directly involved in the conversion from acetyl-coA it is of great interest to disclose the regulatory mechanisms of this enzyme. In a first approach the total activity of CS was measured by using total protein extract from heterotrophic, mixotrophic and photoautotrophic grown cells. It cannot be ruled out that the activity is partially caused by the mitochondrial CS but as no changes on either total protein level nor on lysine acetylation could be observed it can be assumed that the changes in activity were mainly caused by the peroxisomal CIS2. The changes in activity were in concert with the changes on total protein level. Hence under heterotrophic conditions the CIS2 activity was 2.2 times higher compared to photoautotrophic grown cells (heterotrophic/photoautotrophic, 2.5log_2FC), 1.65 times higher by comparing Hetero with mixotrophic (heterotrophic/ mixotrophic o, 1.46log_2FC) and 1.33 times higher by comparing CS activity of mixotrophic versus photoautotrophic grown cells (mixotrophic/photoautotrophic, 0.9log_2FC). Thus the main activity changes were provoked by total protein changes. To analyze the effect of lysine acetylation on the CIS2 activity, the protein extracts were treated either with DI to avoid deacetylation or with the recombinant YHDZ from *E.coli*. The activity of the DI sample was set to 100% as it represents the present lysine acetylation state of the protein extract in comparison to the YHDZ treated samples which showed a decrease in lysine acetylation (supplementary). In general a decrease in CS activity was observed after deacetylation, especially the photoautotrophic grown cells exhibit a significant difference (Fig. 5B).

Discussion

In recent years, several global acetylome characterizations have been reported both in prokaryotes and eukaryotes. This study presents the first *C. reinhardtii* acetylome in combination with stable dimethyl labelling including a comparison between heterotrophic, photoautotrophic and mixotrophic growth conditions. It is known that mixotrophic conditions promote the optimal lipid formation which is important for large-scale generation of algal-based biofuels (Sager and Granick, 1953; Work et al., 2010). To further improve the biofuel production it is of great interest to understand the basic principles of metabolic pathway regulation in *C. reinhardtii*. The acetylome data which we present here provide new possible regulatory mechanisms based on PTM. The use of three biological replicates and accurate similar growth parameters for the different growth conditions gave the base for a significant and meaningful acetylome study. Additionally they provide the first total proteome analysis between heterotrophic, mixotrophic and photoautotrophic conditions. By using this straightforward method we identified an overall number of 3422 proteins with 311 lysine acetylation sites which belong to 165 unique protein groups (Tab.1). Our data show that diverse metabolic pathways as well as functional categories possess lysine acetylated proteins which is in concert with several other organisms (Cobbold et al., 2016; Fang et al., 2015; Finkemeier et al., 2011; Henriksen et al., 2012; Liu et al., 2014; Liu et al., 2016). This widespread distribution represent the vital biological processes of this PTM which include metabolic enzymes involved in protein biosynthesis, photosynthesis as well as several important carbon utilization pathways including glycolysis, TCA and GC pathway. Exceptional for our study is the comparison of differential lysine acetylation within three growth conditions, heterotrophic, mixotrophic and photoautotrophic growth by dimethyl labeling. These three conditions imply the parameter acetate, light and dark. Hence it was possible to investigate which parameter influences lysine acetylation the most as well as to uncover the possible regulatory function. Our data suggest that acetate has the main influence on lysine acetylation which can be explained by the enormous changes in lysine acetylation by comparing either heterotrophic versus photoautotrophic as well as mixotrophic versus photoautotrophic whereas heterotrophic versus mixotrophic just show minor changes in lysine acetylation (Fig. 4). As acetate can be converted to acetyl-coA which is the donor of the acetyl group in lysine acetylation it can be speculated that lysine acetylation plays a major role in directing metabolic enzymes via activity changes in the acetate converting direction.

The link between acetate metabolism and lysine acetylation is well studied and is mainly established because of the known activity regulation of ACS by lysine acetylation in several organisms (Crosby et al., 2010; Gardner et al., 2006; Starai et al., 2002). However as well in our data the ACS2, which provides the entrance of acetate into the GC cycle, was acetylated on several K residues and additionally upregulated in lysine acetylation in acetate containing growth media (Fig. 4). Additional to the ACS2 all enzymes of the GC beside the ACH1 were lysine acetylation (Fig. 6). The presence as well as the control of ICL activity by lysine acetylation was shown in *Escherichia coli* cells and the pathogen *Mycobacterium tuberculosis* (Castano-Cerezo et al., 2014; Xie et al., 2015). Nothing was known so far about the regulation of CS by lysine acetylation. Therefore a spectrometric CS assay was performed which gave insight into activity changes triggered by lysine acetylation. Total protein extract was treated with and without an external added deacetylase, the YHDZ. The most striking CS activity changes were based on total protein amount (Fig. 5A) but nevertheless also significant differences after the deacetylation reaction through the YHDZ could be observed (Fig. 5B). For future experiments it would be interesting to perform site directed mutagenesis of the different lysine acetylation sites to achieve a better insight into the regulatory function of CS by lysine acetylation. Beside the lysine acetylation of all GC enzymes as well a strong increase on total proteome level for all GC enzymes supports the hypothesis that acetate enters via the GC pathway in peroxisomes which is the preferred carbon utilization pathway especially for heterotrophic grown cells (Fig. 6). By further analyzing the changes on proteome level it became obvious that heterotrophic grown cells depend on fermentative pathways like the PRF1 which process H_2 as end product (Mus et al., 2007; Noth et al., 2013). A possible imaginable pathway is that OAA from the GC is exported out of the peroxisome into the cytosol where it is converted to Phosphoenolpyruvate (PEP) by the PCK and further metabolized to pyruvate (PYR) by the Pyruvate phosphate dikinase (PPD1). PYR can then enter the chloroplast and is used to produce acetyl-coA by the PFR1 and is finally used by the ADH1 to produce ethanol. The ADH1 is not just upregulated on total proteome but as well shows increased acetylation. This fermentative pathway is an important renewable energy supply as hydrogen is, beside the divergent possibilities of producing energy, one of the cleanest and therefore most attractive products as it has no carbon emission (Hemschemeier et al., 2009). Therefore it is of great interest that for the mitochondrial aldehyde dehydrogenase 2 (ALDH2) a functional relevance of lysine acetylation was already observed (Xue et al., 2012). Photoautotrophic grown cells produce their energy by the fixation of inorganic CO_2

fixation via the RUBISCO. The RBCL has three lysine acetylation sites and one which shows a down regulation under photoautotrophic conditions compared to heterotrophic and mixotrophic. This downregulation is probably in concert with an increase in activity as reported for *Arabidopsis thaliana* (Gao et al., 2016). In general, photoautotrophic grown cells mainly had PS related proteins which were differentially lysine acetylation (PsaB, PsaD, PdbR, AtpB, petA and CHLM) (Fig. 4; Fig. 6). Additionally several CO_2 accumulating enzymes are upregulated on protein level (Fig. 2B, D) which provide adequate CO_2 supply (Tirumani et al., 2014; Ynalvez et al., 2008). The same proteins were upregulated in lysine acetylation for mixotrophic conditions compared to heterotrophic which shows that mixotrophic grown cells not just receive energy by the conversion of acetate but that simultaneously the PS is active (Fig. 4B; Fig. 6). As *C. reinhardtii* is an important organism for alternative biofuels it is of great interest to further improve its growth potential. Therefore the metabolic key points have to be discovered and itemized. Genetic manipulations could now offer an efficient tool to modulate and regulate metabolic flow. One possibility to alter lysine acetylation is site directed mutagenesis by exchanging the responsible lysine residue by a glutamate or an arginine which either simulates the acetylated or the non-acetylated state respectively. Functional relevance of lysine acetylation of photosynthetic proteins beside the RBCL could not be shown so far. Improvement of PS is as well a common topic and therefore analysis of the different lysine acetylation sites within the PS apparatus would be of great relevance.

References

Agrimi, G., Russo, A., Scarcia, P., and Palmieri, F. (2012). The human gene SLC25A17 encodes a peroxisomal transporter of coenzyme A, FAD and NAD+. Biochem J *443*, 241-247.

Anoop, V.M., Basu, U., McCammon, M.T., McAlister-Henn, L., and Taylor, G.J. (2003). Modulation of citrate metabolism alters aluminum tolerance in yeast and transgenic canola overexpressing a mitochondrial citrate synthase. Plant Physiol *132*, 2205-2217.

Atteia, A., Adrait, A., Brugiere, S., Tardif, M., van Lis, R., Deusch, O., Dagan, T., Kuhn, L., Gontero, B., Martin, W., *et al.* (2009). A proteomic survey of Chlamydomonas reinhardtii mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the alpha-proteobacterial mitochondrial ancestor. Mol Biol Evol *26*, 1533-1548.

Boersema, P.J., Raijmakers, R., Lemeer, S., Mohammed, S., and Heck, A.J. (2009). Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. Nat Protoc *4*, 484-494.

Bohne, A.V., Schwarz, C., Schottkowski, M., Lidschreiber, M., Piotrowski, M., Zerges, W., and Nickelsen, J. (2013). Reciprocal regulation of protein synthesis and carbon metabolism for thylakoid membrane biogenesis. PLoS Biol *11*, e1001482.

Castano-Cerezo, S., Bernal, V., Post, H., Fuhrer, T., Cappadona, S., Sanchez-Diaz, N.C., Sauer, U., Heck, A.J., Altelaar, A.F., and Canovas, M. (2014). Protein acetylation affects acetate metabolism, motility and acid stress response in Escherichia coli. Mol Syst Biol *10*, 762.

Catalanotti, C., Yang, W., Posewitz, M.C., and Grossman, A.R. (2013). Fermentation metabolism and its evolution in algae. Front Plant Sci *4*, 150.

Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., and Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. Science *325*, 834-840.

Choudhary, C., and Mann, M. (2010). Decoding signalling networks by mass spectrometry-based proteomics. Nat Rev Mol Cell Biol *11*, 427-439.

Choudhary, C., Weinert, B.T., Nishida, Y., Verdin, E., and Mann, M. (2014). The growing landscape of lysine acetylation links metabolism and cell signalling. Nat Rev Mol Cell Biol *15*, 536-550.

Cobbold, S.A., Santos, J.M., Ochoa, A., Perlman, D.H., and Llinas, M. (2016). Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. Sci Rep *6*, 19722.

Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J., and Gevaert, K. (2009). Improved visualization of protein consensus sequences by iceLogo. Nat Methods *6*, 786-787.

Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. Molecular & cellular proteomics : MCP *13*, 2513-2526.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

Crosby, H.A., Heiniger, E.K., Harwood, C.S., and Escalante-Semerena, J.C. (2010). Reversible N epsilon-lysine acetylation regulates the activity of acyl-CoA synthetases involved in anaerobic benzoate catabolism in Rhodopseudomonas palustris. Mol Microbiol *76*, 874-888.

Erickson, E., Wakao, S., and Niyogi, K.K. (2015). Light stress and photoprotection in Chlamydomonas reinhardtii. Plant J *82*, 449-465.

Fang, X., Chen, W., Zhao, Y., Ruan, S., Zhang, H., Yan, C., Jin, L., Cao, L., Zhu, J., Ma, H., *et al.* (2015). Global analysis of lysine acetylation in strawberry leaves. Front Plant Sci *6*, 739.

Finkemeier, I., Laxa, M., Miguet, L., Howden, A.J., and Sweetlove, L.J. (2011). Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. Plant Physiol *155*, 1779-1790.

Galdieri, L., Zhang, T., Rogerson, D., Lleshi, R., and Vancura, A. (2014). Protein acetylation and acetyl coenzyme a metabolism in budding yeast. Eukaryot Cell *13*, 1472-1483.

Gao, X., Hong, H., Li, W.C., Yang, L., Huang, J., Xiao, Y.L., Chen, X.Y., and Chen, G.Y. (2016). Downregulation of Rubisco Activity by Non-enzymatic Acetylation of RbcL. Mol Plant.

Gardner, J.G., Grundy, F.J., Henkin, T.M., and Escalante-Semerena, J.C. (2006). Control of acetyl-coenzyme A synthetase (AcsA) activity by acetylation/deacetylation without NAD(+) involvement in Bacillus subtilis. J Bacteriol *188*, 5460-5468.

Glasser, C., Haberer, G., Finkemeier, I., Pfannschmidt, T., Kleine, T., Leister, D., Dietz, K.J., Hausler, R.E., Grimm, B., and Mayer, K.F. (2014). Meta-analysis of retrograde signaling in Arabidopsis thaliana reveals a core module of genes embedded in complex cellular signaling networks. Mol Plant *7*, 1167-1190.

Goodenough, U., Blaby, I., Casero, D., Gallaher, S.D., Goodson, C., Johnson, S., Lee, J.H., Merchant, S.S., Pellegrini, M., Roth, R., *et al.* (2014). The path to triacylglyceride obesity in the sta6 strain of Chlamydomonas reinhardtii. Eukaryot Cell *13*, 591-613.

Hartl, M., Konig, A.C., and Finkemeier, I. (2015). Identification of lysine-acetylated mitochondrial proteins and their acetylation sites. Methods Mol Biol *1305*, 107-121.

Hayashi, Y., Sato, N., Shinozaki, A., and Watanabe, M. (2015). Increase in peroxisome number and the gene expression of putative glyoxysomal enzymes in Chlamydomonas cells supplemented with acetate. J Plant Res *128*, 177-185.

Hemschemeier, A., Melis, A., and Happe, T. (2009). Analytical approaches to photobiological hydrogen production in unicellular green algae. Photosynth Res *102*, 523-540.

Henriksen, P., Wagner, S.A., Weinert, B.T., Sharma, S., Bacinskaja, G., Rehman, M., Juffer, A.H., Walther, T.C., Lisby, M., and Choudhary, C. (2012). Proteome-wide analysis of lysine acetylation suggests its broad regulatory scope in Saccharomyces cerevisiae. Molecular & cellular proteomics : MCP *11*, 1510-1522.

Hershko, A., and Ciechanover, A. (1998). The ubiquitin system. Annu Rev Biochem *67*, 425-479.

Hoober, J.K. (1989). The Chlamydomonas Sourcebook. A Comprehensive Guide to Biology and Laboratory Use. Elizabeth H. Harris. Academic Press, San Diego, CA, 1989. xiv, 780 pp., illus. $145. Science (New York, NY) *246*, 1503-1504.

Hosp, F., Lassowskat, I., Santoro, V., De Vleesschauwer, D., Fliegner, D., Redestig, H., Mann, M., Christian, S., Hannah, M.A., and Finkemeier, I. (2017). Lysine acetylation in mitochondria: From inventory to function. Mitochondrion *33*, 58-71.

Jensen, O.N. (2006). Interpreting the protein language using proteomics. Nat Rev Mol Cell Biol *7*, 391-403.

Jing, E., O'Neill, B.T., Rardin, M.J., Kleinridders, A., Ilkeyeva, O.R., Ussar, S., Bain, J.R., Lee, K.Y., Verdin, E.M., Newgard, C.B., *et al.* (2013). Sirt3 regulates metabolic flexibility of skeletal muscle through reversible enzymatic deacetylation. Diabetes *62*, 3404-3417.

Johnson, X., and Alric, J. (2013). Central carbon metabolism and electron transport in Chlamydomonas reinhardtii: metabolic constraints for carbon partitioning between oil and starch. Eukaryot Cell *12*, 776-793.

Jones, C.S., and Mayfield, S.P. (2012). Algae biofuels: versatility for the future of bioenergy. Curr Opin Biotechnol *23*, 346-351.

König, A.C., Hartl, M., Boersema, P.J., Mann, M., and Finkemeier, I. (2014). The mitochondrial lysine acetylome of Arabidopsis. Mitochondrion *19 Pt B*, 252-260.

Kunze, M., Pracharoenwattana, I., Smith, S.M., and Hartig, A. (2006). A central role for the peroxisomal membrane in glyoxylate cycle function. Biochim Biophys Acta *1763*, 1441-1452.

L'Hernault, S.W., and Rosenbaum, J.L. (1983). Chlamydomonas alpha-tubulin is posttranslationally modified in the flagella during flagellar assembly. J Cell Biol *97*, 258-263.

L'Hernault, S.W., and Rosenbaum, J.L. (1985). Chlamydomonas alpha-tubulin is posttranslationally modified by acetylation on the epsilon-amino group of a lysine. Biochemistry *24*, 473-478.

Lauersen, K.J., Willamme, R., Coosemans, N., Joris, M., Kruse, O., and Remacle, C. (2016). Peroxisomal microbodies are at the crossroads of acetate assimilation in the green microalga Chlamydomonas reinhardtii. Algal Res *16*, 266-274.

Ledford, H.K., Chin, B.L., and Niyogi, K.K. (2007). Acclimation to singlet oxygen stress in Chlamydomonas reinhardtii. Eukaryot Cell *6*, 919-930.

Leite, G.B., Abdelaziz, A.E., and Hallenbeck, P.C. (2013). Algal biofuels: challenges and opportunities. Bioresour Technol *145*, 134-141.

Liu, F., Yang, M., Wang, X., Yang, S., Gu, J., Zhou, J., Zhang, X.E., Deng, J., and Ge, F. (2014). Acetylome analysis reveals diverse functions of lysine acetylation in Mycobacterium tuberculosis. Molecular & cellular proteomics : MCP *13*, 3352-3366.

Liu, L., Wang, G., Song, L., Lv, B., and Liang, W. (2016). Acetylome analysis reveals the involvement of lysine acetylation in biosynthesis of antibiotics in Bacillus amyloliquefaciens. Sci Rep *6*, 20108.

Lundby, A., Lage, K., Weinert, B.T., Bekker-Jensen, D.B., Secher, A., Skovgaard, T., Kelstrup, C.D., Dmytriyev, A., Choudhary, C., Lundby, C., *et al.* (2012). Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. Cell Rep *2*, 419-431.

Macek, B., Mann, M., and Olsen, J.V. (2009). Global and site-specific quantitative phosphoproteomics: principles and applications. Annu Rev Pharmacol Toxicol *49*, 199-221.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol *6*, 838-849.

Melo-Braga, M.N., Verano-Braga, T., Leon, I.R., Antonacci, D., Nogueira, F.C., Thelen, J.J., Larsen, M.R., and Palmisano, G. (2012). Modulation of protein phosphorylation, N-glycosylation and Lys-acetylation in grape (Vitis vinifera) mesocarp and exocarp owing to Lobesia botrana infection. Molecular & cellular proteomics : MCP *11*, 945-956.

Merchant, S.S., Kropat, J., Liu, B.S., Shaw, J., and Warakanont, J. (2012). TAG, You're it! Chlamydomonas as a reference organism for understanding algal triacylglycerol accumulation. Curr Opin Biotech *23*, 352-363.

Mo, R., Yang, M., Chen, Z., Cheng, Z., Yi, X., Li, C., He, C., Xiong, Q., Chen, H., Wang, Q., *et al.* (2015). Acetylome analysis reveals the involvement of lysine acetylation in photosynthesis and carbon metabolism in the model cyanobacterium Synechocystis sp. PCC 6803. J Proteome Res *14*, 1275-1286.

Moulin, M., Nguyen, G.T., Scaife, M.A., Smith, A.G., and Fitzpatrick, T.B. (2013). Analysis of Chlamydomonas thiamin metabolism in vivo reveals riboswitch plasticity. Proc Natl Acad Sci USA *110*, 14622-14627.

Mus, F., Dubini, A., Seibert, M., Posewitz, M.C., and Grossman, A.R. (2007). Anaerobic acclimation in Chlamydomonas reinhardtii: anoxic gene expression, hydrogenase induction, and metabolic pathways. J Biol Chem *282*, 25475-25486.

Noth, J., Krawietz, D., Hemschemeier, A., and Happe, T. (2013). Pyruvate:ferredoxin oxidoreductase is coupled to light-independent hydrogen production in Chlamydomonas reinhardtii. J Biol Chem *288*, 4368-4377.

Rardin, M.J., Newman, J.C., Held, J.M., Cusack, M.P., Sorensen, D.J., Li, B., Schilling, B., Mooney, S.D., Kahn, C.R., Verdin, E., *et al.* (2013). Label-free quantitative proteomics of the lysine acetylome in mitochondria identifies substrates of SIRT3 in metabolic pathways. Proc Natl Acad Sci USA *110*, 6601-6606.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research *43*, e47.

Sager, R., and Granick, S. (1953). Nutritional studies with Chlamydomonas reinhardi. Ann N Y Acad Sci *56*, 831-838.

Schmidtmann, E., Konig, A.C., Orwat, A., Leister, D., Hartl, M., and Finkemeier, I. (2014). Redox Regulation of Arabidopsis Mitochondrial Citrate Synthase. Mol Plant *7*, 156-169.

Scranton, M.A., Ostrand, J.T., Fields, F.J., and Mayfield, S.P. (2015). Chlamydomonas as a model for biofuels and bio-products production. Plant J *82*, 523-531.

Seidel, J., Klockenbusch, C., and Schwarzer, D. (2016). Investigating Deformylase and Deacylase Activity of Mammalian and Bacterial Sirtuins. Chembiochem *17*, 398-402.

Smyth, G.K., Michaud, J., and Scott, H.S. (2005). Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics *21*, 2067-2075.

Starai, V.J., Celic, I., Cole, R.N., Boeke, J.D., and Escalante-Semerena, J.C. (2002). Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science *298*, 2390-2392.

Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. Microbiol Mol Biol Rev *64*, 435-459.

Subramanian, V., Dubini, A., Astling, D.P., Laurens, L.M.L., Old, W.M., Grossman, A.R., Posewitz, M.C., and Seibert, M. (2014). Profiling Chlamydomonas Metabolism under Dark, Anoxic H-2-Producing Conditions Using a Combined Proteomic, Transcriptomic, and Metabolomic Approach. J Proteome Res *13*, 5431-5451.

Terashima, M., Specht, M., Naumann, B., and Hippler, M. (2010). Characterizing the anaerobic response of Chlamydomonas reinhardtii by quantitative proteomics. Molecular & cellular proteomics : MCP *9*, 1514-1532.

Tirumani, S., Kokkanti, M., Chaudhari, V., Shukla, M., and Rao, B.J. (2014). Regulation of CCM genes in Chlamydomonas reinhardtii during conditions of light-dark cycles in synchronous cultures. Plant Mol Biol *85*, 277-286.

Velasquez-Orta, S.B., Curtis, T.P., and Logan, B.E. (2009). Energy from algae using microbial fuel cells. Biotechnol Bioeng *103*, 1068-1076.

Weinert, B.T., Scholz, C., Wagner, S.A., Iesmantavicius, V., Su, D., Daniel, J.A., and Choudhary, C. (2013). Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. Cell Rep *4*, 842-851.

Weinert, B.T., Wagner, S.A., Horn, H., Henriksen, P., Liu, W.R., Olsen, J.V., Jensen, L.J., and Choudhary, C. (2011). Proteome-wide mapping of the Drosophila acetylome demonstrates a high degree of conservation of lysine acetylation. Sci Signal *4*, ra48.

Winck, F.V., Paez Melo, D.O., and Gonzalez Barrios, A.F. (2013). Carbon acquisition and accumulation in microalgae Chlamydomonas: Insights from "omics" approaches. J Proteomics *94*, 207-218.

Wisniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nat Methods *6*, 359-362.

Work, V.H., Radakovits, R., Jinkerson, R.E., Meuser, J.E., Elliott, L.G., Vinyard, D.J., Laurens, L.M., Dismukes, G.C., and Posewitz, M.C. (2010). Increased lipid accumulation in the Chlamydomonas reinhardtii sta7-10 starchless isoamylase mutant and increased carbohydrate synthesis in complemented strains. Eukaryot Cell *9*, 1251-1261.

Xie, L., Wang, X., Zeng, J., Zhou, M., Duan, X., Li, Q., Zhang, Z., Luo, H., Pang, L., Li, W., *et al.* (2015). Proteome-wide lysine acetylation profiling of the human pathogen Mycobacterium tuberculosis. Int J Biochem Cell Biol *59*, 193-202.

Xue, L., Xu, F., Meng, L., Wei, S., Wang, J., Hao, P., Bian, Y., Zhang, Y., and Chen, Y. (2012). Acetylation-dependent regulation of mitochondrial ALDH2 activation by SIRT3 mediates acute ethanol-induced eNOS activation. FEBS Lett *586*, 137-142.

Yan, W., and Chen, S.S. (2005). Mass spectrometry-based quantitative proteomic profiling. Brief Funct Genomic Proteomic *4*, 27-38.

Ynalvez, R.A., Xiao, Y., Ward, A.S., Cunnusamy, K., and Moroney, J.V. (2008). Identification and characterization of two closely related beta-carbonic anhydrases from Chlamydomonas reinhardtii. Physiol Plant *133*, 15-26.

Zhang, K., Chen, Y., Zhang, Z., and Zhao, Y. (2009). Identification and verification of lysine propionylation and butyrylation in yeast core histones using PTMap software. J Proteome Res *8*, 900-906.

FIGURE 1

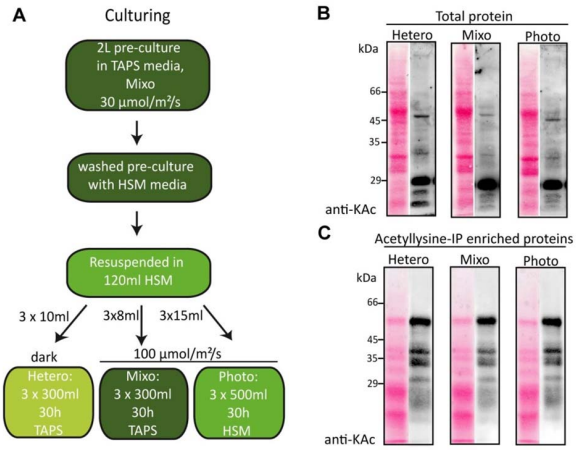


Figure 1. Overview of Chlamydomonas culture conditions and Western blot analysis of lysine-acetylated proteins. (**A**) Cells were grown in 2 L TAPS media to a density of ~ 5 x 10^6 cells/mL, washed in HSM to remove external acetate and resuspended in 120 mL HSM. To compensate for different growth rates under the applied conditions and to reach approximately the same cell densities after 30 h of growth, three cultures for each condition (heterotrophic, Hetero; Mixotrophic, Mixo; photoautotrophic, Photo) were inoculated with the indicated volumes of the resuspended pre-culture. (**B**) Western-blot analysis of total proteins and (**C**) immunoprecipitated lysine-acetylated proteins (acetyllysine IP) from cells grown under heterotrophic (Hetero), mixotrophic (Mixo), and phototrophic (Photo) conditions using the anti-acetyllysine antibody (anti-KAc). Left columns next to the blots show the Ponceau S staining as loading control.

FIGURE 2

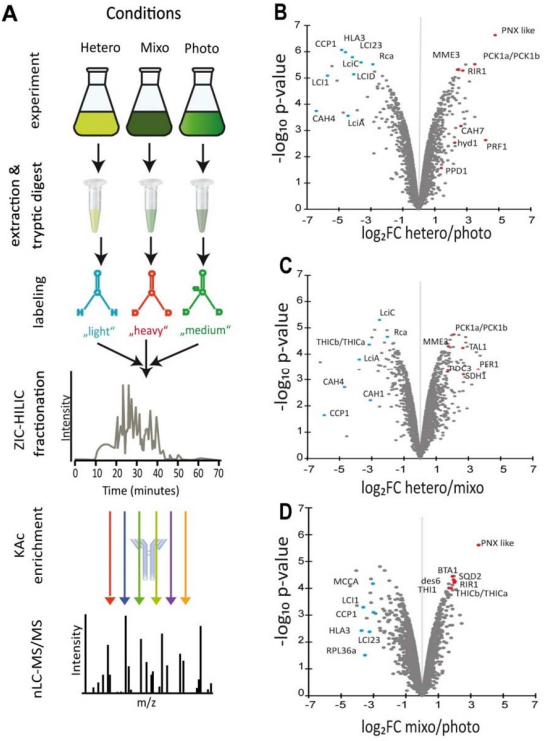


Figure 2. Experimental set-up for proteome-wide identification of lysine-acetylated peptides in Chlamydomonas. (**A**). Proteins were extracted and subjected to a tryptic digestion. Peptides from different growth conditions were pooled after triplex dimethyl-labeling and fractionated by ZIC-HILIC chromatography. The fractions were enriched for lysine-acetylated peptides using an anti-acetyllysine agarose, and eluted and desalted peptides were analysed by mass spectrometry. (**B-D**) Scatter plots of proteome changes observed between the different growth conditions (B. Hetero versus Photo; C. Hetero versus Mixo; D. Mixo versus Photo). The LIMMA log10 p-values were plotted against $\log_2$ LFQ intensities (threshold for significance: $-\log_{10}$ p-value ≥ 1.3; $\log_2$FC ≥ 1).

FIGURE 3

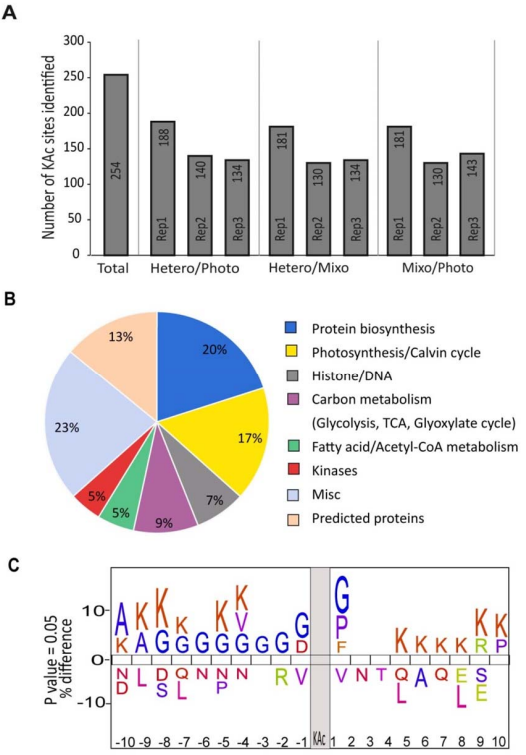


Figure 3. **Overview of lysine-acetylated proteins identified in Chlamydomonas.** (**A**) The graph shows the total number of lysine acetylation sites identified in all three experiments as well as the number of KAc sites identified in each of the three experiments. Rep, replicate. (**B**) Pie chart of functional categories of all identified lysine-acetylated proteins. (**C**) Sequence probability logos of significantly enriched acetylation site motifs for ±10 amino acids around the lysine acetylation sites. The sequence logo was generated using the IceLogo tool (http://http://iomics.ugent.be/icelogoserver/).

FIGURE 4

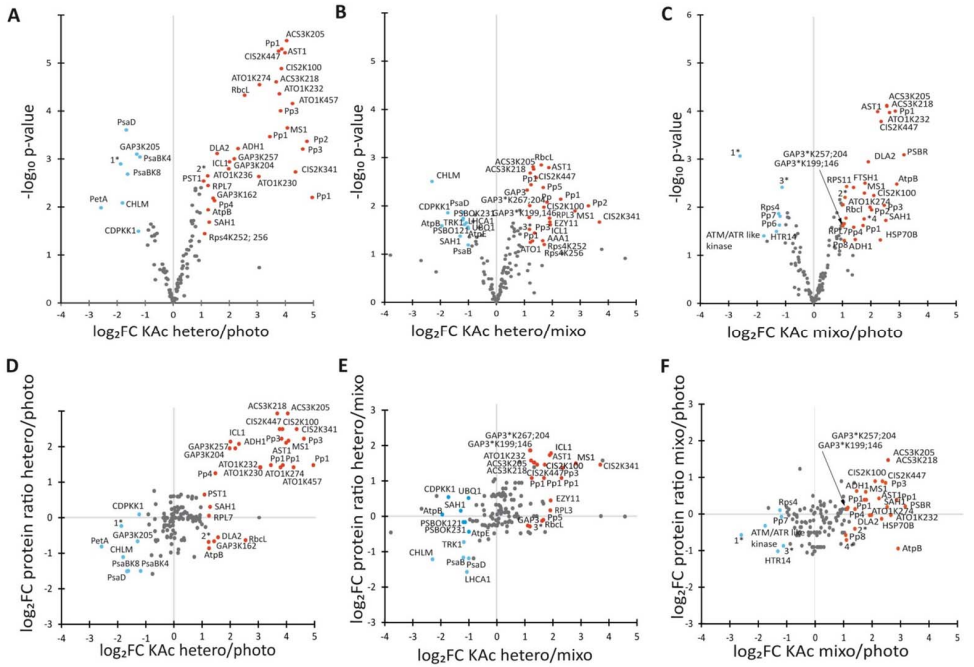


FIG. 4. **Differential KAc between the different growth conditions.** A-C. Scatter plot of KAc differences between different growth conditions. $-\log_{10}$ p-value was plotted against the KAc $\log_2$FC between the different conditions. D-E. Scatter plot KAc differences including total proteome changes of proteins. The $\log_2$FC of proteome changes is plotted against the $\log_2$FC of KAc. (A, D. Hetero versus Photo; B, E. Hetero versus Mixo; C, F. Mixo versus Photo). Red dots symbolize significant KAc upregulation and blue dots represent significant KAc downregulation (threshold for significance: $-\log_{10}$ p-value ≥ 1.3; $\log_2$FC ≥ 1).

FIGURE 5

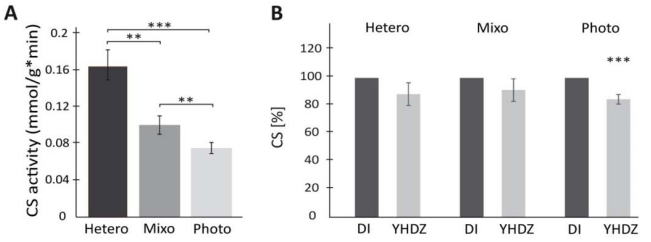


FIG. 5. **Citrate synthase activity influenced by KAc.** A**.** Citrate synthase activity of total protein extract of Hetero, Mixo and Photo conditions. B. Enzyme activity of citrate synthase of each growth condition after treatment with or without deacetylase. Extract which was treated with DI was set to 100% activity. Asterisks indicate significant differences (** $P < 0.01$; *** $P < 0.001$, *t* test). YHDZ, NAD-dependent protein deacetylase; DI, deacetylase inhibitor.

FIGURE 6

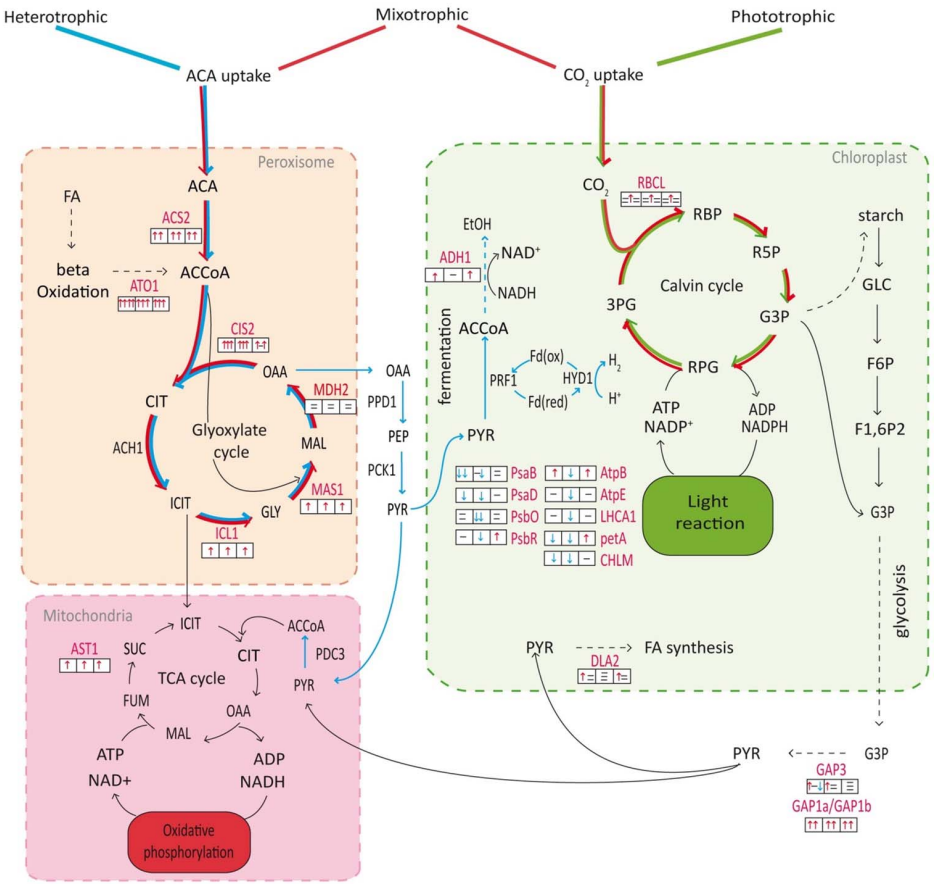
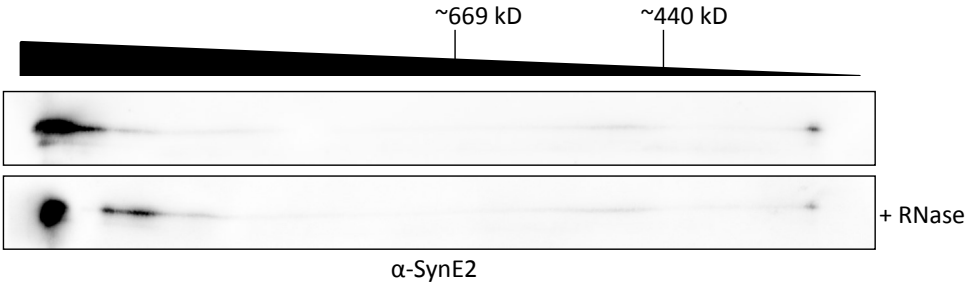


FIG. 6**. KAc within different metabolic pathways in Chlamydomonas *reinhardtii*.**
Central carbon metabolism of Chlamydomonas *reinhardtii*. Blue lines represent preferred carbon utilization under Hetero conditions, red lines under Mixo conditions and green lines under Photo conditions. Red enzymes represents KAc enzymes which were significantly regulated within the different growth conditions. KAc up or down regulation is depicted in the three boxes next to the KAc enzymes (left box; Hetero versus Photo, middle box; Hetero versus Mixo, right box; Mixo versus Photo, red arrow; KAc upregulated, blue arrow; KAc down regulated, minus; not significantly regulated between corresponding conditions). Acetate assimilation mainly takes place in the glyoxysomes and can feed either into the mitochondrial TCA cycle or into fermentative pathways localized in the chloroplast. Carbon dioxide fixation takes place in the chloroplast and energy is achieved by photosynthesis .

6.2 Supplemental figure



Supplemental Figure 1: RNase-sensitive complex in Synechocystis. Thylakoids from the wild-type strain *Synechocystis* sp. PCC 6803 were solubilized and either treated with RNase or incubated on ice for the same time without the addition of RNase before 2D-BN-PAGE was performed. Second dimensions were subjected to immunoblots.

List of publications

Kleinknecht, L., Nickelsen, J., and Bohne, A.-V. (2017). Unanticipated T7 RNA polymerase activity using annealed oligonucleotides as transcription template. Endocytobiosis and Cell Research 28, 33-37.

Torabi, S., Umate, P., Manavski, N., Plöchinger, M., **Kleinknecht, L.**, Bogireddi, H., Herrmann, R.G., Wanner, G., Schroder, W.P., and Meurer, J. (2014). PsbN is required for assembly of the photosystem II reaction center in *Nicotiana tabacum*. Plant Cell *26*, 1183-1199.

Kleinknecht, L., Wang, F., Stübe, R., Philippar, K., Nickelsen, J., and Bohne, A.-V. (2014). RAP, the sole octotricopeptide repeat protein in Arabidopsis, is required for chloroplast *16S* rRNA maturation. Plant Cell *26*, 777-787

Acknowledgement

First of all, I would like to thank Professor Nickelsen for giving me the opportunity to work on a number of interesting topics under his supervision. I started my work in his group already with a fascination for the chloroplast that even expanded while understanding more and more of the sophisticated processes inside it. Prof. Nickelsen was always there when problems arose and gave me helpful and instructive advice to carry on my research.

Special thanks are given to Prof. Dr. Peter Geigenberger for taking over the second position in my exam commission.

I want to thank the *Studienstiftung des Deutschen Volkes* for the financial support in form of a PhD scholarship during the first three years of my work in the lab.

My special thanks go to Daniel and Anne, who worked on the DLA2 project with me at different times. Working together and discussing our work was very pleasant and certainly helped to develop new ideas. And I shouldn't forget to mention Daniel's excellent music selection for our lab, which certainly helped to create a good mood in the Chlamy lab even in times that were not easy. Thank you for making my time in the lab a lot more entertaining and making me laugh quite often!

I'm also very thankful to the "old" Syni lab crew, Anna and Steffen, for discussing experiments with me and eating salad together a lot of times! I want to thank Steffen especially for correction of my thesis and finding all the weird comma errors and the nonsensical phrases. Thank you also for answering all my questions, when I started my Synechocystis project, you helped me to overcome many obstacles!

I also want to sincerely thank Alex for her constructive suggestions, open-minded discussions and for showing me how to perform certain experiments like sequencing gels and dot-blots.

Many thanks go to Karin for doing all the organizational stuff in the lab, like ordering everything we wished for (sometimes urgently) and for keeping a back-up of all the strains I worked with.

Moreover, I am very thankful to all other members of the lab Montse, Tsong, and Julia and of course former group members as well, for answering many questions, helping me out and for having quite a few nice parties together.

Last but not least, I would like to thank most of all my family and friends who have been always there for me and supporting me whenever I needed them. Without you it would not have been possible to finish this thesis! Special thanks go to Jaya for not giving up on asking when I will finally hand in my thesis ;)

Eidesstattliche Erklärung und Erklärung

Eidesstattliche Erklärung

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

München, den 15. Februar 2018

Laura Kleinknecht

Erklärung

Hiermit erkläre ich, dass die Dissertation nicht ganz oder in wesentlichen Teilen einer anderen Prüfungskommission vorgelegt worden ist. Des Weiteren habe ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen.

München, den 15. Februar 2018

Laura Kleinknecht