



Rubisco biogenesis and assembly in *Chlamydomonas reinhardtii*

Wojciech Wietrzynski

► To cite this version:

Wojciech Wietrzynski. Rubisco biogenesis and assembly in *Chlamydomonas reinhardtii*. Molecular biology. Université Pierre et Marie Curie - Paris VI, 2017. English. NNT : 2017PA066336 . tel-01770412

HAL Id: tel-01770412

<https://theses.hal.science/tel-01770412>

Submitted on 19 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



PhD Thesis of the Pierre and Marie Curie University (UPMC)

Prepared in the Laboratory of Molecular and Membrane Physiology of the Chloroplast, UMR7141, CNRS/UPMC

Doctoral school: Life Science Complexity, ED515

Presented by **Wojciech Wietrzynski**
for the grade of Doctor of the Pierre and Marie Curie University

Rubisco biogenesis and assembly in *Chlamydomonas reinhardtii*

Defended on the 17th of October 2017

at the Institut of Physico-Chemical Biology in Paris, France

Phd Jury:

Angela Falciatore, CNRS/UPMC, president

Michel Goldschmidt-Clermont, University of Geneva, reviewer

Michael Schroda, University of Kaiserslautern, reviewer

Cecile Raynaud, CNRS/IPS2, examiner

Steven Ball, University of Lille, examiner

Francis-André Wollman, CNRS/UPMC, supervisor

Katia Wostrikoff, CNRS/UPMC, supervisor

Acknowledgements

You all already know that I'm not very good at it but I'll try...

Thanks to Mama and Papa for always letting me choose whatever I want to do, going to France included;

To Francis-André for letting me stay in France in his laboratory, but foremost for all advices and critical comments and for making the lab above all, a place of discussion (not only scientific);

The biggest thanks to Katia for having me work with her for those 4 years! For defending me, and in general being patient despite me being difficult at times (most of the time);

Thanks to all that shared the bureau with me: Loreto, Han-yi, Domitille, Benjamin, Sandrine and the Ficus. Especially, to Sheriff Sandrine for keeping me in line and for all the food she shared with me!

I'd like to thank all the members: past and present of Francis-André's lab, everyone was special and memorable. I would be too long to evoke all the good memories I have with you. Just a special mention to The Axis: Stefania, Marina and Felix for trying to win this time. And Stephan, especially at the beginning, for always having time for a beer and talk;

To follow, thanks to all the colleagues I have found in the Institute: Marcello, Max, Justin and all the others;

Thanks to my friends Larissa and Wojtek for all the support. Wojtek in particular, for being my adversary for such a long time!

And last but not least, for my special little cat, Nathalie: for teaching me French, how to take care of cats, how to be more social and for being there for me!

It was a pleasure.

Table of contents

| | |
|--|----|
| Abbreviations..... | 8 |
| Opening words..... | 9 |
| 1. General Introduction..... | 11 |
| 1.1. Part I - Photosynthesis, endosymbiosis and consequences..... | 11 |
| 1.1.1. <i>Chloroplasts-a specialized organelle for oxygenic photosynthesis</i> | 11 |
| 1.1.2. <i>Chloroplast genome organization</i> | 13 |
| 1.1.3. <i>Endosymbiotic gene transfer: facts and hypotheses</i> | 15 |
| 1.1.4. <i>Protein rerouting to the chloroplast</i> | 17 |
| 1.2. Part II Nucleus and chloroplast crosstalk..... | 18 |
| 1.2.1. <i>Chloroplast transcription and translation apparatus and regulation</i> | 19 |
| 1.2.1.1. Transcription in chloroplast | 19 |
| 1.2.1.2. mRNA degradation in chloroplast | 19 |
| 1.2.1.3. Translation machinery | 20 |
| 1.2.1.4. Organization of translation | 21 |
| 1.2.2. <i>Nucleus-encoded regulators of chloroplast gene expression</i> | 22 |
| 1.2.2.1. TPR | 22 |
| 1.2.2.2. PPR | 23 |
| 1.2.2.3. OPR | 24 |
| 1.2.2.4. mTERF | 25 |
| 1.2.2.5. Convergent evolution- convergent role? | 26 |
| 1.2.2.6. OTAFs in chloroplast | 27 |
| 1.2.3. <i>Retrograde signaling</i> | 28 |
| 1.3. Part III: Regulatory processes involved in photosynthetic complex assembly..... | 31 |
| 1.3.1. <i>Concerted accumulation of subunits</i> | 31 |
| 1.3.2. <i>Proteolysis</i> | 31 |
| 1.3.3. <i>The CES process</i> | 32 |
| 1.3.3.1. Mechanism | 33 |
| 1.3.3.2. Special case of ATPsynthase | 35 |
| 1.3.3.3. CES as a regulatory process | 36 |
| 1.3.3.4. CES in higher plants | 36 |
| 1.3.3.5. CES and anterograde regulation | 37 |
| 1.3.3.6. Additional remarks | 38 |
| 2. Rubisco..... | 39 |
| 2.1. Rubisco evolution and clades..... | 39 |

| | |
|--|----|
| 2.2. Subunits and structure (Rubisco Type I) | 40 |
| 2.3. Rubisco reaction | 43 |
| 2.3.1. Oxygenation | 44 |
| 2.3.2. Unproductivity | 45 |
| 2.3.3. Efficiency | 45 |
| 2.4. Rubisco biogenesis | 46 |
| 2.4.1. Expression and regulation | 46 |
| 2.4.2. Folding and assembly | 51 |
| 2.4.2.1. CPN60 complex | 52 |
| 2.4.2.2. BSD2 | 53 |
| 2.4.2.3. Rubisco assembly chaperones | 53 |
| 2.4.2.4. RBCX | 54 |
| 2.4.2.5. RAF1 | 56 |
| 2.4.2.6. Comparison | 58 |
| 2.4.2.7. RAF2 | 58 |
| 2.4.2.8. SSU folding | 60 |
| 3. Interlude | 60 |
| 4. Chapter I : MRL1 role in <i>rbcL</i> translation regulation | 62 |
| 4.1. Introduction | 62 |
| 4.2. Part I - Searching for new OTAFs involved in <i>rbcL</i> gene expression | 63 |
| 4.2.1. Genetic approach: Mutagenesis | 63 |
| 4.2.1.1. Negative selection screen | 63 |
| 4.2.1.2. <i>rbcL</i> regulatory sequence drives cytosine deaminase expression | 65 |
| 4.2.1.3. Mutagenesis of CRE1 | 66 |
| 4.2.2. Biochemical approach: Co-immunoprecipitation of MRL-HA | 68 |
| 4.2.2.1. Generation of a tagged MRL1 strain | 68 |
| 4.2.2.2. ImmunoPrecipitation and Mass-spectrometry | 70 |
| 4.3. Part II - Is MRL1 alone? | 72 |
| 4.3.1. MRL1 level limits <i>rbcL</i> mRNA accumulation and is required for LSU | 72 |
| 4.3.2. MRL1 is a stable protein | 73 |
| 4.3.3. MRL1 is required for <i>rbcL</i> translation | 74 |
| 4.4. Discussion and perspectives | 77 |
| 4.4.1. In a search of <i>rbcL</i> T-factor | 77 |
| 4.4.2. M-factor's dual function? | 79 |
| 4.4.3. Insight in MRL1 mode of action? | 79 |
| 4.4.4. MRL1, a regulator of Rubisco accumulation? | 80 |

| | |
|--|-----|
| 5. Chapter II - Rubisco LSU synthesis depends on its oligomerization state in <i>Chlamydomonas reinhardtii</i> | 82 |
| 5.1. Abstract | 83 |
| 5.2. Introduction | 84 |
| 5.3. Results | 86 |
| 5.3.1. <i>rbcL</i> downregulation of synthesis in <i>Chlamydomonas</i> RBCS mutant | 86 |
| 5.3.2. LSU initiation of translation is impaired in absence of SSU | 86 |
| 5.3.3. Translation initiation is inhibited by unassembled LSU | 87 |
| 5.3.4. Assembly intermediates accumulate in absence of SSU | 88 |
| 5.3.5. CES regulation no longer occurs in Rubisco oligomerization mutants | 89 |
| 5.3.6. Rubisco assembly intermediates in the LSU ₈ mut oligomerization mutant reveal the LSU CES repressor | 91 |
| 5.4. Discussion | 92 |
| 5.4.1. LSU CES results from a negative autoregulation on translation initiation | 92 |
| 5.4.2. Fate of unassembled LSU | 94 |
| 5.4.3. Insights into Rubisco assembly pathway | 94 |
| 5.4.4. Tentative identification of the repressor form | 96 |
| 5.4.5. Evolutionary conservation of Rubisco CES process | 97 |
| 5.5. Materials and Methods | 98 |
| 5.6. Figure legends | 101 |
| 5.6.1. Fig. 1 LSU synthesis rate and accumulation in absence of its assembly partner | 101 |
| 5.6.2. Fig. 2 Swapping of <i>rbcL</i> regulatory sequences impairs the CES regulation | 101 |
| 5.6.3. Fig. 3 Expression of <i>cyt. f</i> is inhibited in the absence of Rubisco small subunit | 101 |
| 5.6.4. Fig. 4 CES regulation no longer occurs in the absence of LSU accumulation | 102 |
| 5.6.5. Fig. 5 LSU assembly intermediates accumulate in the SSU-lacking strain | 102 |
| 5.6.6. Fig. 6 LSU ₂ mutations alter LSU accumulation and CES regulation | 102 |
| 5.6.7. Fig. 7 Disruption of LSU oligomerization alters LSU CES regulation | 103 |
| 5.6.8. Sup. Fig. 1 Anti-RAF1 antibody | 103 |
| 5.6.9. Sup. Fig. 2 RAF1 accumulates in the absence of LSU | 103 |
| 5.6.10. Sup. Fig. 3 Close-up of the mutated residues in LSU structure | 103 |
| 5.7. Bibliography | 104 |
| 5.8. Figure list | 107 |
| 6. Chapter III - Further exploration of limitations in Rubisco biosynthesis – Supplementary results and discussion | 113 |
| 6.1. Additional results for: “Rubisco LSU synthesis depends on its oligomerization state in <i>Chlamydomonas reinhardtii</i> ” | 114 |

| | |
|---|-----|
| 6.1.1. <i>Generation of $\Delta RBCS$;5'UTRpsaA:rbcL strain and its additional phenotype</i> | 114 |
| 6.1.2. <i>Reporter gene accumulation is altered in $\Delta RBCS$;5'UTRpsaA:rbcL</i> | 115 |
| 6.1.3. <i>What happens with LSU when Rubisco assembly is perturbed?</i> | 117 |
| 6.1.4. <i>Stability of LSU</i> | 117 |
| 6.1.5. <i>Fractionation</i> | 118 |
| 6.1.6. <i>EPYC1 accumulation is decreased in Rubisco deficient strains</i> | 121 |
| 6.1.7. <i>Quantification of LSU and RAF1 levels</i> | 121 |
| 6.2. Discussion | 124 |
| 6.2.1. <i>LSU is relatively stable when chaperone-bound</i> | 124 |
| 6.2.2. <i>Limiting steps of LSU assembly</i> | 124 |
| 6.2.3. <i>Mutated LSU is directed to aggregates</i> | 125 |
| 6.2.4. <i>EPYC1 accumulates in coordination with Rubisco</i> | 126 |
| 6.2.5. <i>Conclusions from LSU and RAF1 quantifications</i> | 127 |
| 6.3. Additional comments | 127 |
| 7. General discussion | 129 |
| 8. Conclusions | 134 |
| 9. Materials | 135 |
| 10. Methods | 137 |
| 10.1. Cultures | 137 |
| 10.2. Genetics methods | 137 |
| 10.3. Biophysics methods | 138 |
| 10.4. Molecular biology methods | 139 |
| 10.5. Biochemical analysis | 140 |
| 11. Bibliography | 145 |

Abbreviations

5FC – 5-fluorocytosine
5FU – 5-fluorouracil
aa – Amino-acid(s)
AA – acrylamide
ACA – aminocaproic acid
ATP - Adenosine triphosphate
BDS2 - Bundle sheath defective 2
CD – Cytosine deaminase
CDS – Coding sequence
CES – Control by epistasy of synthesis
Co-IP – Co-immunoprecipitation
FUD - Fluorouridine
GOI – Gene of interest
HMWC – High molecular weight complex
LMWC – Low molecular weight complex
LSU – Rubisco large subunit
MIN – Tris-phosphate minimum medium
mTERF – Mitochondrial transcription terminator factor
NADP⁺/H⁺ - Nicotinamide adenine dinucleotide phosphate
OPR - Octotricopeptide repeat protein
ORF – Open reading frame
OTAF – Organellar trans-acting factors
PBS – Phosphate-buffered saline
PCR – Polymerase chain reaction
PPR – Pentatricopeptide repeat protein
PS – Photosynthesis/Photosynthetic
PVDF – Polyvinyl fluoride
RAF1 – Rubisco Accumulation factor 1
RLP – Rubisco-like protein
RuBP - d-ribulose-1,5-bisphosphate
SSU – Rubisco small subunit
TAP – Tris-Acetate-Phosphate medium
TPR - Tetratricopeptide repeat protein

Opening words

D-ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is the “first” enzyme of the Calvin Benson-Basham cycle that allows the entry of the atmospheric carbon (in form of CO₂) into biological cycle. It is the only significant way of biomass production and at the end, sustaining of the Earth’s food-chain as we know it. That makes it, probably the most important existing enzyme (a ranking, which could be contested only by another photosynthetic complex – Photosystem II). In addition, it is arguably the World’s most abundant protein as it is present in all photosynthetic organisms: cyanobacteria, algae and plants (also other, non-photosynthetic organisms such as bacteria and archaea) and can account for more than 50% of the leaves soluble proteins¹⁴. On the other side, as acclaimed as it is, Rubisco is also a very inefficient enzyme. Each molecule of it fixes only about forty CO₂ particles every second. Moreover it is prone to energetically uneconomic miss-reactions with oxygen as a substrate and to frequent “blocking” of its active sites by, for example its misfired reactions’ products. Considering its importance for the global agriculture and food production, it is not surprising that since many years it has been the object of vivid scientific interest. During my thesis more than 400 papers mentioning Rubisco in their titles were published. The majority of them try to highlight Rubisco evolution, its diversity; but also the mechanisms of its formation, regulation and its response to changing environmental conditions; and finally, how to improve its efficiency, carboxylation kinetics and CO₂ specificity. I hope that this little piece will also contribute to this global effort.

Table 1.1: Composition of photosynthetic complexes of *Chlamydomonas reinhardtii*

| Complex | Gene | Subunit | Origin |
|-----------------------|---------------|-----------------------|-------------|
| PSII | <i>psbA</i> | D1 | Chloroplast |
| | <i>psbB</i> | CP47 | |
| | <i>psbC</i> | CP43 | |
| | <i>psbD</i> | D2 | |
| | <i>psbE-F</i> | Cyt. b ₅₅₉ | |
| | <i>psbI</i> | PsbI | |
| | <i>psbJ</i> | PsbI | |
| | <i>psbH</i> | PsbH | |
| | <i>psbL</i> | PsbL | |
| | <i>psbM</i> | PsbM | |
| | <i>psbN</i> | PsbN | |
| | <i>psbT</i> | PsbT | |
| | <i>PSBO</i> | OEE3 | Nuclear |
| | <i>PSBP</i> | OEE2 | |
| | <i>PSBQ</i> | OEE1 | |
| | <i>PSBR</i> | PSBR | |
| | <i>PSBX</i> | PSBX | |
| | <i>PSBW</i> | PSBW | |
| Cyt. b ₆ f | <i>petA</i> | Cyt. <i>f</i> | Chloroplast |
| | <i>petB</i> | Cyt. b ₆ | Nuclear |
| | <i>petD</i> | Subunit IV | |
| | <i>PETC</i> | Rieske | Chloroplast |
| | <i>petG</i> | PetG | Nuclear |
| | <i>petL</i> | PetL | |
| | <i>PETM</i> | PetM | Nuclear |
| | <i>PETN</i> | PetN | |
| PSI | <i>psaA</i> | PsaA | Chloroplast |
| | <i>psaB</i> | PsaB | |
| | <i>psaC</i> | PsaC | |
| | <i>psaI</i> | PsaI | |
| | <i>psaJ</i> | PsaJ | |
| | <i>psaM</i> | PsaM | |
| | <i>PSAD</i> | PSAD | Nuclear |
| | <i>PSAE</i> | PSAE | |
| | <i>PSAF</i> | PSAF | |
| | <i>PSAG</i> | PSAG | |
| | <i>PSAH</i> | PSAH | |
| | <i>PSAK</i> | PSAK | |
| | <i>PSAL</i> | PSAL | |
| ATPase | <i>atpA</i> | CF ₁ α | Chloroplast |
| | <i>atpB</i> | CF ₁ β | |
| | <i>atpE</i> | CF ₁ ε | |
| | <i>atpF</i> | CF ₀ I | |
| | <i>atpH</i> | CF ₀ III | |
| | <i>atpI</i> | CF ₀ IV | |
| | <i>ATPC</i> | CF ₁ γ | Nuclear |
| | <i>ATPD</i> | CF ₁ σ | |
| | <i>ATPG</i> | CF ₀ II | |
| Rubisco | <i>rbcL</i> | LSU | Chloroplast |
| | <i>RBCS</i> | SSU | Nuclear |

1. General introduction

1.1. Part I - Photosynthesis, endosymbiosis and consequences

1.1.1. Chloroplasts-a specialized organelle for oxygenic photosynthesis

Life of the majority of Earth's organisms ultimately depends in way or the other on photosynthesis – a process that allows the conversion of light energy (photons' kinetic energy) into electrochemical energy of electrons which ends up in the storage of the biologically useful reductant equivalent NADPH. Because of the proton transfer coupled to electron movement through the membrane, an electrochemical gradient is created and can be used (via action of ATP synthase) to produce ATP – cell's energy molecule. Those molecules (NADPH and ATP) can be subsequently used in Calvin-Benson-Basham cycle to reduce and fix atmospheric carbon dioxide that at the end gives rise to the sugar molecules that are used in the cell's metabolism (Fig. 1.1) (reviewed in¹³). Photosynthesis in general is then a process in which a short-lived energy (photons) is captured and transferred in a step-wise manner to create a more accessible and universal long-lived energy carrier (sugar). The first photosynthesis-like process involving one photosystem complex (ancestor of PSI) appeared in purple, extremophile bacteria living close to oceanic hot springs and chimneys. In an anoxygenic reaction they used hydrogen sulfide as a donor of electrons for NADPH production. Today's photosynthesis as we know it, appeared in cyanobacteria around 3.5 billion years ago¹⁵. It is mainly an oxygenic process using water as an electron donor, and evolving oxygen as its byproduct. Sometime later, around 1-1.5 billion years ago, a cyanobacteria was engulfed and retained by heterotrophic, eukaryotic cell in a primary endosymbiosis event. It is mostly accepted that this rare event gave birth to all primary plastids of Archaeplastida. It has been also recently suggested that a third party – pathogenic Chlamydiae could have contributed to successful retention of the symbiont¹⁶. On the other hand, it has been proposed that other events of primary endosymbiosis might have occurred, as exemplified by chromophores found in the amoeba *Paulinella chromatophora*^{17,18}. With the rising of high-throughput sequencing technologies and new genomes being sequenced, discussion is still ongoing about exact origin of the closest plastid ancestor¹⁹⁻²¹. Furthermore, secondary and tertiary endosymbiosis took place by acquisition of primary plastids-containing eukaryotes by other, heterotrophic eukaryotes spreading plastids between lineages, the history of which remains an open and dynamic debate (reviewed in²²).

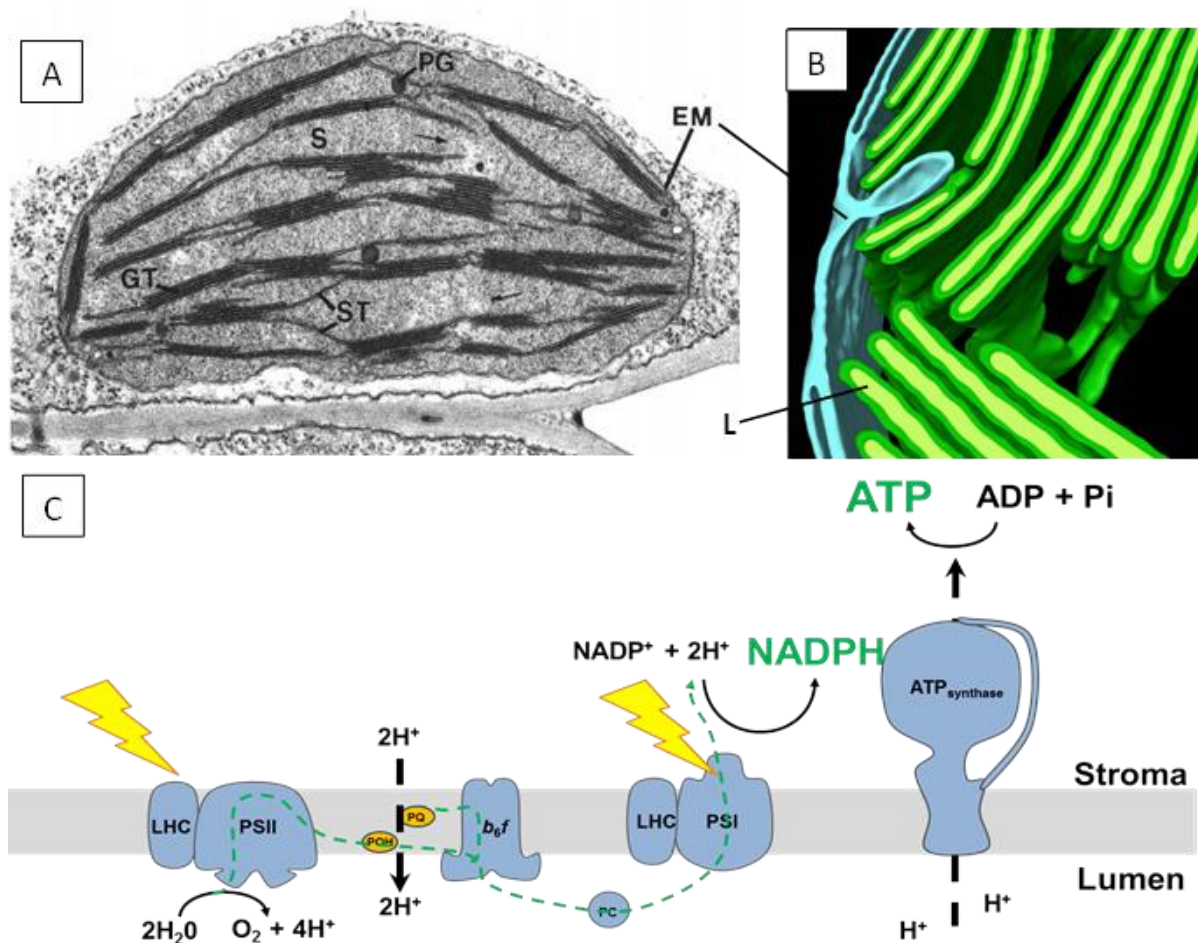


Fig. 1.1: Structure of the chloroplast

A) Electron micrograph of a young tobacco chloroplast showing structuration of thylakoids. S – Stroma; GT – grana thylakoids; ST – Stroma thylakoids = lamellae; EM – 2-layer chloroplast envelope membrane; PG – plastoglobuli. Adapted from: Staehelin, 2003⁴.

B) Tomography reconstruction *Chlamydomonas* chloroplast fragment showing thylakoids as elongated vesicles. L – Lumen. Adapted from: Engel et al. (2015)⁷.

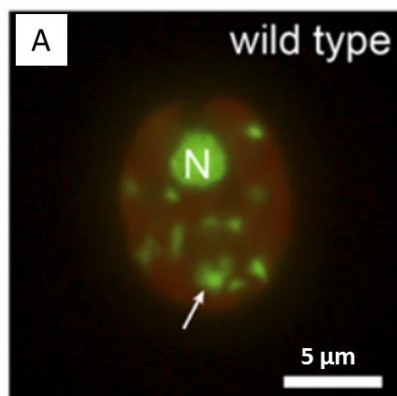
C) Main thylakoid complexes implicated in the photosynthesis (Electron transport chain). Light absorption takes places in two photosystems PSI and PSII surrounded by chlorophyll-rich antennae proteins that increase light absorption surface. Electrons coming from water oxidation travel through the membrane as indicated by dashed, green arrow. Final electron acceptor and the light-photosynthesis main product – NADPH marked in green. ATP is produced via ATP synthase activity that uses proton gradient generated on both sides of the membrane by the action of PSII and cytochrome b6f. For more details see Eberhard et al 2008¹³. LHC – antennae light harvesting complex; PSII – photosystem II; PSI – photosystem I; PQH/PQ – plastoquinol/plastoquinone; PC – plastocyanin; ATP synthase.

Plastids form a family of organelles which consists of chloroplast – organelles where photosynthesis takes place, amyloplasts – a starch storing compartment and chromoplasts that contain large amount of carotenoids and give their distinctive colors to flowers and fruits. The main member of the family and the subject of our interest is chloroplast, an organelle of a significant size that resembles its cyanobacterial ancestor. It is coated with 2- to 4- membrane envelope (number depending on the endosymbiosis events²²) and is filled with membranes – thylakoids. Thylakoids are organized in tight packs called grana or unstacked parts called lamellae. Internal thylakoid membranes are surrounded by stroma - a cytoplasm-like medium, and they themselves enclose part of it creating vesicles whose interior is called lumen (see Fig. 1.1). This compartmentation allows the generation of the electrochemical gradient between the stroma and lumen that is used to generate ATP. Thylakoids contain a multitude of proteins, most importantly the complexes involved in the photosynthetic light reactions and the chlorophyll binding proteins responsible for light absorption and the characteristic, deep green color of chloroplasts (Fig. 1.1). Stroma is a protein-dense hydrophilic medium containing, mainly Calvin Benson Basham cycle enzymes responsible for the "light-independent" part of the photosynthesis - namely carbon fixation, reduction and substrate regeneration, as well as proteins composing the transcription and translation machinery, among which the abundant chloroplast ribosomes.

1.1.2. Chloroplast genome organization

Chloroplast contains also a circular genome (linear forms can also exist²³), a vestige of its ancestor's chromosome (as first shown by Margulis in 1971²⁴). It is usually organized in a 4-part structure with a large single copy (LSC), a small single copy (SSC) parts divided by two inverted repeats (IRA and IRB). It is present in up to 100 copies in each of the cell's chloroplasts (different species containing from 1 to 100 chloroplasts) and is organized in nucleoids attached to internal membranes of the chloroplast and containing in addition multitude of proteins responsible for cpDNA organization, transcription etc.^{25,26} (Fig. 1.2). Nowadays, depending on the species, chloroplast genome contains from ~20 to up to 209 genes (~20 in some Dinoflagellates and 251 in *Porphyra purpurea*). They are mostly organized in polycistronic units, resembling bacterial operons, but contrary to bacteria, the polycistrons usually do not represent functional units (with some exceptions like some ribosomal encoding genes). Rather, they are post-transcriptionally cleaved and proteins are being synthesized from monocistronic units. The core of chloroplast genome consists of i) photosynthesis related genes coding for subunits from photosystem I, II, cytochrome b6f, ATP synthase and large subunit of Rubisco (LSU) (Fig. 1.2), ii) genes coding for subunits of the translation

machinery iii) at least three genes of plastid encoded, bacterial-like RNA polymerase subunits, iv) rRNA genes and v) 27+ tRNA genes²⁷ (can be less in fragmented genomes of diatoms). In addition, the chloroplast genome contains genes involved in lipid biosynthesis, post-transcriptional modifications: *matK* in mustard²⁸, protein assembly and quality control (in *Chlamydomonas* *cemA*, *ccsA*, *clpP* and *ycf3*, *ycf4*) and a number of chloroplast open reading frames (cpORFs) of no assigned function (and who may not be transcribed). Finally, hypothetical chloroplast open reading frames (*ycf*) encoding for proteins of (mostly) unknown function can be found in plastids and cyanobacteria. Most notably, two conserved *ycf* genes (*ycf1* and 2), making for the biggest ORFs in the chloroplast genome are present in green lineage but are absent in the rest of Archaeplastida and in most Monocots. As shown in *Chlamydomonas*²⁹ and tobacco³⁰ those two genes were indispensable for cell survival, yet their function remained uncovered. Only recently was it proposed that Ycf1 (renamed TIC214) could be the only plastid-encoded subunit of the translocation machinery (creating a new category of chloroplast encoded-genes)³¹.



| B | <i>C reinhardtii</i> | | <i>N. punctiforme</i> |
|-----------------|----------------------|-------------|-----------------------|
| | Genome | | |
| | Chloroplast | Nuclear | |
| Size [bp] | 203 395 | 120 404 952 | 89 000 000 M |
| # of genes | 99 | 14 560 | 7 432 |
| Organisation | Circular | Linear | Circular |
| # of copies | Up to 100 | 1* | 1 |
| Introns | Yes | Yes | No |
| % of coding DNA | 38 | | 71 |
| % GC content | 34,6 | 64 | 41,5 |

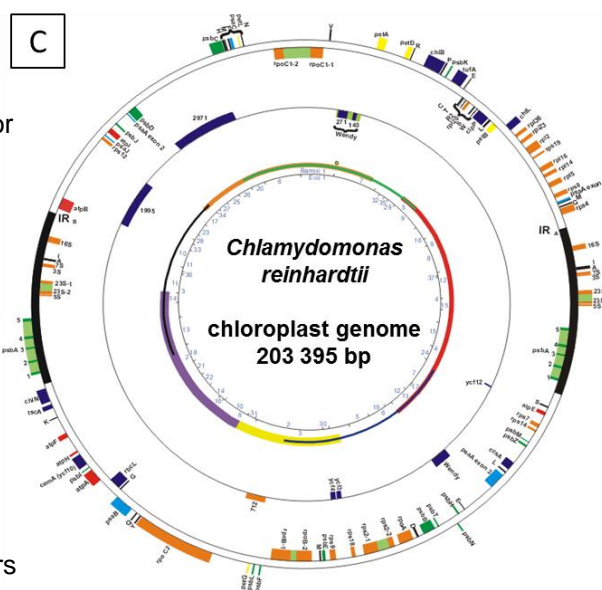
* on 17 chromosomes

Fig. 1.2:
Chloroplast genome of *C. reinhardtii*

A) Representative cell stained with DAPI for DNA highlighting (green). Image merged with the fluorescence of the chlorophyll (deep red). White arrow points one of the nucleoids present in the chloroplast; N – Nucleus.

B) Comparison between genomes of *C. reinhardtii* and the one of the cyanobacteria phylogenetically closely related to the chloroplast.

C) Map of *Chlamydomonas* chloroplast genome with all 104 genes marked in colors (for detailed description see: <http://chlamycollection.org/chloro/genome>)



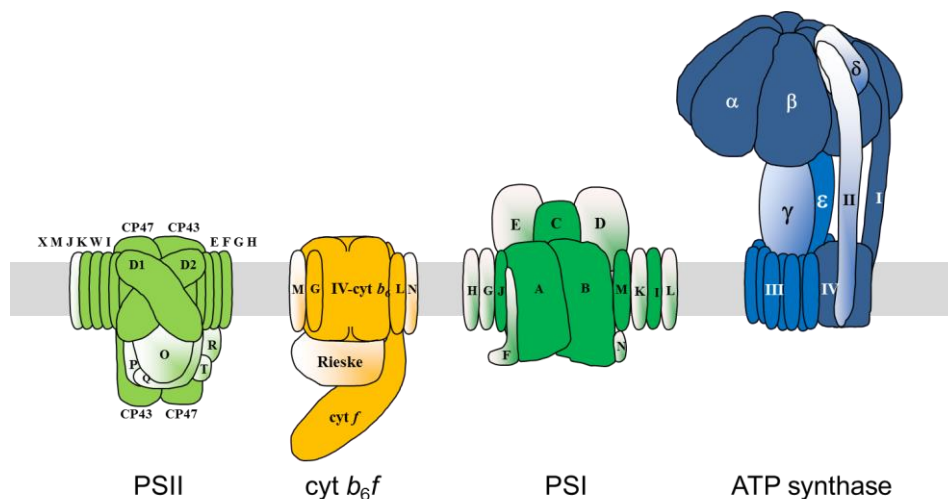


Fig. 1.3: Photosynthetic complexes are mosaics of proteins of dual origin.

In *Chlamydomonas reinhardtii* four main protein complexes of photosynthetic light reaction are constituted in total of more than 50 subunits. Precise regulation is needed to express each subunit in correct amount and assembly them all in a correct manner. Here in full color, subunits coded in the chloroplast; in light gradient subunits originated from nuclear genes.

1.1.3. Endosymbiotic gene transfer: facts and hypotheses

Chlamydomonas reinhardtii chloroplast genome contains 73 putative protein-coding genes, 5 rRNA genes, 30 tRNA genes and 1 small RNA-coding gene³². It stands in bright contrast with cyanobacteria suggested to be the closest chloroplast ancestors: *Nostoc punctiforme*'s genome contains 7500 ORFs³³ and the recently discovered *Gloeomargarita lithophora* genome has 2990 ORFs²⁰. It might be partially explained by the fact that the engulfed symbiont does not need to maintain many genes necessary for its autonomy as it is protected by its host and can sap from its metabolism (e.g. genes coding for the proteins of cell-wall synthesis). Moreover both host and the symbiont had to share certain genes which after endosymbiosis become redundant (e.g. protein quality control genes). However, the most important reason for the chloroplast genome reduction is the export of many of its genes to the nuclear genome. This transfer is supposed to be advantageous as it allows simpler regulation of gene expression and reduces the energetic costs of the gene expression. On the other hand, gene transfer is most probably challenging for the cell as the transferred genes, had to acquire the features necessary for the eukaryotic expression system to be retained (capture of a suitable promoter, and of a downstream AT-rich region for proper transcription termination)³⁴. Today, Arabidopsis nuclear genome contains up to 1700 genes that most

probably originated from the symbiont and are targeted back to the chloroplast³⁵. As shown by the phylogenetic analysis of both genomes, this transfer was most probably done sequentially and the acquisition was very variable in terms of the number of cpDNA fragments and their size³⁶. Interestingly, plastid gene gain has been observed for only three cases already mentioned above: *matK*, *ycf1* and *ycf2* that seemingly appeared in the common ancestor of green algae and plants^{32,37}.

A question arises then: why did all the plastid genes have not been exported to the nucleus? First, translation and following assembly of some subunits of photosynthetic complexes might be coupled and thus would require the presence of a given gene in the chloroplast^{5,38} (see CES process). Similarly, certain proteins may require, for their correct assembly the simultaneous presence of their cofactors (e.g. pigments) that are synthesized in the chloroplast³⁹. Protein synthesis from photosynthetic complexes might also require fast adaptation to the redox state changes within the organelle, denoted as the CoRR hypothesis (Colocation of Redox Regulation). Retaining them within the chloroplast would allow rapid sensing and response to the stress⁴⁰. This might simplify the regulation of these proteins' expression in organisms, like higher plants which contain multiple chloroplasts, by directly influencing only the concerned organelles without a need to send signals to the nucleus⁴¹. Another hypothesis is that, plastid gene expression has to be retained as it has now another function unlinked to photosynthesis as exemplified by *Rhizanthella gardneri*, a parasitic, non-photosynthetic plant which kept now the smallest chloroplast genome of land plants (33 active ORFs)⁴². Despite all that, RNA translation, maturation and processing have been detected suggesting that an "active" plastid is necessary for plants' viability. In photosynthetic organisms many genes retained in the chloroplast encode for subunits of photosynthetic complexes. The majority of those subunits are highly hydrophobic, membrane proteins which would be difficult to import to the chloroplast⁴³. Similar conclusion was drawn for mitochondria where the two genes conserved in all known mitochondria encode for hydrophobic apocytochrome *b* and subunit I of cytochrome oxidase. Elegant experiment demonstrated that import of the proteins to the mitochondria is perturbed after reaching a threshold of local hydrophobicity within the protein molecule⁴⁴. Exceptions in the chloroplast, e.g. highly hydrophobic light harvesting complex antennae proteins are efficiently transported to the thylakoid membranes from the cytosol does not exclude this hypothesis as they do not contain a large total number of hydrophobic transmembrane helices. On the other hand, soluble bacterial RNA polymerase subunits, subunit P1 of the plastid *CtpP* protease and similarly, soluble large subunit of Rubisco are still being encoded in the chloroplast. The same goes for PSII single transmembrane helix-subunits

PsbN and PsbH that are retained in the chloroplast but whose artificial transfer to the nucleus in tobacco and Arabidopsis allows complementation of mutants, unable to express these genes^{45,46}. And finally, the nucleus gene transfer might be a long process that is still ongoing. The first example of an organism without cpDNA has been reported recently, and further analysis of its plastome and nuclear genome might shed light on what a “minimal” plastid is⁴⁷.

1.1.4. Protein rerouting to the chloroplast

With the endosymbiotic gene transfer (EGT), a second mechanism had to evolve to allow relocation of transferred gene products back to the chloroplast to retain its proper functioning. Chlamydomonas chloroplast might contain more than 3000 proteins (~3000 predicted and 996 Mass-spec detected to be chloroplast localized) compared to 73 protein coding genes⁴⁸. This is achieved by the establishment of transport machinery through a combination of prokaryotic and eukaryotic components allowing protein targeting from cytoplasm to the chloroplast. Two multimeric complexes now form a major, import channel in the chloroplast envelope: TOC (Translocon at outer envelope membrane) and TIC (Translocon at internal envelope membrane) respectively (see^{49,50} and reference within). Nevertheless, the origins and the mechanism that allowed creation of the plastid translocation process is still debated, as the exact functions of the progenitors of the TIC/TOC complexes in the symbiont and host are unknown. In parallel, host cells developed a chloroplast protein addressing system. Most of the chloroplast proteins found nowadays exhibit a short, cleavable, N-terminal sequence called a transit peptide that is recognized by the chloroplast's envelope GTP-dependent TOC receptors that direct them to the translocation complexes⁵¹. Despite not being conserved on the sequence level, transit peptides share common features: a similar size ranging from 20 to 100 residues and amino acid constitution (rich in hydroxylated residues and bearing a relatively low percentage of acidic residues) and ability to form amphipathic helices while in contact with lipid bilayers⁵². After serving its purpose, the signal peptide is cleaved in the stroma by processing peptidase and the precursor is further sorted according to its internal function⁵³. Because of their high diversity little is known about the transit peptides' origin. Nevertheless they share common features with antimicrobial peptides constituting the cells' immune system which might suggest that they could have arisen from a priori hostile interactions between symbiont and host⁵⁴.

1.2. Part II Nucleus and chloroplast crosstalk

Integration of the prokaryotic symbiont into a eukaryotic cell gave rise to a bi-original system with spatially separated genomes that in the end will built a photosynthetic machinery (Fig. 1.3) of extreme importance (among others), suggesting a requirement for a tight control of expression to produce the correct amounts of the necessary proteins. Therefore mechanisms of nucleus-to-organelle (anterograde) and organelle-to-nucleus (retrograde) perception and signaling exist to maintain the coordinated expression and function of compartmentalized genomes (Fig. 1.4). First, as a large majority of plastid proteome is encoded in the nucleus, its abundance and regulation is directly dependent on nuclear transcription and cytoplasmic translation. Next, as the chloroplast genome does not contain any gene expression regulators, it therefore relies on the nuclear-encoded factors that are imported to control gene expression at translational and post-translational levels. Finally nuclear-encoded factors are also responsible for protein modifications, transport and assembly of protein complexes as well as regulation of their activity. Occurrence of these regulatory steps is, in utmost case represented by the differentiation of organisms' plasmids into chloro-, chromo- or amyloplasts that *a priori* contain the same gene set^{55,56}.

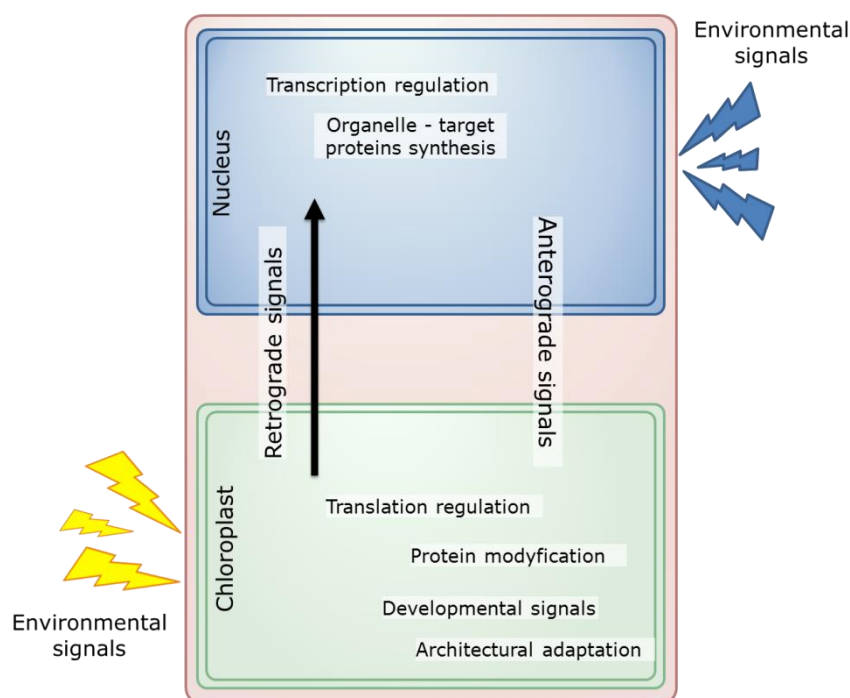


Fig. 1.4: Nucleus-chloroplast crosstalk

Environmental signals (light, nutrients, temperature, developmental cues) sensed by the cell influence expression of nuclear genes by either modulating production of subunits of chloroplast complexes or proteins that will be targeted to chloroplast to affect chloroplast gene expression.

1.2.1. Chloroplast transcription and translation apparatus and regulation

Transcription in chloroplast

Contrary to bacteria, chloroplast transcripts are rather long-lived as demonstrated by Eberhard et al.; despite inhibition of chloroplast transcription through a rifampicin treatment, the global transcripts' abundance (and translation rate) does not change throughout the duration of the experiment⁵⁷. As demonstrated in *Chlamydomonas*, changing amounts of organelle transcripts rarely is reflected at protein levels pointing to the fact that the major anterograde regulation is exerted post-transcriptionally^{57,58}. Specific regulation of chloroplast gene expression at a transcriptional level is limited, but it has however major implications during chloroplast and plant development⁵⁹. In *Chlamydomonas*, mRNA changes are due to the activity of the plastid encoded RNA polymerase (PEP) – the sole enzyme responsible for chloroplast transcription, which is controlled by a single nuclear-encoded sigma factor, *SIG1*⁶⁰, whose expression follows the life cycle's changes and nutritional unbalance^{60,61}. This suggests that transcription's gene specificity is rather limited and undergoes global changes. In contrast to unicellular alga, chloroplasts from higher plants develop from (non-photosynthetic) proplastids, a process which is accompanied with massive production of proteins to form the photosynthetic apparatus. In maize chloroplast, this differentiation parallels with an increase in mRNA levels and translation rates of a multitude of genes, which together account for the increased production of a given protein⁶². Moreover, a lower DNA amount observed in young maize chloroplast from a chloroplast DNA polymerase mutant was correlated with lower RNAs amount. This was followed by decreased expression of photosynthetic proteins. This might be the result of rRNA titration during intensive translation that accompanies chloroplast development as the mutant phenotype is less severe in the developed part of a leaf⁶³. According to the authors' interpretation, an increase in the mRNA level of photosynthetic genes would be necessary in the biogenesis phase but of lesser importance in mature chloroplast. In addition higher plants' PEP integrates multiple sigma subunits⁶⁴ and operates in parallel with a second, nucleus-encoded RNA polymerase (NEP). The relative NEP and PEP activities would be key determinants in the transition between plastid development and differentiation⁵⁶.

mRNA degradation in chloroplast

Similarities of the chloroplast mRNA degradation machinery to cyanobacterial ancestors are represented by the conservation of orthologs of bacterial RNases found in the chloroplast. A general feature of chloroplast mRNA is their intensive 5' and 3' mRNA processing as most of transcripts undergo maturation. In *Chlamydomonas* only three

photosynthesis-related genes are present as primary transcripts: *atpH*, *petA* and *rbcL*. Both 5'-3'⁶⁵, 3'-5'⁶⁶ exoribonucleolytic processing were reported. The 3'-5' employs the exonucleolytic activity of polynucleotide phosphorylase (PNPase)⁶⁷ or starts with an endonucleolytic cleavage to eliminate possible staling mRNA secondary structures. Ends of the 3' processing can be defined by the same secondary structures or proteins binding to the RNA that block the progression of the enzyme. Polyadenylation of the structurally-blocked RNA can reinitiate the decay. The 5' processing would also result from a 5'-3' exonucleolytic process. Experimental evidence for the participation of OTAFS in the 5' maturation was obtained by the insertion of poly(G) cage in *Chlamydomonas* 5'UTR mRNA sequences, which were found to stabilize the mRNA in OTAF mutant background (most probably by blocking the RNase). The combined action of endo- and exonucleolytic activity of RNases is also necessary for the intercistronic cleavage of polycistronic transcripts. Again, the activity or processivity of normally unspecific RNases is limited by the RNA-binding proteins (OTAFs; see below)⁶⁸. For the RNases and details see⁶⁹ and reference within.

Translation machinery

At the same time, the translation machinery of the chloroplast is much conserved and most of its components are homologous to their cyanobacterial counterparts. Translation is initiated by 30S ribosomal subunit binding to the translation initiation site (AUG is the predominant initiation codon). Contrary to bacteria however, two thirds of plastid genes lack Shine-Dalgarno sequence⁷⁰ and thus require a different initiation mechanism, which is however still not well understood. One of the possibilities is that nucleus-encoded gene specific factors could be required to bind the transcript and mediate the interaction with 30S subunit or translation initiation factors (see OTAFs). Otherwise they could reshape the mRNA (secondary structures) of those transcripts to facilitate the targeting to the ribosome entry site (see OTAFs). This proposition is tempting in the light of the observed structural differences between bacterial and plastid ribosomes, as the chloroplast ribosome entry site is narrower than in bacteria due to extended (compared to bacteria) polypeptide chains of some subunit and by that, may accept only unstructured mRNA⁷¹. Initiation and further elongation is assisted by homologs of bacterial initiation and elongation factors which are imported from the nucleus⁷², with the exception of the chloroplast-encoded algae-specific elongation factor- Tu (*tufA* gene)⁷³. Plastid ribosomes share high structural similarity with those of bacteria with some small but significant differences⁷⁴. They are composed of the 50S and 30S subunits constituting a 70S ribosome⁷⁵. Each subunit consists of many proteins and rRNAs. Ribosomal 30S subunit contains one, 16S rRNA whereas 50S subunit includes

23S, 5S and 4.5S rRNA molecules. All of them are encoded by chloroplast genes situated in the inverse repeat (IR) regions of the genome. The 4.5S rRNA particle is not present in bacteria, however is highly homologous to the C-end of bacterial 23S rRNA. 24 and 33 proteins are part of 30S and 50S ribosomal particle respectively^{76,77}. They are distributed between both, nuclear and chloroplast genome. Five additional ribosomal proteins were identified in the plastid as compared to *E.coli* (PSRP2-6) (two are part of 30S and 3 of 50S particles). Moreover, some of the ribosomal proteins carry N- and C-terminal extensions that stabilize the structure of both subunits and account for differences at the entry and exit sites of the ribosome⁷⁴. All transfer RNAs (tRNAs) are spread in the chloroplast genome, that however contains only up to 30 tRNAs (30 in *Chlamydomonas*), which is less than the minimal set of 32 (due to the wobble). In addition, there was no demonstration of the import of the missing tRNAs to the chloroplast. Notwithstanding, plastids are able of synthesizing their proteins with a reduced tRNA set due the superwobbling phenomenon – i.e. the ability of a given tRNA to read all four codons sharing their first two nucleotides (see^{78,79}). This is due to the possibility of the relatively small uridine (U) to interact weakly with all four base pairs. Consequently, the minimal set of tRNAs would be reduced to 25, which is less than the smallest number found in photosynthetic algae and plants (27). At the same time in non-photosynthetic organisms this set is even less reduced (to 4 in *Rhizanthella gardneri*)⁴².

Organization of translation

Chloroplast translation localization has been since years a matter of discussion. Initial works reported that translation always starts at stromal ribosomes and some of them are later attached to the membrane via nascent polypeptide⁸⁰. It was suggested later that, some proteins, are translated on membrane associated ribosomes⁸¹. In a recent study authors used a ribosome profiling technique to analyze the transcripts associated to membrane-bound ribosomes. They concluded that out of 37 chloroplast-encoded membrane proteins in maize, 19 are co-translationally targeted to the membrane (translation is initiated on free ribosomes that then bind to the membrane)⁸². In *Chlamydomonas* it was proposed that zones of intense membrane translation exist in the chloroplast, from whom the biogenesis of PS complexes would originate. These so called T-zones have been identified by fluorescence microscopy probing for mRNA and subunits of PSII. They are localized at the outer periphery of the pyrenoid and after biochemical evidence contain Ribosome-dense chloroplast translation-dedicated membranes (CTM)⁸³. However their exact nature has not been completely elucidated^{83,84}. Similarly, the translation of the pyrenoid localized LSU was shown to be localized to the basal region of the chloroplast⁸⁵. Neither T-zones nor localized

translation of LSU were observed in higher plants – maybe because of different biogenesis of the chloroplast and/or lack of the structural organization conferred by the pyrenoid.

1.2.2. Nucleus-encoded regulators of chloroplast gene expression

Each chloroplast gene takes a long road till its product can fulfill its function. After correct transcription, usually, a polycistronic, primary transcript needs to i) be cleaved into mono- or di-cistronic mRNAs, ii) get its introns spliced out, iii) be trimmed (at 5' and 3' ends) and sometimes iv) edited⁸⁶. A mature transcript undergoes translation that needs to be i) timely and ii) efficient enough to match the cell's needs. It is generally accepted nowadays that nucleus-encoded RNA binding factors play a role and can regulate all of these processes^{69,87}. Classical RNA-binding proteins containing: Zinc-finger domains⁸⁸, RRM domains⁸⁹, or CRM domains are usually responsible for stabilization, splicing (e.g. APO1 type II intron splicing factor⁸⁸) and editing of RNA⁸⁶ (for more see^{90,91}). A vast number of chloroplast RNA-associated processes have now been connected to the action of proteins belonging to a helical-repeat proteins super-family and now coined "regulators of organelle gene expression "(ROGEs)"⁸⁷ or organelle specific gene expression factors⁹², hereafter called OTAF (for organellar Trans-Acting Factors). Most of them form tandems of α -helices organized in two- or three-fold that form a solenoid-like platform for nucleic acid and/or protein interaction⁹³. Although evolutionary independent, they are characterized by the presence of a degenerated, repeated, helical motif of different lengths: tetratricopeptide proteins (TPR) having a module of 34 aa length, pentatricopeptide proteins (PPR; 35 aa per helix), octatricopeptide proteins (OPR) of 38 aa and functionally, closely related mitochondrial transcription termination factors (mTERF).

TPR

Through broad genomic database analysis, it was shown that TPRs are widely present in photosynthetic organisms with 29 TPRs discovered in *Synechocystis* and more than one hundred in *Chlamydomonas* and *Arabidopsis*⁹⁴. However, most of them localize in the cytoplasm as on average only one fifth of them could be predicted with available softwares to be targeted to the chloroplast⁹⁴. Helical motif repeats are found from 1-32 times but on average TPRs contain 2 helical repeats per protein in *Synechocystis* and 3 repeats in photosynthetic eukaryotes⁹⁴. Alone or as part of multimeric complexes TPRs are principally involved in a variety of protein-protein interactions,⁹⁵ and are engaged in a

multitude of processes, such as: transcriptional regulation, RNA metabolism, folding and protein transport⁵¹ (see also^{95,96}). Due to their significant implication in RNA processing it was suggested that some TPR proteins could interact directly with RNA molecules⁹² (R-TPR or RNA-TPR)⁹⁷. This led to the discovery and description of different TPR subfamilies containing variants of a classical TPR motif: cl-/crnTPR (crooked neck-like TPR)⁹⁸ or HAT (half-a-TPR)⁹⁹. The function of TPR proteins in the chloroplast is illustrated by (at least) 4 factors required for RNA processing discovered in *Chlamydomonas* (Table 1.2) or by the *Arabidopsis* HCF107 HAT protein, which can bind RNA *in vitro*, in a sequence-specific manner illustrating its role in the regulation of chloroplast gene expression¹⁰⁰.

PPR

Closely related to TPR, PPR proteins have been first described in *Arabidopsis* through genome and proteome analysis¹⁰¹. They are almost exclusively found in eukaryotic organelles: mitochondria and plastids (on average up to 30 PPRs in non-photosynthetic organisms)^{102,103}. Exception to that would be some pathogenic/symbiotic prokaryotes that most probably acquired some PPR-coding genes through horizontal gene transfer from their hosts¹⁰⁴. This limited presence points that they probably emerged at the beginning of eukaryotic evolution. Higher plants' PPR are particularly striking, as within land plants, the PPR family has greatly expanded with more than 400 members reported up to date¹⁰² (with more than 1000 genes present in some genomes, as compared to e.g. 14 PPRs in *Chlamydomonas*¹⁰⁵) suggesting the family's proliferation during plants' colonization of the land¹⁰⁶. Indeed through RNA stabilization, processing, splicing and editing (reviewed in¹⁰⁷) different PPRs affect photosynthetic apparatus formation, plant development, pollen fertility, and stress sensing (reviewed in¹⁰⁸). PPR-RNA interaction has been demonstrated through biochemical (Co-IP¹⁰⁹, *in vitro* binding^{109,110}, RIP-Chip¹¹¹), genome-wide ribosome profiling¹¹²) and computational approaches¹¹³. *In silico* analysis allowed discovering correlations between the identities of amino acid residues in the PPR protein and the mRNA nucleotides it recognizes. Recoding of maize PPR protein (PPR10) and using designed PPRs confirmed the predictions and established this "PPR code" (for P-PPR¹¹⁴ and S-PPR¹¹⁵) that proposes a matrix of single aa- single nucleotide pairings that define PPR's sequence specificity^{68,108}. Structural proof of the aa-nucleotide recognition and combinations was obtained through RNA-PPR co-crystallization^{115,116}. However, the code is not perfect, partially degenerated so that different aa specify the same nucleotide but can bind it with different affinities. This limits the further development of designed PPRs⁶⁸.

Two different subfamilies of PPR were described up to now. Canonical PPRs called P-class PPRs contain 2-30 repetitions of a classical 35 aa motif (with some special cases of proteins with additional domains¹¹⁷). This class is mainly involved in the transcript's maturation⁴⁶, stabilization^{118,109}, splicing^{119,120}, and translation^{111,121,122,123,124} processes. Another, PLS-class is present only in *Streptophyta* clade (with the exception of some parasitic protists¹²⁵) and would be involved more specifically in the editing process as site recognition factors^{126,127}. Proteins from this subfamily contain additional, degenerated variants of the P motif (L-motif: 35-36 aa; S-motif: 31 aa) each with different conservation of amino acids positions within the repeat¹²⁸. This subfamily usually possesses also an E or DYW domains at the C-terminus. As the DYW bears a cytosine deaminase motif signature, DYW-containing PPR have been proposed as candidates for deamination of cytidines in RNA strands¹²⁹. This hypothesis is still a matter of debate as the DYW was found to be sometimes dispensable¹³⁰, although this might be due to trans-complementation. Yet definite biochemical proof of a cytosine deaminase activity is still missing. The editosome would also consist of other non-PPR proteins associated to the PPR recognition factors like proteins from the RIP (RNA-editing Interacting Protein)/MORF (multiple Organellar RNA-editing Factors) family, and proteins from the ORRM or OZ families (reviewed¹²⁷). Importantly, the number of editing sites in land plants (e.g. 488 and 34 in Arabidopsis mitochondria and chloroplast respectively⁸⁶) is positively correlated with the number of PLS PPRs in those organisms, which may imply a specific adaptation of the nuclear genome for the control of the organelle gene expression in the harsher land environment¹³¹ (and may explain the higher number of PPRs in plants as compared to other eukaryotes). However it still unresolved why there is a difference in the number of PPRs acting on RNA editing (~200) and the much higher total number of editing sites (~600)¹²⁷ (this observation also contradicts the model of a gene-specific mode of action of OTAFs: see below). Also, despite their overall high number in higher plants' genome, no redundancy of PPRs has been observed as it has been proposed that PPRs serve as regulators of expression of specific, organelle genes. Subsequently, via analysis of mutants obtained by forward genetics, a wide range of PPRs' physiological and molecular functions has been described. Most PPR mutants in plants have a strong specific effect on gene expression, leading to the idea of a PPR code. Less specific PPRs having multiple targets reveal target sequence similarity, or decreased specificity due to a small number of PPR motifs within the protein^{111,122,132}

OPR

The less-described octotricopeptide repeat protein family contains a degenerated motif of 38 aa that is predicted to arrange, as in the case of TPRs and PPRs into α -

helical tandems (2-24 per protein) that again serve as ligand-binding platforms. The first OPR protein, Tbc2, has been discovered in *Chlamydomonas reinhardtii* chloroplast and was found to be required for the translation initiation of *psbC* gene¹³³ – providing another example of tandem-repeat family protein involvement into organelle gene expression. Interestingly, OPRs are found mostly in unicellular algae (*Chlamydomonas reinhardtii*, *Volvox carterii*), parasitic protists (Apicomplexans) and parasitic bacteria (*Coxiella burnetii*), where they undergo rapid expansion, as witnessed by a NCC-like cluster of 32 OPR paralogs on *Chlamydomonas* chromosome 15¹³⁴. In contrast, only one (as compared to more than 120 in *Chlamydomonas*¹³⁴) record of OPR was discovered in higher plants where it is necessary for 16S (chloroplast) rRNA maturation¹³⁵. OPR are mostly predicted to be targeted to the organelles: chloroplasts and mitochondria. They operate in posttranscriptional regulation of organelle gene expression in translation initiation (Tab1¹³⁶, Tbc2¹³³, Tda1¹³⁷), and RNA maturation with the best described involvement of OPRS in the two-step trans-splicing of the chloroplast *psaA* gene¹³⁸⁻¹⁴⁰ (Splicing supercomplex(es) responsible for this process containing at least 5 OPR proteins¹⁴¹). From the published experiments and structure predictions, OPRs' molecular mechanism of interaction seems to be similar to the PPR proteins. They would interact directly with RNA, as proven in a few cases^{136,142} with a modular: one nucleotide-one aa binding mode, and by such would in some cases provide protective caps against exonucleases and release sRNA footprints, as seen for CrMCG1¹⁴³. The combinatorial code would however differ from the PPR one. Linking sRNA footprints to specific OTAF should help to solve this OPR code. The target specificity could be further provided by additional domains such as FAST and RAP which could provide further RNA interaction as proposed by Eberhard et al.¹³⁷. Sixteen *Chlamydomonas* OPRs have at least one Fas-activated kinase-like domain (FAST) that is found in proteins involved in RNA processing in mitochondria¹⁴⁴. RAP binding domain (RAP: RNA binding abundant in Apicomplexans) has been also found in some OPR (sometimes with the conjunction with the adjacent FAST domain) and has been suggested to interact with RNA as well and to possess nucleolytic activity^{144,145} due to structural homologies to endonucleases¹³⁴. In *Chlamydomonas* OPR presenting an association to the RAP domain form a subfamily of 38 paralogs called the NCL subfamily¹³⁴.

mTERF

mTERF proteins (from: mitochondrial transcription termination factors) were first proposed to play a role in mammalian mitochondria gene expression almost 28 years ago¹⁴⁶. Now it is known that they act on mitochondrial transcription¹⁴⁶, DNA-related functions¹⁴⁷ and organelle ribosome biogenesis¹⁴⁸. A first report of an mTERF from a

photosynthetic organism was published in 2004¹⁴⁹. Nowadays, up to 30 mTERF proteins can be found in higher plant proteomes, a number much higher than in other metazoans¹⁵⁰. They are almost exclusively predicted to be targeted to the mitochondria or chloroplast with no characterized example of cytoplasmic mTERFs. Their main characteristic is, similarly to other families described above, the presence of a degenerated amino acid motif of around 30 residues that forms a pair of antiparallel α -helices with usually an adjacent helix that works as a ligand-recognition scaffold¹⁵¹. Usually, tandems of mTERF repeats organize into a luna/croissant shape, however within the family, their arrangement can vary significantly¹⁵². Knowledge of the role of mTERFs in photosynthetic organisms is still limited to few examples from *Chlamydomonas*¹⁴⁹, maize^{153,154} and *Arabidopsis*^{155,156}. In the two latter cases plastid-targeted mTERF have been shown to be involved in global transcript abundance and splicing, leading to the suggestion that proteins from this family may act on the post-translational regulation of chloroplast gene expression⁹². This conclusion is further supported by the finding of a significant number of mTERF protein associated to the chloroplast nucleoids¹⁵⁷.

Convergent evolution - convergent role?

Altogether, all these families, despite probably not being evolutionary connected (except TPR and PPR sub-families) share structural similarity. They are organized into solenoid-like structure of repeated motifs that sets the base for interaction with other factors. Unfortunately, due to technical difficulties to produce antibodies for modular proteins, a physical proof of this interaction has been made mainly for PPRs¹⁵⁸ (reviewed^{108,159}) and only for few members of TPR^{8,100}, OPR¹³⁶ and mTERF¹⁵¹ families. Despite the difficulties, enormous progress in OTAFs research has been made, mostly due to the analysis of photosynthesis deficient mutants of maize, *Arabidopsis* and *Chlamydomonas*. From the published results emerges the view that plastid RNA metabolism is highly dependent of the nucleus (proving again the vast posttranscriptional regulation of organelle gene expression) and engages proteins of all mentioned sub-families. It is tempting to conclude that members of all those groups could interact directly with RNA and that their role is convergent. A question remains: why is the distribution of e.g. OPRs and PPRs so different: hundreds of PPRs in plants and only one OPR, versus the opposite situation in *Chlamydomonas*. With the PPR code being consistently polished (and emerging OPR code) more mechanistic proofs of protein-RNA interactions are to come soon. Because of PPRs regular modularity of helical-repeat proteins it will be possible to predict their RNA targets *in vivo* to more easily determine their function. With that in hand we could have an insight into the PPRs' advantage over

the classical RNA-binding domain proteins (zinc-finger, RRM domain etc.), that usually consist of multiple, flexibly connected globular domains which makes it difficult to predict their specificity and to modulate it¹⁶⁰, and even engineer artificial RNA-recognition molecules of broad application¹⁶¹.

OTAFs in chloroplast

Based on the analysis of the mutant phenotypes and the characterization of the individual proteins' behavior *in vitro* and *in vivo*, chloroplast OTAFs (whether TPRs, OPRs, PPRs or mTERFS) were found to act on different levels of posttranscriptional gene expression.

First, they play a role in RNA maturation, as caps physically protecting transcripts from exonucleolytic degradation (both at 5'- and 3'-ends) and by that define mature transcript ends. The best described example is the PPR10 protein from Maize chloroplast that binds specific intergenic sequences of the *atpl-atpH* and *psaJ-rpl33* polycistronic transcripts thereby defining the 5' and 3' terminus of the monocistronic unit by preventing their exonucleolytic digestion^{121,162}. OTAFs sometimes even promote the transcript maturation via initiation of endonucleolytic cleavage^{68,163}. Multiple OPRs are involved in *psaA* trans-splicing process in *Chlamydomonas*¹⁴¹. Those OTAFs that have a direct role in mRNA metabolism have traditionally been called "M" factors, a nomenclature particularly used in *Chlamydomonas*.

Next, OTAFs can act on a specific mRNA, usually 5'UTR, to further recruit additional effectors such as ribosome 30S subunit (see Ribosome structure) that is required for translation¹²². They can also themselves, "activate" the transcript by reshaping the RNA structure to allow translation initiation¹⁰⁰. In these cases where their role is linked to translation activation, they were called T-factors. Note that OTAFs are in large part represented by the members of the super helical protein family, but are not restricted to these families, and can also be pioneer proteins. Such translational activators, required for proper translation initiation of a given transcript, have been abundantly described in *Chlamydomonas*. They target the 5'UTR of a given gene¹⁶⁴ either by direct binding or via another, sequence recognizing protein (see above). An example of it is the *tca1* (see Table 1.2) mutant, where the *petA* mRNA accumulated normally but neither translation nor *cyt f* protein could be observed suggesting its implication in the early stage of *petA* translation initiation¹⁶⁵. Similar observations were made for TAB1¹³⁶, TAB2¹⁶⁶, TBC2¹³³ and TDA1¹³⁷. The latter has been also found in dense ribonucleoprotein complexes suggesting another function in sequestering *atpA* transcripts into un-translated "reserve" pool¹³⁷.

Notably, few instances were described where an M factor stabilizing the mRNA cooperates with a T-factor. Genetic analyses suggest that this is the case for the *Chlamydomonas psbD* translation system requiring both Nac2 (MBD1) and RBP40 proteins¹⁶⁷, as well as for the MCA1 –TCA1 couple. Currently, no such coupled activities have been highlighted in higher plants.

OTAFs are in some cases true regulators of gene expression, whose abundance are directly linked to the abundance of their target product, as exemplified for MCA1 (and less-so for TCA1) linked to *cyt. f* accumulation¹⁶⁸. Another example comes from the study of the MAC1 factor whose accumulation level dictates *psaC* transcript accumulation¹⁶⁹. OTAFs' regulatory function is further exemplified in response to changing environments, such as in nutrient starvation as seen for MCA1, TCA1, TAA1¹⁷⁰ or MAC1¹⁶⁹. This point will be further discussed in Chapter 1.

A more comparative analysis is required to assess the evolutionary importance of OTAFs, especially considering the convergent distribution of different sub-families in plants and algae which is coupled to high conservation of their target sequences.

1.2.3. Retrograde signaling

Plastids transfer information about their developmental state and their physiological condition to the nucleus via retrograde signaling. This allows modifying the expression of the nuclear genes which in turn influences plastid gene expression (by an anterograde communication). Chloroplasts are sites of photosynthesis, a process that is heavily influenced by environmental cues (especially light intensity and quality) and as such need to respond rapidly to their fluctuations. Furthermore many other metabolic reactions take place within plastids such as starch and lipid production, amino acid biosynthesis, production of hormones etc., and it is thus obvious that they need to adapt to the physiological changes sensed by the chloroplast. From this perspective, chloroplast evolved not only as a “photosynthesis factories” but also as sensors for plants. As nicely demonstrated by Feng et al. chloroplasts play major function at the onset of *Arabidopsis* flowering by perceiving light necessary for this transition¹⁷¹.

The multitude of reactions taking place in the plastids and their evident interconnectivity suggest that different retrograde signals exist and cross-talk to communicate the various needs of physiological rearrangements. The first observation of retrograde signaling in plants dates back to the work of Bradbeer et al., when authors observed that in plants containing undeveloped plastids, synthesis of photosynthesis-related nuclear genes was inhibited¹⁷². Nowadays, a broad range of retrograde signals (more than 40) have been

discovered (reviewed¹⁷³) and shown to operate in plastid biogenesis (biogenic signals), photosynthesis regulation (operational signals), and individual chloroplast quality control¹⁷⁴. A genetic screen designed to pinpoint retrograde signaling factors by using a carotenoid biosynthesis inhibitor (norflurazon), allowed to discover six *gun* (genomes uncoupled) mutants that do not exhibit retrograde repression of a specific set of nuclear genes (coined by the authors PhANGS: Photosynthesis-associated nuclear genes) in the presence of the inhibitor¹⁷⁵. Five of those mutants: *gun2-6* lack proteins involved in tetrapyrrole biosynthesis pathway, suggesting that tetrapyrroles and/or their derivatives are implicated in the plastid to nucleus signaling. Interestingly, the last mutant is affected in the GUN1 protein, a chloroplast PPR-SMR protein showing a role in chloroplast translation and linking retrograde tetrapyrrole signals to anterograde signaling¹⁷⁶. It was shown to associate with nucleoids in the chloroplast stroma and to interact with plastid ribosomal protein S1 (PRPS1) and thus, might be implicated in transcriptional and/or posttranslational regulation of gene expression. Besides those, redox signals (oxygen radicals)(reviewed in¹⁷⁷), secondary metabolites¹⁷⁸ and transcription factors¹⁷⁹ have been shown to act in the chloroplast-nucleus communication. All in all, the exact role and mechanism of action of many of those factors is not yet elucidated but it is possible that expression of OTAF could be the target for retrograde signaling cascades, allowing this way to re-address the signal to the chloroplast¹⁷³.

Table 1.2: M and T factors found in *Chlamydomonas reinhardtii*.

| Complex | Chloroplast gene | M factor | Protein family | RNA target | Ref. | T factor | Protein family | Ref. |
|----------------------|------------------|----------|----------------|---------------------|---------|----------|----------------|------|
| PSI | psaA | RAA1 | OPR | psaA Intron 1 and 2 | 140 | TAA1 | OPR | 170 |
| | | RAA2 | hPUS* | psaA Intron 2 | 180 | | | |
| | | RAA3 | OPR | psaA Intron 1 | 181 | | | |
| | | RAA4 | | | 142 | | | |
| | | RAA7 | pioneer | psaA Intron 2 | 182 | | | |
| | | RAA8 | OPR | psaA Intron 1 | 139 | | | |
| | | RAA6 | pioneer | psaA Intron 2 | 141 | | | |
| | | RAA7-13 | - | | | | | |
| | | RAT1 | hpAP | tscA | 138 | | | |
| | | RAT2 | OPR | | 138 | | | |
| | psaB | MAB1 | mTERF | 5'UTR | | TAB1 | OPR | 136 |
| Cyt b ₆ f | psaC | MAC1 | TPR | 5'UTR | 169 | TAB2 | | 166 |
| | petA | MCA1 | PPR | 5'UTR | 165 | TCA1 | pioneer | 164 |
| | petB | MCB1 | PPR | - | 6 | MCB1 | PPR | 6 |
| | petD | MCD1 | OPR | 5'UTR | 183,184 | | | |
| | petG | MCG1 | OPR | | 143 | | | |
| PSII | psbA | RBP63 | - | - | 185 | | | |
| | psbB | MBB1 | TPR | 5'UTR | 186 | | | |
| | psbC | MBC1 | OPR | - | a | TBC1 | OPR | 133 |
| | | | | 5'UTR | | TBC2 | | |
| | psbD | MBD1 | TPR | 5'UTR | 187 | RBP40 | - | 167 |
| | psbH | MBB1 | TPR | | 188 | | | |
| ATPase | psbI | MBI1 | OPR | | 143 | | | |
| | atpA | MDA1 | OPR | | a | TDA1 | OPR | 137 |
| | atpB | MDB1 | OPR | | b | | | |
| | atpE | MDE1 | - | | c | | | |
| Rubisco | rbcL | MRL1 | PPR | | d | | | |
| | | | | | 1 | | | |

hPUS – homologous to pseudouridine synthase

hpAP – homologous to poly(ADP-ribose) polymerase

a- Viola S., et al. (in prep)

b- Cavaiuolo M., et al. (in prep)

c- Drapier D., et al. (unpublished)

d- Osawa S-I., et al. (in prep)

1.3. Part III: Regulatory processes involved in photosynthetic complex assembly

1.3.1. Concerted accumulation of subunits

In cyanobacteria, most photosynthetic genes form operons and most-but not all- subunits of a same protein complex are synthesized simultaneously in a given ratio. Because of this “central” regulation, accumulation of individual subunits does not necessarily affect the fate of the others. For example: D2 subunit of PSII was shown to accumulate partially in the absence of another core-subunit D1¹⁸⁹. As well, in *Synechocystis 6803*, the absence of the peripheral subunits of PSII (PsbH, PsbI, PsbK) did not result in the loss of phototrophy or drastic decrease in the accumulation of the core of the complex¹⁹⁰⁻¹⁹³. In eukaryotic cells however, genes coding for subunits of the same complex are encoded in two different genomes (Fig 1.3). As a result they are spatially (different cellular compartments: chloroplast and nucleus) and temporarily (their synthesis is not simultaneous) separated. In addition, photosynthetic complexes require for their proper biogenesis a multitude of cofactors, pigments and ions. Mis-assembly may lead to significant energy waste, unnecessary protein aggregation, or production of dangerous radical species. Although chloroplast genes are submitted to a nuclear control by the action of nucleus-encoded factors that are necessary for the translation of chloroplast-encoded subunits (Table 1; Fig 1.3) those factors do not account (directly) for the coordination of the production of different subunits nor for their subsequent assembly with their nucleus-encoded partners to form functionally active complexes. Yet, concerted accumulation of all the given subunits is ensured by the equilibrated action of two other, distinctive mechanisms.

1.3.2. Proteolysis

First, because of the dual origin of their proteome, chloroplasts (and mitochondria) possess much more elaborated proteolytic machinery than cyanobacteria. It contributes to maturation of both chloroplast-encoded (by f-Methionine deformylation and cleavage, N-terminal processing) and imported proteins (by transit peptide removal, post-translocation processing). Proteases also remove mis- or unfolded proteins in response to environmental stress, contribute to plastid development (especially in plants), amino-acids recovery and cell's senescence (reviewed in¹⁹⁴). In *Chlamydomonas* two proteases complexes in particular contribute to the maintenance of the photosynthetic apparatus: ClpP and FtsH complexes target soluble and membrane subunits of photosynthetic

proteins (reviewed in¹⁹⁵). Proteolytic regulation of the photosynthetic subunits accumulation was already proposed in 1983. Using pulse-chase labelling experiments authors have shown that in *Chlamydomonas*, Rubisco small subunit is probably rapidly proteolized when not assembled with LSU¹⁹⁶. Further research extended this hypothesis for PSI, II, ATPsynthase and *b₆f* complexes suggesting that all photosynthetic multimeric complexes undergo regulation via degradation of unassembled subunits: e.g. turnover of the PSII reaction center protein D2 is highly accelerated in the mutants unable to accumulate its partner D1¹⁹⁷. A similar situation was observed for subunit IV and *b₆* of the *b₆f* complex – in mutants lacking other major subunits, their degradation rate was increased (with unchanged rate of synthesis)¹⁹⁸.

1.3.3. The CES process

Secondly, chloroplast displays a unique-to-organelle process that links post-transcriptional and post-translational regulation of gene expression. In the absence of some crucial/core subunits, translation of their chloroplast-encoded partners is ceased (!). This process can be exemplified by the behavior of *b₆f* complex: translation of *petA* gene (coding for cytochrome *f* – one of cytochrome *b₆f* core subunit's) is diminished to about 10% of the wild-type rate in the absence of either cytochrome *b₆* or subunit IV. At the same time, the level of *petA* mRNA is only slightly decreased which suggests that *petA* expression is limited at the translation level when other *b₆f* core subunits are absent¹⁹⁸. This assembly-dependent regulation of translation was named: control by epistasy of synthesis (CES). This one and similar examples were reported for subunits of all photosynthetic complexes of *Chlamydomonas reinhardtii*, so that throughout the years it became the CES model organism (Table 1.3) (see below: Special case of ATP synthase; Rubisco described in: Regulation of Rubisco expression).

CES cascades follow the assembly process. More details about the physiological importance of this process came with the research on photosystems I and II. Mutants defective in the accumulation of PSI PsaB subunit show virtually no synthesis of PsaA subunit (PsaA is then a CES subunit in the process)¹⁹⁹. At the same time as elegantly shown, restoration of PsaB accumulation in a suppressor strain results in the parallel restoration of PsaA synthesis which demonstrates that PsaB is required for efficient translation of its partner²⁰⁰. Moreover, absence of the PsaA subunit leads to reduced synthesis of PsaC, but not PsaB. On the other hand however, null *psaC* mutant show normal rate of translation for both core subunits PsaA and B, which afterwards are rapidly proteolized²⁰¹, as they cannot further assemble (see proteolysis). An analogous situation exists for PSII complex in which D2 subunit is necessary for D1 translation

which in turn is essential for efficient translation of *psbB* (coding for CP47 core antenna) mRNA^{197,202,203}. Both PSI and PSII are assembled in a step-wise manner with individual subunits being added to a preformed core of the complex. By negative feedback loops, the CES process regulates the formation of downstream subunits when the upstream assembly-intermediate is not formed (described in^{204,205}). Thus, CES subunits' expressions follow the assembly pathway preventing unnecessary production of "CES subunits" (but not of the most upstream ones).

Mechanism

Mechanistically, the mode of regulation seems to be universal for all described instances. In none of the above-mentioned cases, an influence on the transcript level of the CES subunit could be observed in the absence of its regulator (except a slight effect on *petA* mRNA which is however, much lower than on protein level). Radioactive pulse labeling experiments demonstrated that synthesis rate of CES subunits is impacted at a posttranscriptional level either via a direct impact on their translation or by rapid degradation of the polypeptides^{198,204-208}.

A two-way experiment allows distinguishing between those two alternatives. First, swapping the 5'UTR of a CES-regulated gene of interest (GOI) to an unrelated, 5'UTR sequence was sufficient to escape the inhibition, which indicates, that the GOI 5'UTR is the cis-factor of the regulation. Next, by using constructions where the 5'UTRs of the CES-subunit is coupled to a reporter coding sequence, it was possible to show that CES 5'UTR are sufficient to confer translation inhibition to an unrelated reporter protein²⁰⁸. It is highly improbable that unrelated genes would share translation trans-factors that trigger their co-translational degradation thus; with the two observations in consideration, the actual mechanism leading to lower accumulation of CES-subunits is the regulation (inhibition) of translation initiation of their genes^{208,209}. Since its first description for *petA* gene of cytochrome *b₆f* complex, similar observations were made for all the above-mentioned examples. Nevertheless, a translation limitation could still be explained either by the hypothesis that the inhibition requires the CES-subunit as an effector (auto-inhibition) or that an assembly partner is necessary for the activation of the CES-subunit's translation (trans-activation). Assessment of the valid hypothesis was possible by combining the use of the reporter gene driven by a 5'UTR of a CES subunit with the deletion/removal of both the CES-subunit and its assembly partner (Fig 1.5). Insensitivity of the reporter gene accumulation to the presence of the CES-subunit would mean that CES-subunit is trans-activated in normal conditions as its regulation should be solely

regulated by the activating partner. On the other hand, for the autoregulation to occur, the CES-subunit needs

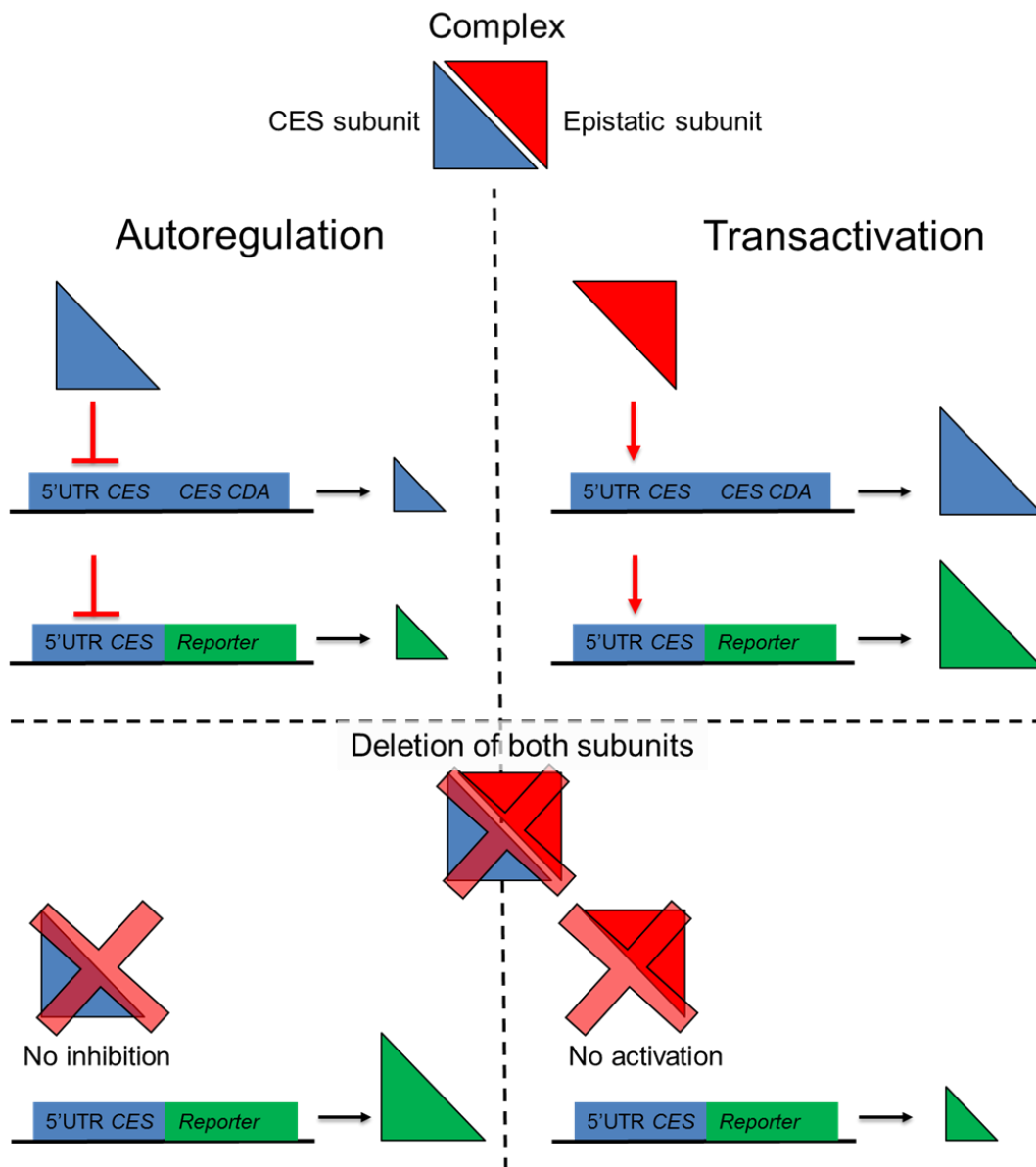


Fig. 1.5: Possible mechanism of CES regulation

Top – CES subunits' regulation can be explained by two mechanisms: Autoregulation (left) and Transactivation (right). When unassembled with its partner, a CES subunit can either inhibit its own translation (CES subunit) and the translation of a 5'UTR CES-driven reporter gene (left) or its translation can be activated by epistatic subunit when its unassembled (right).

Bottom – By deletion of both subunits one can discriminate between the two mechanisms. Absence of both subunits should result in different levels of accumulation of the reporter gene depending on the occurring mechanism⁵.

to be present (even as trace amounts). In its absence, the feedback should be released and the reporter would no longer mimic the CES protein behavior (and would accumulate normally) (Fig. 1.5). At the same time, care should be taken to keep the number of CES-related 5'UTR unchanged in situations where reporter gene synthesis has to be compared. Complete deletion of a CES subunit could result in different OTAFs' titration and lead to an artificial increase in the reporter gene expression. To avoid that, strains with truncated, unstable versions of the CES-subunits were used. This way, the transcript level of the gene of interest and thus its 5' UTR stays unaffected while its active product is still missing. This strategy led to the conclusion that CES subunits of PSI, PSII and *cyt.b₆f* appeared not to be trans-activated by their partners but rather auto-regulated in an assembly-dependent manner^{204,205,208}.

Nevertheless, the exact physical mechanism of this regulation is still not completely uncovered. Only in the example of cytochrome *f* was it shown that the stroma-exposed, C-terminal domain (15 aa) of the protein is indispensable for the regulation to occur^{208,210}. Further characterization demonstrated that this "repressor domain" contains key residues that control the auto-regulatory properties of cytochrome *f* and allows its interaction with trans-acting factors^{209,211} (see below).

Special case of ATPsynthase

ATP synthase manifests a particular case of the CES process. The soluble CF₁ (catalytic part of the enzyme) is constituted of subunits α , β , γ , δ and ϵ (assembling in a 3:3:1:1:1 ratio). Similarly to both photosystems, its assembly is sequential and regulated by a CES cascade: β subunit expression is controlled by the presence of nucleus-encoded subunit γ , and subsequently is obligatory for the translation of the subunit α ²⁰⁶. At the same time, because of the uneven ratio in the CF₁ between subunits α , β and all nucleus-encoded ones (3:1), a more sophisticated regulation has been developed to ensure correct synthesis of the chloroplast encoded α and β subunits.

First, a reporter gene regulated by *atpB* 5'UTR is weakly translated in the absence of γ subunit (epistatic subunit). At the same time however this repression is lacking in the combined absence of subunits γ and α which points that not only β itself but also α is required for the (auto)inhibition. Indeed, α_3/β_3 or α/β intermediates accumulate in the absence of γ . Knowing the discrepancy of α and β ratio to the nucleus-encoded subunits (3:1), it is reasonable that α_3/β_3 intermediates would regulate β translation in the absence of γ . This way, the sequestration of γ by the α_3/β_3 could ensure the correct stoichiometry between chloroplast and nuclear subunits.

Additionally, regulation of α subunit is also unique: a 5'UTR *atpA*-driven reporter gene stays weakly synthesized when neither of the subunits (α itself and β) is accumulating. This is different from all above-mentioned cases of assembly-dependent autoinhibition, and rather points to a trans-activation hypothesis where α synthesis is trans-activated by its partner - β . This additional control probably prevents the accumulation of the self-aggregation prone α subunit²¹².

CES as a regulatory process

A question arises: what is the physiological purpose of this CES process as it was observed almost exclusively in deletion mutants, which in natural conditions would be counterselected and would rapidly disappear? Some indications came from the observations that in some cases, when a CES-subunit was not present (so there could be no repression), the CES-5'UTR-driven reporter gene's accumulation was observed to be higher than in WT^{204,205}. In *cyt f* mutants lacking the C-terminal repressor domain for example (see below), *cyt f* synthesis could reach up to 300% of the WT level^{209,210}. This means that in normal, controlled situation, *cyt f* expression is not maximal, but rather actively limited to match the expression levels of other subunits of the complex. This would apply to all CES-implicated units of the given complex and would represent a mean for optimization of the energy spent for the synthesis of very abundant proteins. Indeed, the step-wise assembly of PSI and PSII consists of loops of feedback regulation that prevent excessive translation of CES-subunits when previous assembly steps have not been finished^{204,205}. On the other hand, upstream, epistatic subunits, a priori do not undergo any active control however, as mentioned above, their lifetimes are drastically reduced when unassembled. CES-subunits, as was shown for *cyt f* in the absence of subunit IV¹⁹⁸ and D1 and CP47 from PSII¹⁹⁷, accumulate to very low levels without their partners (due to translation repression) but remain stable as "auto-repressors". It is then possible that during the biogenesis of complexes, they could interact with their upstream partners to protect them from proteolytic degradation and/or allow for the re-initiation of their own translation.

CES in higher plants

After its discovery in *Chlamydomonas*, the question arose if CES is conserved in higher plants. The first convincing example of a CES process in higher plants was through the observation of the regulation of Rubisco large subunit's synthesis. It was demonstrated that, in RNAi-generated SSU knock-down lines of tobacco, LSU synthesis was decreased while the level of its mRNA stayed unaffected (yet less associated with polysomes) and that the mechanism involved an autoregulation^{213 214}. The same

observation was later repeated in maize²¹⁵. Some other hints pointing towards the occurrence of CES is suggested by the behavior of the viridis115 mutant of *Hordeum*. In this line, a defect in D1 translation was accompanied by a parallel inhibition of CP47 expression²¹⁶. Similarly, in the Arabidopsis hcf107 mutant affected in a TPR required for the *psbH* mRNA stabilization, the PsbH absence results subsequently in reduced synthesis of CP47 suggesting a CES interaction between those subunits⁴⁶. Despite accumulating evidence, stronger proofs are needed to confirm the CES-behavior of PSII, cyt. *f* and other complexes in higher plants.

CES and anterograde regulation

CES has not been observed in cyanobacteria. E.g. mutations preventing accumulation of subunits (D2) from Photosystem II complex have no effect on the translation rate of other subunits¹⁸⁹. The key residue of the C-terminal regulatory loop of cyt. *f* F₃₀₇ is not conserved in cyanobacteria suggesting that CES regulation of *petA* was developed after the initial endosymbiosis²⁰⁹. Therefore, it is probable that the CES process coevolved with additional, nucleus encoded/eukaryotes specific trans-factors required for the regulation. In fact, no other CES-subunits but Rubisco LSU²¹⁷ have been shown to contain RNA-binding domains. Neither was it demonstrated that they could directly interact with RNA. With the numerous OTAFs found to be responsible for chloroplast gene expression it is tempting to also propose their involvement in the CES process. In particular, the T factors – gene-specific proteins responsible for translation initiation, are likely candidates to be co-effectors of the regulation. Indeed, in the CES regulation of *Chlamydomonas* *petA* gene the participation of both its M and T factors has been proposed. According to the model, in normal situation, MCA1 (M-factor) and TCA1 (T-factor) would form a tertiary complex capable of interacting with *petA* mRNA ensuring its translation. Newly synthesized cyt. *f* would be, however quickly sequestered into *b₆f* complex preventing interactions with MCA1 and leaving it available for translation. Noteworthy, unassembled cytochrome *f* with “inactive” repressor domain harbor a high level of MCA1, suggesting that the cyt. *f*-MCA1 interaction is prevented. This would make MCA1 constantly available for cytochrome *f* translation, resulting in overexpression (up to 300%). In the absence of its partners however, long-lived un-assembled cyt. *f* would bind (directly or not) free MCA1 molecules and target them to degradation. In this situation *petA* translation would solely depend on the newly imported MCA1 (not on the reusable MCA1), a process which is ten times less efficient, and this way limit *petA* expression²¹¹.

Additional remarks

We lack evidence for both the exact mechanism of this interaction and other examples to conclude whether M and T factors participation is a general feature of the CES regulation. It is evident however, that CES represents one of the anterograde pathways for chloroplast gene expression. It most probably evolved following the endosymbiosis event and nowadays harbors nucleus-encoded factors. In this work, I will focus on the example of Rubisco to demonstrate that expression of the nuclear subunit is a rate-limiting step (controlled step) for the expression of the chloroplast-encoded Rubisco large subunit. In this light, appearance of CES-regulation a posteriori to the endosymbiosis would represent the need for the regulatory mechanism to control the expression of two, now spatially separated genes.

Additionally, CES mechanism plays a major role in the regulation of expression of mitochondria respiratory complexes: cytochrome oxidase (COX1), cytochrome b (CoB) and ATPsynthase. E.g. COX1 is a multimeric complex containing subunits of nuclear and mitochondrial origin. It undergoes a step-wise assembly process where each step is regulated by the inhibitory loop on COX1 genes and engaging the assembly intermediates interacting with 5'UTR of COX1 transcript, situation that mirrors the CES cascades seen in the chloroplast (see²¹⁸ and references within). Most notably, this process is coordinated by the nuclear chaperones and trans-factors (especially: Mss51 – an unique protein that is responsible for COX1 mRNA and protein intermediates stabilization²¹⁹, Pet309 – a PPR protein²²⁰ and Ssc1 – an Hsp70 chaperone²²¹). Those strongly suggest that chloroplast translation regulation should also represent interplay between OTAFs-assisted CES and nucleus encoded assembly machinery (chaperones) (see Rubisco folding).

Table 1.3: CES regulation in *Chlamydomonas reinhardtii*. (list may not be complete)

| Complex | Subunit | CES | Regulation | Reference |
|--------------|------------------|-------------------|-----------------------------|-------------|
| PSII | D2 | Epistatic subunit | Autoinhibition | 202,203,205 |
| | D1 | CES-subunits | | |
| | CP47 | | | |
| Cyt b6f | Subunit IV | Epistatic subunit | | 208,209 |
| | cyt. f | CES-subunit | Autoinhibition | |
| PSI | PsaB | Epistatic subunit | | 199,204 |
| | PsaA | CES-subunits | Autoinhibition | |
| | PsaC | | | |
| ATP synthase | Subunit γ | Epistatic subunit | | 206,222 |
| | Subunit β | CES-subunits | Autoinhibition | |
| | Subunit α | | Trans-activation by β | |
| Rubisco | SSU | Epistatic subunit | | 207 |
| | LSU | CES-subunit | Autoinhibition | |

2. Rubisco

In this last introductory part I would like to focus on the model protein I have been working on during my thesis. D-ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is one of the main photosynthetic complexes that catalyzes the reaction of carbon dioxide reduction into organic carbon. It played a major role in the regulation of the atmosphere composition, from the time of its appearance in the Archean era and forth on. During its long history different classes of Rubisco have evolved and now are present in most autotrophic organisms (phototrophic and chemoautotrophic alike). This heritage made Rubisco an object of researchers' interest since its discovery as Fraction I protein seventy years ago²²³

2.1. Rubisco evolution and clades (forms)

Rubiscos and Rubisco-like proteins (RLP) originated from Rubisco-like precursor of anoxic, methanogenic archaea. Rock sediments suggest that they diverged around 3.8 Gya to give rise to the oldest, form III Rubisco found today in archaea and form IV (RLP) of anoxic bacterial lineage²²⁴. During the Mid-Archean non-oxygenic photosynthesis correlates with the apparition of form II, diversified from form IV Rubisco of anoxic photo-bacteria. The first water-oxygenation reactions took place circa 2.9 Gya which suggest, that form I Rubisco (found in oxygenic phototrophs), that made it happen had to evolve from form III before, probably in semi-anaerobic sulphate-rich waters around 3 Gya²²⁴. Nowadays, Rubisco forms three different bona fide clades (Forms I, II and III), which can be distinguished based on the differences in LSU amino acid sequence (Form IV is a Rubisco-like protein clade. RLPs have no carboxylation activity, yet they share structural features and bind RuBP as a substrate). The common feature of all those forms is the catalytic unit which consists of an antiparallel dimer of large subunits (which nevertheless can be structurally different between the forms). The most widely distributed is the "youngest" form I that is found in green lineage phototrophs and a wide range of prokaryotes (Table 2.1). Large subunit sequence is highly conserved within the clade, with approx. 80% aa sequence similarity. However, based on the existing sequence differences, the clade was further classified into four subclasses: IA, IB – green type Rubiscos found in cyanobacteria, green algae and plants and forms IC and ID - red-type enzymes of non-green algae and phototrophic proteobacteria. Rubisco form II, first found in the purple bacteria *Rhodospirillum rubrum*, consists of only a dimer of LSU of weak sequence similarity with form I LSU and represents the simplest Rubisco clade (see Table 2.1) distributed in proteobacteria and Dinoflagellates. It differs structurally from form I by the presence of an additional α -helix at LSU N-terminus.

Frequently, it can coexist with form I and in those instances has rather a regulatory potential in mixotrophic-growth conditions (form I serves as the main carboxylase in carbon-limited conditions)²²⁵. Rubisco form II has lower specificity for CO₂ than form I, however usually compensates by having higher kinetic rate. The ancient form III is found most often in the extremophilic anaerobionts where it acts as a dimer or sometimes as an oligomer (up to 5 dimers) of the dimer unit. Its main role is no longer carbon fixation but rather removal of ribulose-bisphosphate (RuBP), the usual substrate for Rubisco form I and II, but in this case a byproduct of a purine/pyrimidine metabolism that has virtually no other scavenging pathway than to be utilized by Rubisco²²⁶. Recent study on the Rubisco structure of *Methanococcoides burtonii* proposes a new subgroup within this clade -form IIIB that is characterized by the presence of a 29 aa insertion structurally close to the C-terminal domain that may act as a SSU mimic to stabilize the LSU core²²⁷. Form IV is a diverse group of Rubisco-like proteins (RLP) that are distinctive from other classes as they do not have *bona fide* Rubisco activity but share a primary and tertiary structure with them and could be traced back to the same common ancestor²²⁸. RLPs are widely distributed in bacteria but with only a single example in alga and archaea (in addition to a regular type I and type III Rubisco respectively). They catalyze different, frequently non-redundant reactions like: methionine salvage pathway, thiosulphate oxidation etc., however, no precise mechanism of action has been elucidated so far.

Table 2.1: Rubisco types and distribution

| Rubisco Type | Composition | Activity | Distribution |
|--------------|-----------------------------------|--------------|---|
| IA | LSU ₈ SSU ₈ | + | α, β, γ -Proteobacteria, |
| IB | | + | Plants, green algae, cyanobacteria |
| IC | | + | Chlorobacteria |
| ID | | + | <i>Stramenopiles, Rhodophyta, Haptophyceae</i> |
| II | L ₂ /L _n | + | α, β, γ -Proteobacteria, Dinoflagellates |
| III | L ₂ /L _{(2)n} | +/- | Archaea |
| IV (RLPs) | L ₂ /? | - (see text) | α, β, γ -Proteobacteria, Chlorobacteria, Archaea, Algae |

2.2. Subunits and structure (Rubisco form I)

Rubisco (Fig. 2.1) consists of two distinct subunits: a Large subunit (LSU) of 50-53 kDa encoded by *rbcL* gene (or *cbbL* gene in proteobacteria), and an additional, small subunit (SSU) of around 16 kDa encoded by an *RBCS* gene family (or *cbbS* in proteobacteria) (see also Rubisco forms). In green-line eukaryotes, *rbcL* is a chloroplastic gene while *RBCS* is found in the nuclear genome. In (proteo- and cyano-)

bacteria and in non-green algae, both subunits are co-transcribed as a single operon unit. Large subunit consists of two domains: a small 151 aa N-terminal domain build up from β -sheets with α -helices on one side of the fold, and a larger 324 aa (Chlamydomonas) C-terminal domain of a β/α -barrel structure consisting of eight $\beta\alpha$ -units connected with loops that fold into triose-phosphate isomerase (TIM)-barrel domain with a flexible C-terminus of around 15 aa (13 aa in Chlamydomonas).

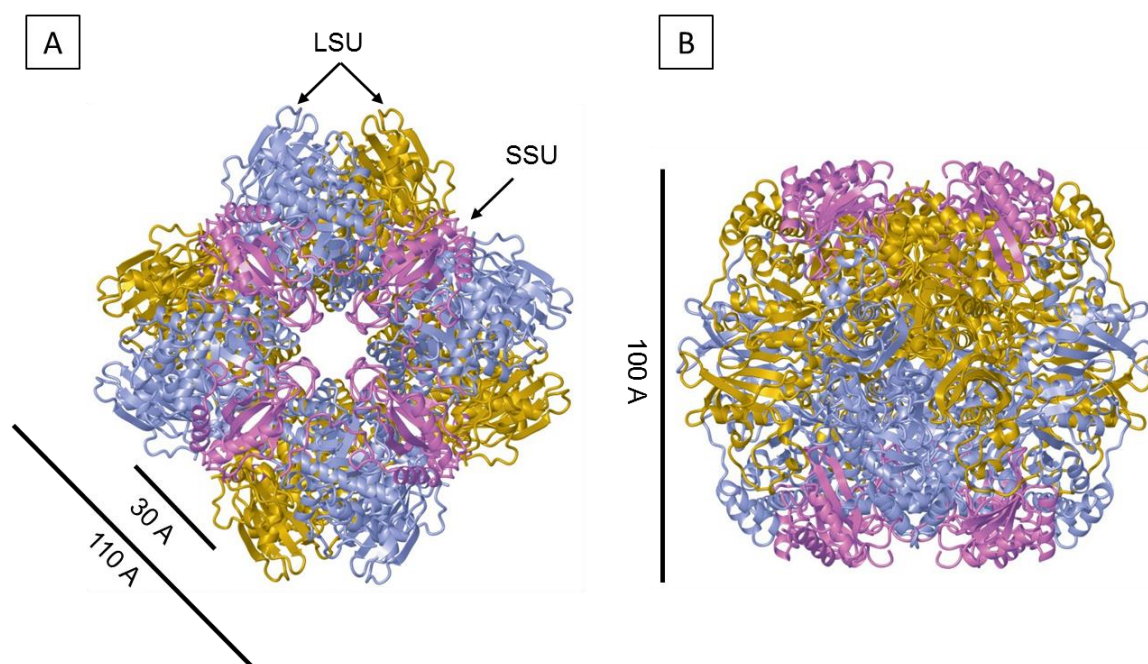


Fig. 2.1: Form I Rubisco structure

Spinach rubisco (PDB: 8RUC) structure. A - top view of the holoenzyme; B - side view. Large subunits in yellow and blue (in each dimer), small subunits in purple; Bars indicate the size of the enzyme and the middle channel. Note the positions of SSUs and protrusions into the channel.

In contrast to LSU present in Rubisco from all clades, the more recently evolved small subunits are found only in Type I Rubiscos. They share much less homology between species (approx. 35% aa sequence similarity) and are not essential for the enzymes activity per se, as the octamer LSU_8 core alone is catalytically active, but improve enzyme's stabilization^{229,230}. Nonetheless, they influence the catalytic activity of the enzyme as the chimeric variants of the enzyme containing heterogeneous SSU were shown to be able to improve Rubisco turnover²³¹. Finally, SSU also define holoenzymes cellular localization, as in Chlamydomonas they are responsible for targeting Rubisco to (and by that leading to the formation of) the pyrenoid²³². On a structural level, small

subunit displays much higher diversity than LSU. The basic scaffold of the protomer is formed by a 4-stranded β -sheet with two helices covering it on one side. Its most variable parts are the loop connecting the first (A) and second (B) β -strands and the very C-terminus of the protein. The length of the loop differs between organisms. In bacteria and non-green algae it consists only of ten residues, whereas in higher plants it has twenty-two aa, and in green algae it extends to twenty eight aa. The shorter loop is compensated by the elongation of the carboxy-terminus that forms a β -hairpin allowing proper positioning to the place where the extended loop of higher plants and green algae is⁹. A longer C-terminus is also present in green algae however it is un-structured and has no attributed function (Fig. 2.2).

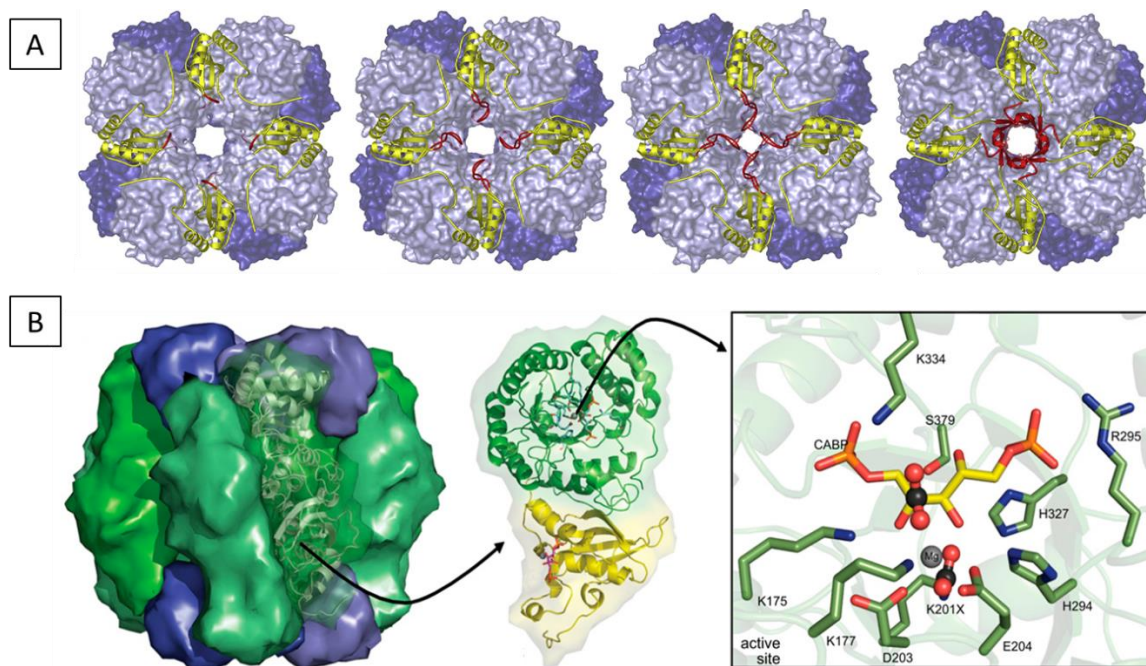


Fig. 2.2: Form I Rubisco structure

A - Variation in the small subunit structures. From left to right: *Synechococcus* PCC6301 PDB: 1RBL; Spinach PDB: 8RUC; *Chlamydomonas* PDB: 1GK8; *Galdieria partita* PDB: 1BWV. Large subunits in blue, small subunits in yellow, red part represents the variable A-B loop.

B - Active center of Rubisco: From left to right: Complete holoenzyme - active centers are placed on around 2/3 of the height on both sides of each LSU dimer (green); (Middle) Zoom on a single LSU, in green C-terminal domain, in yellow N-terminal domain. Residues contributing to the active sites are highlighted. Each subunit takes part in the formation of 2 active sites that are complemented by the other subunit from the dimer; (Right) zoom on one of the active centers of the enzyme. Ten residues of the C-terminal domain that form the core of the site are marked (2 residues from the N-terminus of the second LSU are not shown). Reaction site shown in its "active" state with carbamylated Lys201 coordinating magnesium ion and the C6 intermediate mimic 2-carboxyarabinitol-1,5-bisphosphate (CABP) with the substrate CO₂ molecule in the active site. Modified from⁹

The catalytic unit of the enzyme is a large subunit head to tail dimer. Depending on the Rubisco form, it can reach different oligomerization states (see below). In the form I, Rubisco dimers form a tetrameric “ring” of approximately 100 Å in height and 110 Å in diameter, with a vertical channel of 30 Å between the dimers²³³. Four small subunits stabilize the cylindrical structure on top and at the bottom of the holoenzyme creating a barrel-shaped hexadecamer (Fig. 2.1 and 2.2). Although this kind of symmetrical organization is not unique to Rubisco, in its case, it is characterized by the extreme rigidity of the structure with each active center acting independently. Two active sites are located at the interface of the N-terminal domain of one subunit and of the TIM-barrel domain of the second antiparallel subunit of the dimer. They are positioned on the surface of each dimer so that the entry of the active site is opened to the outside of the enzyme. Each one is constituted mainly from residues from the C-terminal β/α -barrel. Ten polar and charged residues (K_{175} , K_{177} , K_{201} , D_{203} , E_{204} , H_{294} , R_{295} , H_{327} , K_{334} , S_{379} (taken from the *Chlamydomonas* structure)) from the loops connecting the first, second, and fifth to eighth β -strand with respective α -helices are interacting with the substrate/product and/or magnesium ion directly. Two conserved residues (E_{60} and N_{123}) from N-terminal loops of the second LSU of the dimer contribute to the center as well. Loop 6 (connecting the sixth β -strand with the sixth α -helix) regulates the catalysis and specificity²³⁴. Changing its conformation (by the movement of residues 331-338) yields the open and closed states of the active center. In the closed state, substrates are protected from solvents, a situation that is necessary for the catalysis to occur.

2.3. Rubisco reaction

As expected, in the majority of cases Rubisco catalyzes the reduction of atmospheric carbon dioxide into biologically available organic carbon through carboxylation of D-ribulose-1,5-bisphosphate (RuBP). The reaction is a two-step process. First, activation of the reaction center takes place by carbamylation of uncharged group of Lys_{201} (as seen in *Chlamydomonas* and Spinach structure) using a non-substrate CO_2 molecule. The resulting carbamate traps a magnesium ion to one of its carbonyl oxygens that, finally is coordinated to 3 oxygens of aminoacids 201-204²³⁵. The activated enzyme can bind its first substrate, RuBP (and one molecule of water) to the proximity of the carbamate that surrounds the magnesium ion, which assumes its final octahedral coordination (to 3 aa, 2 times to RuBP and to one water molecule). The first actual reaction step is binding of RuBP $C_{3's}$ hydrogen to Lys_{201} and the subsequent negative charge equilibrium shift towards C_2 by the proximity of carbamate oxygen (tautomerisation of RuBP to an enediolate). Next, CO_2 (or O_2) will be fixed in the active site, thereby replacing water. There it gets polarized and mobilized to perform an

electrophilic attack on the enediolate molecule. Displacement of hydrogen atoms within the molecule allows the formation of the covalent bond between RuBP and CO₂ resulting in a six-carbon sugar. In the next step, C₃ of the sugar is hydroxylated (with another molecule of water) with the aid of His₃₂₇ and afterwards deprotonated by Lys₂₀₁. The un-optimal configuration of electrons of C₃ is equilibrated by the cleavage of the C₂-C₃ bond resulting in the formation of the first product molecule: D-3-phosphoglycerate that is liberated from the active site. The rest of the previously carboxylated RuBP (now a C₃ sugar with a spare electron on C₂) is protonated by Lys₁₇₅ which results in the formation of another D-3-phosphoglycerate (3PGA). With its liberation, the reaction cycle is finished²³⁶ (Fig. 2.3).

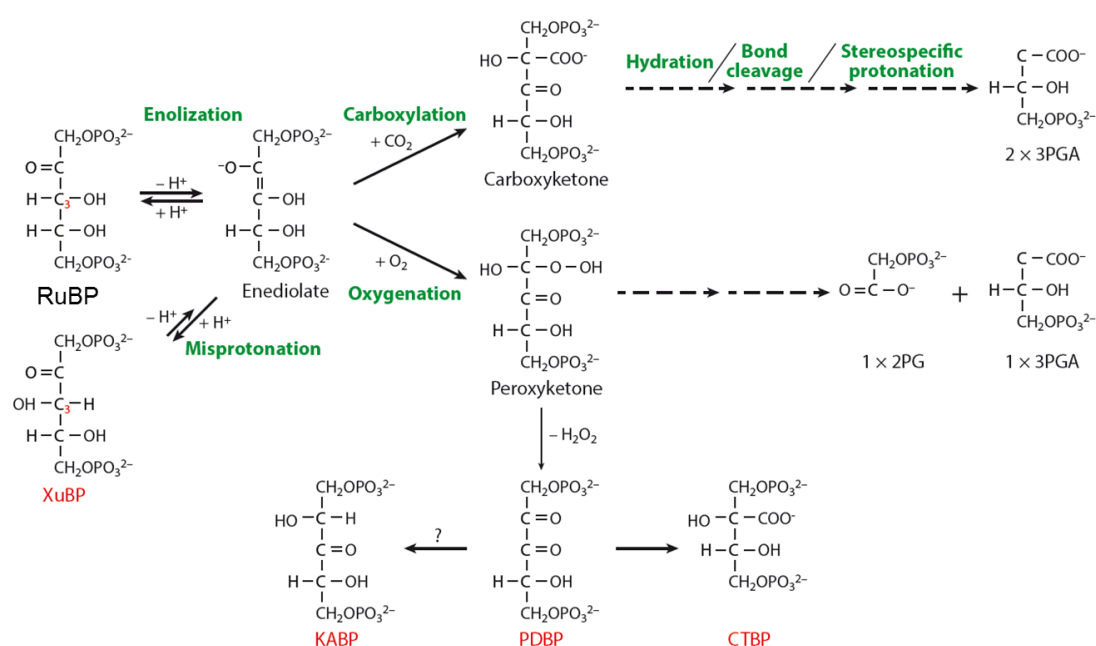


Fig. 2.3: Rubisco-catalyzed reactions

Rubisco substrate, RuBP undergoes 5-step reaction during CO₂ fixation (Enolisation, Carboxylation, Hydratation, Bond Cleavage and Protonation). Other possible mis-reactions that lead to the formation of inhibitory sugar phosphates are shown: misprotonation and Oxygenation. Names of the inhibitors in red. Modified from³³³

2.3.1. Oxygenation

Because CO₂ does not bind to Rubisco itself (to form a Michaelis complex), but directly to RuBP, and due to the fact that CO₂ and O₂ are electrochemical similar, the enzyme has difficulties to distinguish between the two gases. And although its affinity for CO₂ is higher (Form I on average: Km(CO₂)= 9 uM; Km(O₂)= 500 uM), the great discrepancy

between the amounts of the CO₂ and O₂ in the atmosphere makes oxygen an important reaction competitor. Oxygenation of RuBP is then possible. The later leads to the formation of a C₅ sugar, that when subsequently hydrolyzed liberates one molecule of D-3-phosphoglycerate and one molecule of 2-phosphoglycolate (2-PG). 2-PG is unavailable for the further sugar formation and needs to be recycled through an energy consuming photorespiratory pathway. Oxygenation can take up to 25% of all Rubisco turnovers. Because of that, photorespiration drastically diminishes the theoretical, maximal yield of CO₂ fixation as for each oxygenation, ATP needs to be spent and molecular CO₂ is released (Fig. 2.3).

2.3.2. Unproductivity

Due to similarity between substrates and products, as well as the complexity of the catalyzed reaction, Rubisco is very error prone with multiple possible inhibition states. First, its proper substrate RuBP, when bound to the uncarbamylated active centers, forms a tight complex with the enzyme. Next, enediolate - the first reaction intermediate (deprotonation of RuBP)- is very unstable and reactive. Due to the low speed of the carboxylation, a reverse mis-protonation of the enediolate is possible leading to synthesis of xylulose-1,5-bisphosphate, which stays tightly bound to the reaction center inhibiting further reactions. An oxygenation intermediate, peroxyketon, is also an unstable molecule that can turn into other inhibitory sugar phosphates (Fig 2.3). Finally, during night time, or under limiting light conditions plant Rubisco is usually inhibited by 2'-carboxy-D-arabinitol-1-phosphate (CA1P) to limit misproductivity. Reactivation and/or removal of the inhibitors from the active site necessitates conformation changes in the structure (C-terminal end pulling and subsequent movement of loop 6) induced by Rubisco activase(s) – the ATP dependent metabolic chaperon of Rubisco.

2.3.3. Efficiency

The efficiency of Rubisco is described as the real ratio between carboxylation and oxygenation. It is defined by the CO₂ to O₂ concentration ratio in the vicinity of the enzyme multiplied by the specificity factor for those gases. The specificity factor itself is described as the ratio of velocities of concurrent reactions multiplied by the Michaelis constants for CO₂ and O₂ respectively ($\Omega = (V_{CO_2}K_{O_2})/(V_{O_2}K_{CO_2})$). Each Rubisco clade and even enzymes from different organisms within the same clade have vastly different kinetic properties. Ancient Rubiscos, like archaeal Form III have very low specificity for CO₂ - most probably due to their evolution in extreme anoxic conditions. On the other hand, within the oxygenic photosynthesizing organisms, the differences can be also visible with proteobacteria having low CO₂ specificity ($\Omega=30$), green lineage mediocre

specificity ($\Omega=80$) and red algae the highest one ($\Omega=200$). This suggests positive evolution of the enzyme to the most specific one, however, an inverse correlation is observed for the turnover rate of carboxylation (k_{cat}). At the same time, differences lying in the environmental origin of the given organism influence heavily the effective turnover rate of Rubisco. Organisms that live in an environment with high CO_2/O_2 ratio (e.g. *Rhodospirillum rubrum*) can afford Rubisco with low CO_2 specificity. On the other hand, certain organisms (cyanobacteria, algae etc.) living in niches with low CO_2 availability developed carbon concentration mechanisms (CCMs) that actively change the partial gas pressures (at the expense of ATP), thereby improving environment for Rubisco to increase its efficiency. Those organisms contain Rubisco of lower affinity for CO_2 (higher K_{CO_2}) and faster carboxylation rates (higher V_{CO_2}). Conversely, e.g. C3 plants now have higher specificity for CO_2 over O_2 , better affinity for CO_2 but lower carboxylation rates. All in all however, from the available experimental data, carboxylation rates of all Rubiscos have been measured to be in the range of 3-5 CO_2 molecules fixed per active center every second, leading to the conclusion that Rubisco in general is a very slow enzyme. Its inefficiency needs to be compensated by the huge amount of Rubisco synthesized in all photosynthetic organisms. Most probably, its high importance for the photosynthesis and cell survival, rigid structure and dependency on a multitude of factors for proper formation (see below) restricted its evolution and limited the number of positive mutations that could improve it. Artificial improvement of Rubisco to match the global increasing demand for food is a growing trend in the research community. It is out of the scope of this manuscript; nonetheless, multiple reviews on the subject are available²³⁷⁻

239

2.4. Rubisco biogenesis

Two of the main reasons limiting attempts to improve the activity of Rubisco for agricultural purposes are the intrinsic physical proprieties of the enzyme making it difficult to express it *in vitro* (see Folding) and a complicated and still not fully understood biosynthesis pathway leading to its formation. I will focus on the description of Type I Rubisco in eukaryotic organisms as it is of most interest (Fig 2.4).

2.4.1. Expression and regulation

Two subunits of Rubisco are encoded in two separate genomes and thus different processes are responsible for their individual expression that at the end yields Stoichiometric accumulation of the proteins.

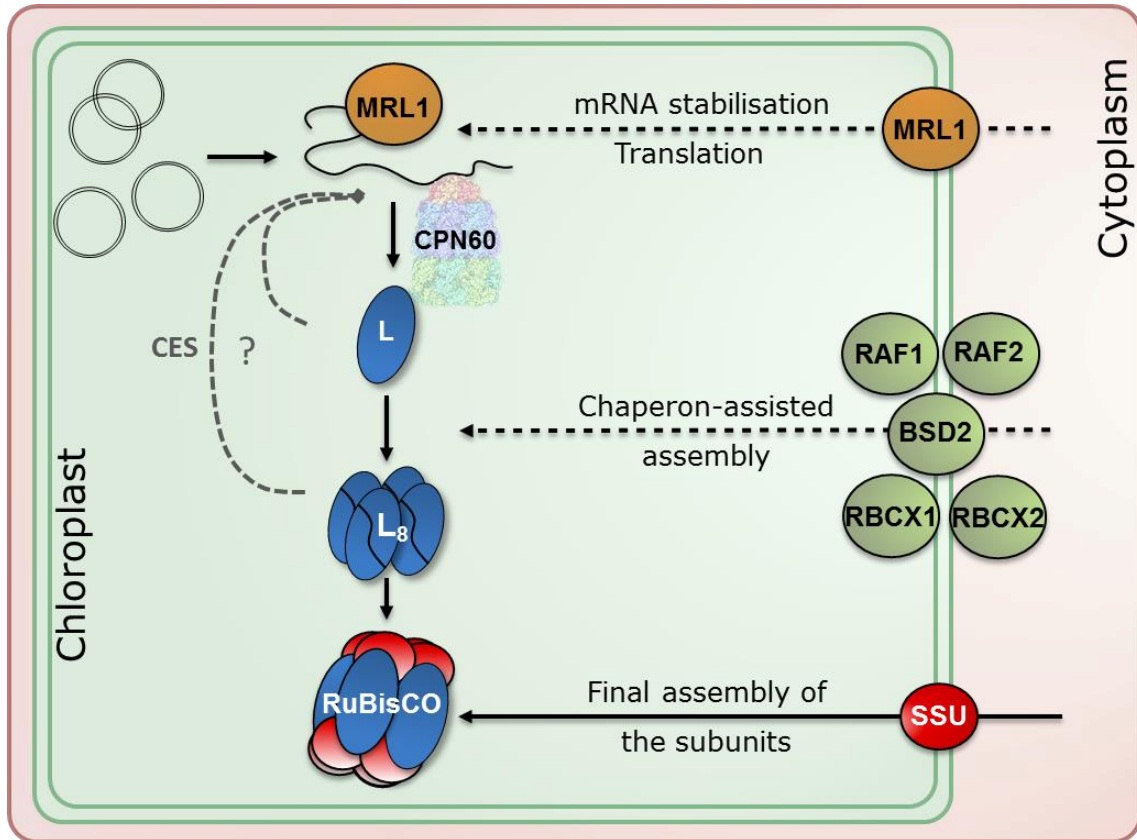


Fig. 2.4: Rubisco biogenesis

rbcL is transcribed from multiple copies of chloroplast genome. Its mRNA is stabilized by MRL1 protein. After proper translation, LSU is folded by a chaperonin system CPN60/20/10. Assembly into Rubisco LSU₈ core is assisted by multiple chaperones. At the end, mature SSU, imported from the cytoplasm, replaces chaperones to form a complete enzyme.

Small subunit is encoded by a *RBCS* gene family of 2 to 12 members (2 in *Chlamydomonas*: *RBCS1* and *RBCS2*). *rbcL*, coding for LSU is a single gene in the chloroplast genome, which due to the multiple copies of the genome can be present in the cell in ~100 copies. Already in 1980s was it suggested that *RBCS* expression is regulated mostly on transcriptional level whereas *rbcL* should be controlled post-transcriptionally. *RBCS* is a PhANG (see Retrograde signaling) gene and responds to a multitude of signals also perceived by the chloroplast. Its mRNA level in the cell depends on: developmental stage, tissue, circadian rhythm, CO₂ level, light, temperature and nutrients. It was demonstrated that *RBCS* promoter region is essential for this diversification however no evident experimental data demonstrated responsible cis- or trans- elements. Additionally, after proper synthesis in the cytosol, transit peptide-containing SSU pre-protein (preSSU) is targeted to the chloroplast. SSU transit peptide

contains multiple motifs of variable influence on chloroplast targeting, suggesting that the translocation process occurs stepwise and requires interaction with various proteins^{240,241}. It was suggested that the cytosolic chaperon machinery (e.g. Hsp70) is assisting in this process but there are no experimental evidence supporting this hypothesis (in contrast, Hsp70 amount did not affect preSSU transport *in vitro*²⁴²). On the other hand it is clear that preSSU is interacting with Tic/Toc translocon complex most probably entering the chloroplast through it (²⁴³ and references within). preSSU is subsequently cleaved of by stromal processing peptidase and N-terminal methionine is (mono)methylated by Rubisco methyltransferase giving rise to the mature SSU peptide^{244,245} (see below: SSU Folding).

Biosynthesis of the large subunit is much more documented. First, as mentioned in part II, regulation of *rbcL* expression is mostly post-transcriptional but in higher plants, during leaf development level, its mRNA amount changes significantly due to the modulated activity of PEP (via the action of sigma factors 1 and 6). It is the highest at the early stage of proplastid to chloroplast transition when the photosynthetic apparatus is being created^{246,247}. In both plants and algae, *rbcL* transcript is one of the most abundant mRNA in the chloroplast (in *Chlamydomonas* ~3000 copies in standard conditions). In higher plants the 180 nt long *rbcL* primary transcript undergoes processing yielding a shorter 60 nt transcript (the length of these transcripts varies between species, a large insertion is found in the primary mRNA in grasses²⁴⁸). Both forms are stable and accumulate (see consequences below). In *Chlamydomonas*, *rbcL* is one of few (others being *petA* and *atpH*) transcript that does not undergo maturation and its primary transcript has a 92 nt long 5'UTR (with a triphosphorylated 5' end¹). A first study indicated that the promoter region of *rbcL* gene ranges from -18 to +63 in regard to transcription site however, it does not confer for full transcription efficiency to a heterologous reporter gene. A canonical promoter sequence (TATAATAT: TATA-box) lies 12 nt upstream of transcription start site. Maximal transcription efficiency is only obtained when a remote enhancer element lying within the coding sequence (somewhere between +126 and +170 from translation initiation site (which is +92 from transcription start)) is present²⁴⁹. Another study highlights that the beginning of the LSU coding sequence (+83 from translation start) increases heterologous protein production²⁵⁰. Later, it was confirmed that *rbcL* 5'UTR does not contain a promoter region as changes in its sequence or its distance to the start codon did not disturb transcription²⁵¹.

rbcL 5'UTR is on the other hand, essential for the transcript stabilization¹⁰. In its native form, the 92 nucleotides *rbcL*-5' UTR would fold into two stem-loop structures, as

determined by alkylation with dimethyl sulfate¹⁰. These stem-loops do not prevent transcript degradation by themselves (by acting as a barrier against exonucleases)²⁵¹, but would provide the necessary local structural arrangement for a binding site located between these two stems to be recognized by a trans-factor crucial for the stabilization¹². Indeed, insertional mutagenesis and screen for Rubisco deficiency has led to the identification in *Chlamydomonas* of MRL1, a nucleus-encoded protein that is required for *rbcL* transcript stabilization¹, MRL1 is a soluble, PPR protein (11 PPR motifs) of 150 kDa found in chloroplast stroma in high molecular weight complexes. It has no other known domains. While MRL1 knock-outs completely abolish *rbcL* mRNA and LSU protein accumulation, knock-down mutants showed lower transcript accumulation which correlated with lower LSU protein amount (see also Results)²⁵². Swapping *rbcL* 5'UTR rescues the mutant phenotype indicating that MRL1 binds to this regulatory sequence. Whether MRL1 truly binds the stability determinant identified by Salvador's group^{12,251} is unknown and questionable as two other studies^{6,8} pinpointed the existence of different footprints at *rbcL* very 5' end and at the second stem-loop downstream of the stability element. Nevertheless, MRL1's function is in line with the proposed role of helical-repeat proteins in organelle transcripts' stabilization^{211,222,253} making MRL1 a M-factor of *rbcL*¹ (see above). In the same study the authors demonstrate that in *Arabidopsis*, MRL1 is required for the accumulation of *rbcL* mRNA processed form. MRL1 thus most probably plays a similar function: it would bind to the 5'UTR of *rbcL*, maybe direct cleavage by an endoribonuclease, and protect the generated 5' termini from exonucleolytic degradation. At the same time, the integrity of *rbcL* processed form is not obligatory for the *rbcL* pre-mRNA to accumulate and is not required for phototrophy which indicates that at least the primary transcript can be stabilized/engaged by other factors making it available for translation¹. In line with this observation is the fact that both forms are polysome-associated²⁵⁴.

rbcL translation was proposed to take place not only on free- but also (mainly) on membrane-bound ribosomes in barley²⁵⁵. However the extent of this membrane associated translation seems to be contradicted by the ribosome-profiling analysis of soluble versus membrane-bound ribosomes⁸². Furthermore, this observation is somewhat in contrast with the more recent FISH experiments in *Chlamydomonas* showing that high amounts of *rbcL* mRNA localize to the basal region of the chloroplast, even though weaker and dispersed *rbcL* mRNA signal is observed throughout the chloroplast as well and proposed to come from membrane bound polysomes⁸⁵. Transcript localization to this region is hypothesized to facilitate co-translational localization of LSU to the pyrenoid. The translation process itself could be regulated at

the initiation step and/or during the elongation. In barley (isolated) plastids, the *rbcL* transcript was found to be associated with polysomes in the dark and its translation respectively arrested or promoted during dark/light transitions^{256,257}. This led authors to propose that *rbcL* translation elongation would be part of the general process of photosynthetic genes' control by the elongation factors (e.g. *tufA*) whose expression also relies on light stimuli²⁵⁸ but experimental evidence for an *rbcL* specific control is still lacking.

More details about *rbcL* regulation of translation were provided with the analysis of SSU knock down/out mutants in *Chlamydomonas* and tobacco. First, SSU depletion resulted in parallel diminishment of LSU accumulation which was the first evidence for assembly-dependent control of Rubisco and coordination of separated subunits^{207,259}. In both cases *rbcL* transcript level was not affected by the absence of SSU, but Rodermel et al. observed that (in tobacco) *rbcL* mRNA polysome loading was shifted towards lighter fractions²¹³. In parallel, radioactive pulse chase experiments demonstrated that LSU synthesis is either inhibited on a translation initiation step or rapidly degraded in the absence of SSU (as is SSU in the absence of LSU)²¹³. SSU could act as a positive regulator for *rbcL* promoting its translation in WT conditions. This activation would then be released in the absence of SSU. Insight into the mechanism involved in this reduced LSU synthesis came with the polysome analysis of a double mutant combining truncated LSU and *RBCS* silencing (siSSU). In the truncated *rbcL* context, *rbcL* mRNA is still accumulating however, the mutated LSU cannot fold properly and is degraded²¹⁴, and thereby is unavailable for any regulation (see CES mechanism). In the double mutant, as expected accumulation of both SSU and LSU are strongly diminished. However, as compared to simple siSSU mutant, *rbcL* transcript showed an increased association to polysomes (as seen in WT). This made authors suggest that, when lacking its assembly partner LSU undergoes a CES process and is responsible for its own translation inhibition – a situation observed in *Chlamydomonas* for other photosynthetic subunits²¹⁴. This was further supported by the experiment in which LSU accumulation was diminished in tobacco plants as a result of virus induced gene silencing of its BSD2 chaperone (see Folding), *rbcL* transcript association to polysomes was even higher than in the control line with almost no free *rbcL* transcript suggesting that loss of the negative regulation (due to instability of LSU in the absence of BSD2) increases translation rate of LSU to levels higher than in WT. A similar effect had already been observed for CES-insensitive cytochrome *f* mutants in *Chlamydomonas*²¹¹ and points to the fact that in normal conditions LSU synthesis may not be maximal. Indeed, in the line with similarly diminished levels of both BSD2 and SSU, the polysome pattern was drastically different

with additional *rbcL* mRNA being found with monosomes. This suggest that contrary to BSD2-depleted lines, in the SSU-limiting situation, some free LSU (resulting from incomplete BSD2 silencing) is available for negative regulation (which leads to free or monosomal associated *rbcL* mRNA). Together those results suggest that LSU synthesis is not maximal in WT, and indicates the presence of some unassembled LSU able to repress its translation to fit SSU levels. More generally, it indicates that SSU is a limiting factor for LSU synthesis however, mechanistically, the regulation itself is SSU independent²¹⁴. This is compatible with previous work showing that *R. rubrum* Rubisco type II synthesis is diminished during an oxidative stress: the translational regulation, (which might be different from CES), cannot be SSU dependent as Type II Rubisco consist only of LSUs. Oxidative stress-triggered inhibition of LSU synthesis was also observed in tobacco and *Chlamydomonas*²⁶⁰⁻²⁶². In the latter case, authors were able to demonstrate that, LSU can interact with RNA via its N-terminal domain (see Structure) which contains a RNA binding domain. This RRM domain would be hidden in the enzyme structure and would be exposed in the stress conditions²¹⁷. RNA binding of LSU was however unspecific which may suggest that whichever the mechanism, most probably additional trans-factors are required for its specificity.

2.4.2. Folding and assembly

Studies on the organelle chaperone system have revealed that chloroplasts and mitochondria have complex systems of folding, quality control and stress response of bacterial and eukaryotic origin^{263,264}. In addition, new specific chaperones are being regularly discovered. Rubisco expression from eukaryotic organisms has been an intense research field since many years. However, the difficulty to properly and efficiently express it *in vitro* has been hampering those efforts. Nowadays, it is known that a multitude of auxiliary proteins are necessary for proper Rubisco biogenesis. It seems logical as Rubisco is an enzyme of dual origin (Type I Rubisco in eukaryotes) which due to spatial separation of its both subunits most probably requires temporal stabilization of oligomerization intermediates and/or subunits alone to yield proper assembly.

First, as classical TIM-barrel domain proteins with patches of highly hydrophobic surfaces, newly synthesized LSU is highly aggregation prone and cannot fold properly without assistance. It was earlier demonstrated *in vitro* that GroEL/ES-type chaperonin system is essential for the proper folding of *R. rubrum* Rubisco (type II)²⁶⁵. Chaperonins are ATP-dependent chaperones that encage the non-native polypeptide thereby providing the required isolated environment for their proper folding. Chaperonins - GroEL/ES in bacteria and CPN60/20/10 in the chloroplast are composed of two

heptameric rings of GroEL or CPN60 subunits organized in a tail to tail manner. The double ring is closed on both sides by two heptameric “caps” of GroES or CPN20-CPN10 heterodimers. Through its exposed hydrophobic residues, the open ring can bind a multitude of neo-nascent proteins that are then trapped inside the ring cavity when enclosed by the GroES ring. Upon hydrolysis of seven ATP molecules, the GroEL ring undergoes drastic conformational changes that help the trapped protein to fold properly. Binding of GroES and ATP to the opposite (trans-) ring of the complex triggers dissociation of the GroES cap from the cis-GroEL heptamer and the release of the folded protein (reviewed in ²⁶⁶).

CPN60 complex

Cpn60 isoforms from *Chlamydomonas* are homologous to GroEL; Cpn60 α together with Cpn60 β were able to functionally replace *E.coli* chaperonin *in vivo*²⁶⁷ and Cpn20 substituted for bacterial GroES in Rubisco refolding²⁶⁸ proving that both complexes play a homologous role and most probably share the same mechanism of action. However, an additional feature of plastid chaperonin complex is the diversity of subunits that it is consisted of with at least two isoforms of CPN60: α and β , which can be encoded by multiple genes (e.g. three isoforms in *Chlamydomonas*: CPN60A, CPN60B1 and 2). The same diversity can be observed for GroES homologs: in addition to CPN10 (GroES-like protein) that can exist in one or two copies in the genome, photosynthetic eukaryotes encode also for one or two CPN20s (practically a dimer of CPN10). Additionally, a CPN20 homolog - CPN23 (with a longer linker between CPN10s) was described in *C. reinhardtii*²⁶⁹. *In vivo*, CPN20 and 10 form heterodimers that mimic the hepta-fold of GroES. Different isoforms are assumed to assemble in variable stoichiometry leading to differences in folding efficiency. This diversity results most probably from an adaptation of the chaperonin to more diverse substrates range in the chloroplast.

As mentioned in the example above, both chaperonin systems interact with LSU as a substrate. GroEL is necessary for type II Rubisco formation and it was demonstrated that it is required for the *in vitro* expression of cyanobacterial type I Rubiscos^{270,271}. *In vivo*, CPN60/20/10 complex definitely interacts with Rubisco at some point of its biosynthesis as it was first discovered in pea extracts as Rubisco-bound protein²⁷². Multiple studies that followed have shown that CPN60 forms a stable intermediate with LSU in maize, pea, spinach and tobacco. It is limiting for Rubisco formation while expressed *in vitro* and, *in vivo*, mutants of CPN60 subunits showed defects in Rubisco accumulation^{254,273,274}. It is then not surprising that nowadays CNP60/20/10 complex is

unanimously accepted to be critical for LSU folding³. However it is not clear yet whether in the chloroplast, CPN60 complex interacts directly with newly synthesized LSU or if there are any chaperones acting ahead of CPN60 complex. Most notably, the DnaK-DnaJ-GrpE chaperone system, that in chloroplast is often collaborating with the chaperonin, may be the best candidate for an early interaction with LSU²⁶⁴. Indeed, expression of Rubisco subunits in an *E.coli dnaK* null mutant leads to extensive aggregation of LSU (as compared to the WT strain) that can be suppressed by co-overexpression of DnaK²⁷⁵.

BSD2

Additionally, bundle sheath defective 2 protein (BSD2) that has similarity to the zinc-finger domain of DnaJ (Hsp40) chaperone was discovered in maize²⁷⁶. Yet it lacks other domains of DnaJ protein family and thus cannot be assigned to this family. DnaJ proteins are co-chaperones of Hsp70, which will bind to the substrates and increase Hsp70's ATPase activity. BSD2 is a nucleus-encoded protein targeted to chloroplast that is necessary for Rubisco synthesis. In both maize and tobacco BSD2 mutant, LSU does not accumulate while *rbcL* transcript is accumulated and associated to polysomes^{214,276}. Along with the similarity to the DnaJ protein it suggests that BSD2 plays a role in LSU posttranslational regulation. It is further supported by the work in *Chlamydomonas* where the putative BSD2 ortholog (Znj2) co-migrates with *rbcL* transcripts on polysomes. *In vitro* assays showed that CrZnj2 also prevents protein aggregation²⁷⁷. BSD2 is therefore assumed to be a first in line LSU-specific chaperone that can interact with translated LSU polypeptide preventing its aggregation/mis-folding or targeting it to the chaperonin complex. On the other hand in the maize BSD2 mutant, CPN60-bound LSU can be observed during radioactive pulse experiments pointing, on the contrary, that BSD2 acts after the chaperonin²⁷⁴.

Rubisco assembly chaperones

LSU folding is only the first step in type I Rubisco assembly. It cannot acquire its final conformation without the additional steps of assisted-oligomerization. This is in contrast to e.g. type II LSU which *in vitro*, after leaving GroEL/ES dimerizes spontaneously to attain its final, active form²⁶⁵. Type I form is thought to assemble in at least a two-step manner. First, LSU₈ enzymatic core is formed which afterwards is stabilized by the fixation of eight SSUs. *In vitro* expression of cyanobacterial Rubiscos is possible but inefficient²⁷⁰. Moreover, *in vitro* reconstitution of *Synechococcus* sp. PCC6301 Rubisco with its subunits and chaperonin alone was shown to lead to aggregation of LSU (with SSU being either assembled in low amount in active complex

or soluble²⁷⁸). Both suggest that most probably i) collision-assembly of the subunits is inefficient, ii) they need to oligomerize sequentially and/or iii) additional factors are required to stabilize the assembly intermediates. Attempts to produce hybrid Rubisco usually result in low yield of synthesized enzyme suggesting that chaperones necessary for LSU assembly coevolve with LSU proving how important is the control over its biosynthesis^{238,279}. Type IB green Rubisco from plants and green algae have been studied most and up to now, three assembly chaperones have been reported.

RBCX

In some cyanobacterial species (β -cyanobacteria) *rbcX* is located in between the *rbcL* and *rbcS* genes in Rubisco operon (*rbcLXS*) and thus was the prime candidate for a non-subunit factor implicated in Rubisco formation²⁷¹. Despite low sequence similarity between β -cyanobacterial species (below 60%), its position in between Rubisco subunits encoding genes in the operon suggests an important function. Indeed, RBCX increased the amount of functional *Anabaena* sp. CA and *Synechococcus* sp. PCC7002 Rubisco produced in *E. coli* system^{280,281}. In the marine *Synechococcus* PCC7002 cyanobacteria, RBCX seems to influence Rubisco formation as a translational frameshift introduced in the *rbcX* gene resulted in lower Rubisco accumulation. However, complete segregation of the mutated copies could not be reached suggesting that either the *rbcX* is essential for cell survival or that the introduced mutation has a pleiotropic effect due to its proximity to *rbcL* or *rbcS* (a possibility that was not examined in the study)²⁸¹. The latter hypothesis seems to be supported by the observation that in the closely related *Synechococcus elongatus* PCC7942, an *RbcX* deletion has no apparent effect on cell phototrophy or on Rubisco accumulation. This functional difference might also be due to the positioning of the *rbcX* gene which in *Synechococcus* PCC7942 is no longer in the *rbcLS* operon but 100kb away²⁸². Nonetheless, in the same work authors confirmed that SynPCC7942 RBCX co-expressed with Rubisco in *E. coli* improves holoenzyme's accumulation yield. At the same time, the very low levels of RBCX protein in the cells (undetectable in cell extracts) and/or its selective expression (in response to precise environmental stresses) prevent drawing conclusions about its role *in vivo*. Green algae and higher plants encode two, chloroplast targeted RBCX proteins: RBCX-I and -II (RBCX-I is closer to cyanobacterial RBCX). In *Chlamydomonas* only RBCX-II was retained and is now present in two sub-isoforms RBCX-IIa and -IIb. RBCX-I and II share low sequence similarity to their cyanobacterial counterpart and to each other (below 30%), however the structure of the RBCX domain is almost identical in all cases (except that in *Chlamydomonas* RBCX-IIb has an un-structured C-terminal extension that doubles the size of the protein). Chloroplast RBCX most probably share the same function in

cyanobacteria as *Chlamydomonas* RBCX-IIa and *Arabidopsis* RBCX-I similarly improve cyanobacterial Rubisco formation in *E.coli*^{283,284}.

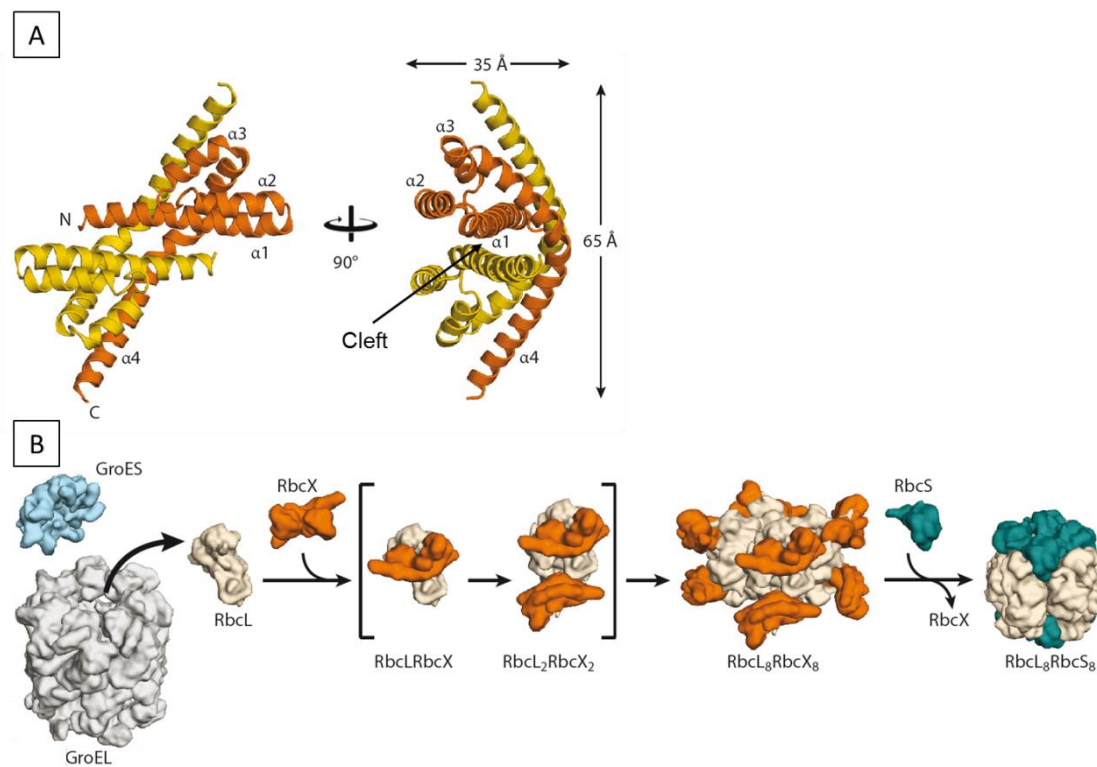


Fig. 2.5: RBCX structure and interaction with LSU

A – Structure of *Synechococcus* sp. PCC7002 RBCX dimer (PDB: 2PEN) in two orientations. Two monomers are shown in yellow and orange. Helices of one of the dimers are annotated. Note the hydrophobic cleft responsible for RBCX-LSU interaction.

B – Model of RBCX interaction with LSU. After leaving chaperonin complex (GroEL/ES) LSU monomer is sequestered by the RBCX dimer preventing its re-trapping by the folding complex. Subsequently LSU dimers are formed and are protected by two RBCX dimers. These units then oligomerize further until Rubisco LSU₈ core is formed. Binding of SSUs changes the conformation of LSU resulting in dissociation of RBCX dimers and final formation of the holoenzyme. Modified from³

RBCX is a 15 kDa protein consisting of three short and one long ($\alpha4$) α -helices organized in a bundle. Long helices are antiparallel with a 60° shoulder in the middle of their length. The active form of RBCX is a homo-dimer of an arc shape with α -helical bundles at opposite ends²⁷⁸ (Fig. 2.5). The crystal structure of *S. elongatus* PCC6301 LSU oligomer complexed with *Anabaena* sp. CA RBCX showed that an RBCX dimer interacts with LSU C-terminal, conserved sequence (EIKFE(F/Y)X). The motif is bound by the central, hydrophobic cleft of RBCX, some peripheral residues also play a role in the stabilization

of the interaction. From the structure and the *in vitro* biochemical experiments, RBCX dimer seems to stabilize the octamer of large subunits before binding of SSUs occurs. This chaperon-LSU interaction is dynamic in nature with different RBCXs having different affinities for large subunit. For example, the *Anabaena* sp. CA RBCX has a 50 fold higher affinity for *S. elongatus* PCC6301 LSU than its cognate RBCX does (a fact which was used by the authors to crystalize the LSU₈RBCX₈ intermediate) and cannot be dissociated by *S. elongatus* SSU. Moreover, in the same work, the authors generated an LSU mutant (LSU_{R212S}) arrested at the dimerization state and could observe that RBCX could form complexes with the LSU dimer stabilizing this intermediate. They propose a model where RBCX could bind to LSU as soon as it leaves GroEL/ES chaperonin thereby preventing its rebinding. RBCX binding would promote LSU dimer formation and further oligomerization by preventing disassembly until LSU₈RBCX₈ is finally replaced by eight SSUs²⁸⁵.

RAF1

Interestingly, another assembly chaperon performing a similar function has been identified in maize. Coined Rubisco Accumulation Factor 1 (RAF1), it is obligatory for Rubisco formation in maize as a RAF1 knock out results in very low LSU accumulation (below 2%) and seedling lethality. *rbcL* and *RBCS* transcripts level and translation were not affected in the mutant, which points to its post-translational role. Radioactive pulse experiment of maize leaves showed retention of the LSU in a high molecular weight complex (HMWC) that authors attribute to CPN60-bound LSU. In-planta cross-linking experiments demonstrated that it can form heavy complexes (approx. 720 kDa) with LSU of unknown stoichiometry²⁷⁴. All these observations point to RAF1 being an LSU chaperone. Further analysis on the crystal structure obtained subsequently, allowed shedding more light on the mechanism of RAF1 interactions. First, RAF1, similarly to RBCX, acts as a dimer of ~45 kDa subunits, that each binds one dimer of LSU in the LSU₈ core. Each protomer is build up by an N-terminal α -helical domain, a C-terminal β -sheet domain and separated by a flexible linker. RAF1 protomer dimerization occurs through interactions of their C-domains that are placed on the horizontal plan of the octamer, facing outside the enzyme. Flexible linkers bind alongside the dimer interface and N-terminal domains contribute to binding of the top and edges of the LSU dimer (Fig. 2.6). *In vitro*, RAF1- as RBCX- acts after LSU release by the chaperonin complex, preventing its rebinding to the folding machinery. In the *in vitro* reconstitution experiments of Syn7942 Rubisco, RAF1 can be mostly found as an LSU₂-RAF1₂ complex. It promotes formation of the LSU octamer and is released by SSU binding. RAF1 is conserved in the green lineage and it seems that it has the same role in both

prokaryotes and eukaryotes^{274,286}. RAF1 most probably co-evolved with LSU in plants and cyanobacteria as was shown by high correlation factor between their phylogenies. In addition, co-expression of *Arabidopsis* LSU with its cognate RAF1 in tobacco plants improves the formation of chimeric AtLSU₈NtSSU₈ Rubisco as compared to the plants not expressing AtRAF1, suggesting high sequence specificity between LSU and its chaperones²⁸⁷. This might explain the low levels of hybrid Rubisco accumulation reported in other publications^{279,288}.

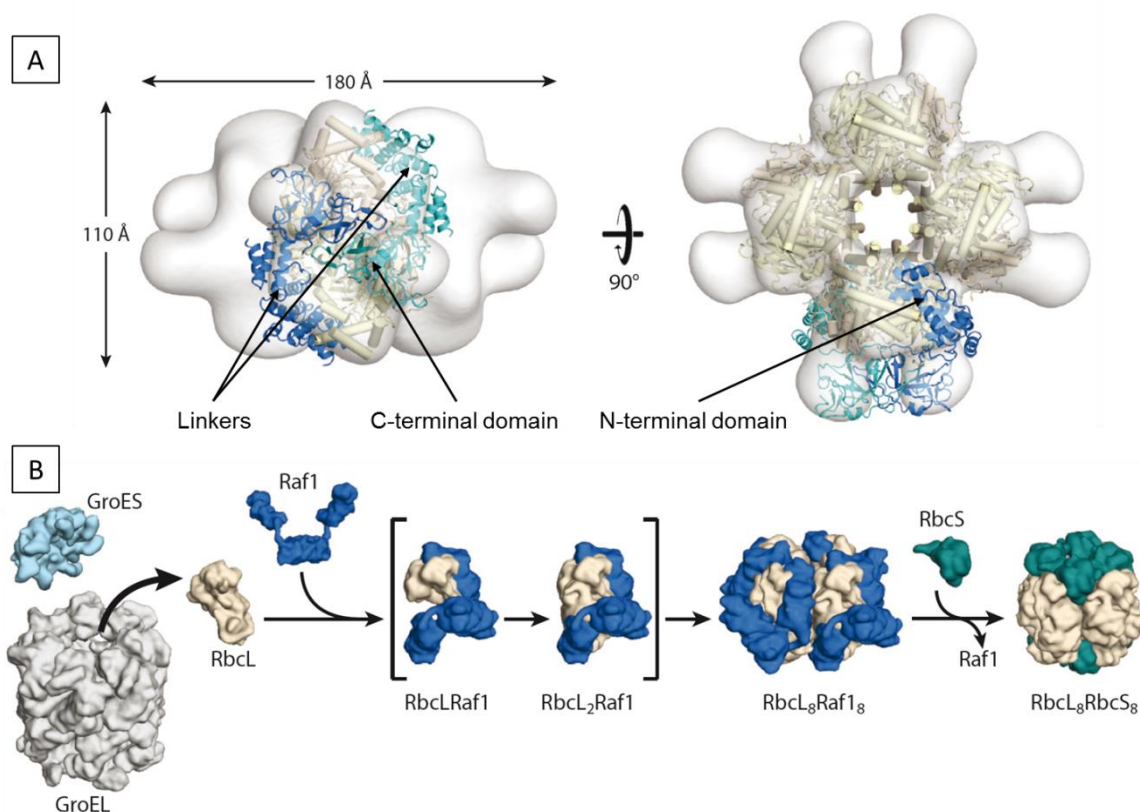


Fig. 2.6: RAF1 structure and interaction with LSU

A- Reconstitution of LSU₈-RAF1₈ complex based on structure of *Arabidopsis thaliana* RAF1-1a (EMBD EMD-3053). Left: side view of the complex with RAF1 dimer in blue and cyan. RAF1s are dimerizing by their C-domains. N-terminal domain binds to top and bottom of the LSU dimer as seen on the top view (right). Linkers “hug” the LSUs.

B – Model of RAF1 interaction with LSU. RAF1 would mimic the behavior of RBCX by binding the GroEL/ES-liberated LSU preventing its rebinding and subsequently stabilizing LSU throughout the oligomerization until LSU₈ core is formed. Finally, SSU would replace RAF1 dimers. Modified from³

Comparison

RAF1 and RBCX coexist in the green lineage and for now their suggested function and mechanism of action are very similar. Nonetheless, both chaperones bind to different places on the LSU dimer surface: RBCX interacts with the C-terminal extension of LSU and contacts the N-terminal domain of the other LSU from the dimer, while RAF1 binds to the top and bottom of a dimer and embraces one side of the dimer. Whether both chaperones are required for Rubisco synthesis *in vivo* is not clear. First, almost no RBCX protein could be detected both in prokaryotes and eukaryotes^{281,282,284} and its absence had no detectable phenotype in cyanobacteria (at least in Syn7942) suggesting that it might have lost its function because of RAF1. At the same time deletion of RAF1 in *Synechocystis* sp. PCC 6803 also appears not to influence Rubisco biogenesis²⁸⁹ which is in bright contrast with maize protein²⁷⁴. This raises the question whether in cyanobacteria both chaperones are redundant and share the same pathway, cooperating and exchanging freely. The evolutionary purpose of this suggested redundancy is not clear, however it is possible that strict control of Rubisco formation was crucial and aimed to prevent accumulation of non-beneficial mutations. On the other hand, RAF1 and RBCX can be redundant but act specifically in certain stress conditions²⁸⁹. It is also possible that in eukaryotic organisms, where RAF1 function would prevail (according to the strong phenotypes of knock-outs), the mechanism of Rubisco assembly has changed due to genetic (no longer in operon) and spatial (chloroplast and nucleus) separation of *rbcL* and *rbcS* genes. In these conditions, maybe only the more stable interaction of LSU and RAF1 stayed beneficial. This of course does not exclude that both chaperones could play specific and for now undiscovered roles and be indispensable in given conditions: for example during arrest of cytosolic translation, signaling etc.

RAF2

Recently, a new Rubisco accumulation factor has been discovered. RAF2 is a conserved protein in autotrophic bacteria, cyanobacteria, algae and plants. It shares homology to pterin-4 α -carbinolamine dehydratase (PCD) but is inactive as such, due to conserved active-site disruption. Additionally, in plants, RAF2 has a conserved, N-terminal extension of unknown function (Fig. 2.7). From the bacterial homolog's structure, RAF2 is assumed to act as a dimer, but could form dimer and tetramer in plants²⁹⁰. In maize RAF2 deletion causes a less pronounced Rubisco-deficiency phenotype (~10% accumulation) than RAF1 but still leads to seedling lethality²⁹⁰. Through cross-linking and immunoprecipitation experiments it was shown to interact with

both LSU and SSU- nevertheless with much higher affinity towards SSU. The prokaryotic (*Halothiobacillus neapolitanus*) homolog of RAF2 (acRAF) increases Rubisco production in *E. coli* which confirms its chaperone-like function²⁹¹. In α -carboxysomes-containing bacteria, RAF2-encoding gene seems always part of the carboxysome operon together with the Rubisco subunit genes (and others) suggesting its importance in their formation. The mechanism of action of RAF2 is only partially unraveled, however it is suggested from co-immunoprecipitation experiments that RAF2 might stabilize SSU in the stroma before Rubisco assembly. It is also proposed that it can interact with RAF1 and BSD2 LSU chaperones, which led authors to suggest that SSU can form dynamic intermediates with the chaperones and/or LSU which would stay in dynamic equilibrium until the holoenzyme is formed. This is in contrast with the cyanobacterial dogma that SSU can fold spontaneously in the cytoplasm and stays stable and soluble before fixing to the enzyme's core.

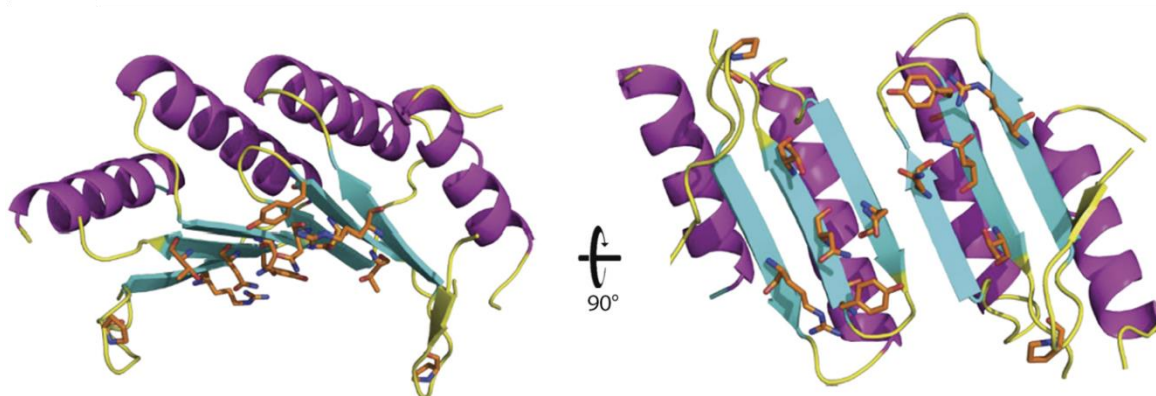


Fig. 2.7: acRAF2 structure

Structure of acRAF2 dimers from *Thiomonas intermedia* (PDB: 4LOW). Conserved, solvent-exposed residues marked as sticks. On the right panel two antiparallel protomers shown separately. Modified from²

SSU folding

As mentioned above, small subunit is thought to fold without any interference. In eukaryotes, SSU is imported through the chloroplast envelope via the TIC/TOC translocon, and undergo transit peptide cleavage followed by methylation of Met₁^{292,293}. It was suggested that it could interact with CPN60 complex at an early stage after the import but to a lower extent than LSU²⁹⁴. It was also suggested to form complexes with LSU during the enzyme's biogenesis²⁹⁵ which would stay in line with the models of LSU oligomerization and its dynamic interactions with assembly chaperones. The recently discovered RAF2 protein was shown to interact with SSU in vitro, maybe to stabilize the small subunit before the docking onto the enzyme's core²⁹⁰. Very interestingly, small subunit itself may participate in LSU folding. No LSU chaperones have been found in organisms carrying Red Type Rubisco (Type IC and ID). Bacterial red-type LSU can be produced in *E.coli* and in in vitro reconstitution experiments with only GroEL/ES. However, the enzyme could properly fold only in the presence of SSU. SSUs of Type IC and ID have C-terminal extensions that fold into β -hairpins that promote their oligomerization and play a crucial role in the enzyme assembly²⁹⁶.

3. Interlude

The aim of this study was to shed light on the complicated process of chloroplast protein biosynthesis. What better subject than Rubisco? It is the most abundant protein on Earth co-responsible for the appearance of oxygen, for the majority of biomass production and for the main sequestration pathway of nowadays highly emitted and popular carbon dioxide. All that makes Rubisco important for evolutionary, economic and ecological reasons. It then has been studied for almost 70 years now and still, its surprising inefficiency remains a mystery. A lot of focus, especially in recent years, has been dedicated to improving Rubisco catalytic activity to match the increasing demand on food supply for an ever-growing human population. The estimation of the real value of those improvements is subject to an open debate which is out of the scope of this manuscript, which would rather touch the technical difficulties encountered. A major obstacle for Rubisco-improvement effort was the difficulty to express it to a satisfactory level in an *in vitro* system. Rubiscos from higher plants seem especially stubborn in this matter. One of the reasons is probably the incompletely described biosynthesis pathway of Rubisco synthesis. As I tried to highlight in the introduction, gene expression, its

regulation and the subsequent assembly is a complicated multilayered process harboring a huge number of factors interacting at different stages and at different places in the cell. At the beginning Rubisco bearing only two subunits that bind in a 1:1 ratio seemed a convenient model to analyze chloroplast translation and protein assembly. Using *Chlamydomonas reinhardtii* we decided to decipher this process for Rubisco form I *in vivo*. As will be apparent further on, it revealed to be, as it usually happens, more complicated than what it first looked like.

In the first part of the introduction I tried to give a general overview on the complexity of chloroplast expression system evolved from two systems to allow the proper production of mostly the photosynthetic machinery. A second part was consecrated to Rubisco regulation in particular as it was our protein of choice to study protein biosynthesis in *Chlamydomonas*. The next three chapters of this manuscript will summarize the work I did during my stay in the laboratory.

1. It has been suggested that nuclear-encoded proteins regulate plastid gene expression on a post-transcriptional level acting in pairs as stabilization (M-) and translation initiation (T-) factors. This overview is changing as new discoveries about gene specific expression peculiarities are being made. The first part of my thesis was consecrated on finding a potential partner of MRL1 – a stabilization factor for *rbcL* transcript. Results of those tryouts are being presented in a first chapter that follows.
2. The main portion of my stay in the laboratory was dedicated to unravel the regulation of *rbcL* gene expression with regard to its assembly process, with particular emphasis on the prevailing CES process that regulates the coordinated assembly of all photosynthetic complexes. Rubisco is the sole soluble enzyme that is recognized to be affected by CES and the one example that might be the most conserved. Results of this part of my work are presented in form of a manuscript in the second chapter of this manuscript.
3. Finally and most importantly, I will present in a last chapter additional results, troubleshooting, discussions, and perspectives for the future.

I hope you will find what follows useful.

4. Chapter I: MRL1 role in *rbcL* translation regulation

4.1. Introduction

Plastid-encoded genes are regulated mostly on post-transcriptional level through the action of nucleus-encoded proteins (Organelle Trans-acting Factors (OTAFs))²⁹². They are mostly constituted of modular proteins of tandem repeats that can interact with RNA and/or other proteins. They are responsible for gene specific post-transcriptional regulation, mostly via, stabilization, processing (maturation, splicing) of mRNA, translation (target recognition, formation of initiation complex and ribosome recruitment)⁶⁸. One such factor is MRL1, a conserved pentatricopeptide repeat protein (PPR) that is necessary for stabilization of the mRNA of gene (*rbcL*) coding for large subunit of Rubisco¹ (Cre06.g298300). MRL1 is a soluble protein containing 11 PPR repeats in its N-terminal part, a conserved domain called C-domain and a long, C-terminal extension with no specified domains. The N-terminus of the C- domain organizes into tandem of α -helices (with no PPR homology) that contribute to the solenoid structure (Fig. 4.1).

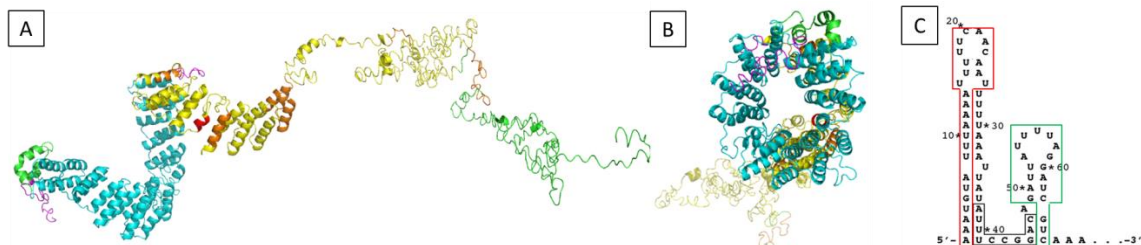


Fig. 4.1: Predicted structure of MRL1

A – Side and B – front view of MRL1 structure prediction. In fluogreen and magenta N-terminal domain (without predicted 20 aa of TP); in cyan PPR domain with 11 PPR repeats; in yellow C-domain (α -helical part and unstructured tail), in green unstructured C-tail; in orange sequence insertions as compared to *Volvox carteri*. Prediction done by I-Tasser using PPR10 structure as a model.

C – Secondary structure of the first 69 nt of *rbcL* transcript. 26 nt MRL1 footprint found in⁶ boxed in red; 18 nt footprint found in⁸ boxed in green; stabilizing sequence described in^{10,11} boxed in black. Modified from¹².

MRL1 would interact with *rbcL* 5'UTR through its PPR-domain, probably with the utmost 5' stem-loop of the 5'UTR region as a corresponding sRNA footprint has been recently discovered and whose sequence exhibits high conservation across the green lineage^{6,8}. MRL1 binding protects *rbcL* mRNA from 5'-3' exonucleolytic degradation¹, and as such is a so called M-factor. It was not shown however what the mode of action of MRL1 is: whether it merely physically protects the 5' end of the transcript or whether its binding changes the loops' structures allowing further translation. Definite characterization of the binding site and MRL1-RNA interaction is still to come. For many photosynthetic related genes, couples of stabilization (M) and translation related (T) factors have been discovered (see introduction: Table 1.2). More and more proofs however suggest that this is not a general rule and that individual trans-factors can play multiple roles in gene expression^{164,187}. We wanted thus to test whether *rbcL* transcript requires additional proteins for proper translation.

4.2. Part I - Searching for new OTAFs involved in *rbcL* gene expression

4.2.1. Genetic approach: Mutagenesis

Forward genetics has been a powerful tool in identifying genes implicated in photosynthesis. *Chlamydomonas reinhardtii* is perfectly suited as a model for such functional genetic studies: it displays facultative autotrophy, a haplobiontic life cycle allowing mutations to be readily expressed and screened for, a sexual reproduction that is easily induced and a zygote stage, from which progenies are readily isolated by dissection, its genetic compartments can be transformed. All these reasons made it an organism of choice for many research groups. Recent efforts to allow reverse genetics were conducted to generate tens of thousands of *Chlamydomonas* insertional mutants, which are nowadays available for the community^{293,294}.

Negative selection screen

A prior genetic screen using insertional mutagenesis conducted in the laboratory generated over 12000 transformants. They were screened for Rubisco deficiency using the in-house developed high-throughput fluorescence-based method²⁹⁵. Rubisco-deficient cells are non-photosynthetic, thus all light energy they absorb at some point is reemitted as heat and fluorescence which can be easily detected. Rubisco defects have a distinctive phenotype in which the level of emitted fluorescence increases slowly during the time of illumination to reach its maximal level (which will depend on light intensity)

(thus, indicating a complete block of the electron transport chain)¹³(Laura Houilles: thesis). Typically, fluorescence induction curves were monitored over 7 minutes of constant illumination. Out of all the transformants, only two were specifically affected in Rubisco accumulation, but at the end were found to be new alleles of the already discovered *MRL1* gene (Up to date, ten mutated *MRL1* alleles have been independently discovered and partially characterized (see Table 4.1). This low success rate raised the questions whether straightforward insertional mutagenesis does not suffer from insertion hot spot bias that limits its utility in the research of new photosynthesis-related trans-factors.

Table 4.1: Known *mrl1* alleles

| Allele | Name | Mutation | Reference |
|----------------|---------|-----------------------------------|-----------|
| <i>mrl1.1</i> | wcf3 | Exon 2 aphVIII insertion | 247 |
| <i>mrl1.2</i> | 75.5EN | Exon-intron splicing junction SNP | 296 |
| <i>mrl1.3</i> | wcf2 | 5'UTR aphVIII insertion | 247 |
| <i>mrl1.4</i> | L11A | 5'UTR aphVIII insertion | |
| <i>mrl1.5</i> | L54B | TOC insertion in exon 2 | |
| <i>mrl1.6</i> | 111480 | CDS CIB1 insertion | |
| <i>mrl1.7</i> | 155641 | Intron CIB1 insertion | 294 |
| <i>mrl1.8</i> | 155641b | CDS CIB1 insertion | |
| <i>mrl1.9</i> | 050384 | 3'UTR CIB1 insertion | |
| <i>mrl1.10</i> | 169685 | CDS CIB1 insertion | |
| <i>mrl1.11</i> | UV1 | UV mutagenesis induced | This work |

mrl1.6 - 1.10 origin from the CLiP library of mutants (prefix: LMJ.RY0402)

We decided to test another screening method by implementing a negative screen recently developed for *Chlamydomonas* based on the use of cytosine deaminase (CD)²⁹⁷. Cytosine deaminase (encoded by the *codA* gene) is an *E. coli* enzyme catalyzing conversion of cytosine to uracil. It can also convert 5-fluorocytosine (5FC) into 5-fluorouracil (5FU), which is cytotoxic for *Chlamydomonas*. Combining the endogenous promoter and 5' regulatory regions of a gene-of-interest (GOI) to the chloroplast-optimized *codA* coding frame would produce a chimeric gene, which can be used for further mutagenesis to search for trans-factors targeting the GOI regulatory sequence. 5-fluorocytosine (5FC) sensitive lines could be rescued after the mutagenesis through a mutation affecting the 5'GOI-*codA* chimeric gene expression. It could be attained either

by mutations in the 5'UTR of GOI (cis-mutations) or by mutations in a nuclear gene, whose product is required for the GOI expression (trans-mutations). Finally, mutations could impact directly *codA* CDS sequence. Those sorts of transformants can be easily discarded by an additional fluorescence-based screen aimed to pinpoint photosynthetic defects. Trans-mutations can be afterwards distinguished from the cis-mutations by backcrossing of the mutated strain to the WT from the observed inheritance pattern. Only half of the progenies should inherit the mutation following mendelian segregation of the nuclear genes (whereas chloroplast cis-mutations would be uniparentally inherited).

Our collaborators (Rosie Young and Saul Purton, University College London) created two strains of *Chlamydomonas* harboring the *codA* gene under the control of i) *rbcL* 5'UTR (CR1 strain) and ii) *rbcL* 5'UTR with its first 81 bp of *rbcL* coding sequence (CRE1 strain). The latter was created in a second round to promote the chimeric gene expression based on the observation that *rbcL* coding sequence enhances transgene expression²⁴⁵) as the CR1 line did not result in the CD expression sufficient enough to confer 5FC sensitivity (Fig. 4.2 C). All these strains were created in a cell-wall deficient background, a necessary requirement for consistent 5FC uptake into *Chlamydomonas* cells.

***rbcL* regulatory sequence drives cytosine deaminase expression**

This second chimera, using *rbcL* 5'UTR with the first 81 bp of *rbcL* coding sequence allows to produce cytosine deaminase to detectable levels, that were shown to be sufficient to confer 5FC sensitivity (Fig. 4.2). CRE1 strain grows on TAP medium to levels comparable with CW15 WT but is unable to grow on TAP supplemented with 5FC (5FC at 2 mg/mL, Fig. 4.2B). Note that, WTs and CR1 (expressing low levels of CD as monitored by immunoblot directed against the HA tag; Rosie Young: personal communication) are insensitive to 5FC which demonstrates that sensitivity is caused by CD expression.

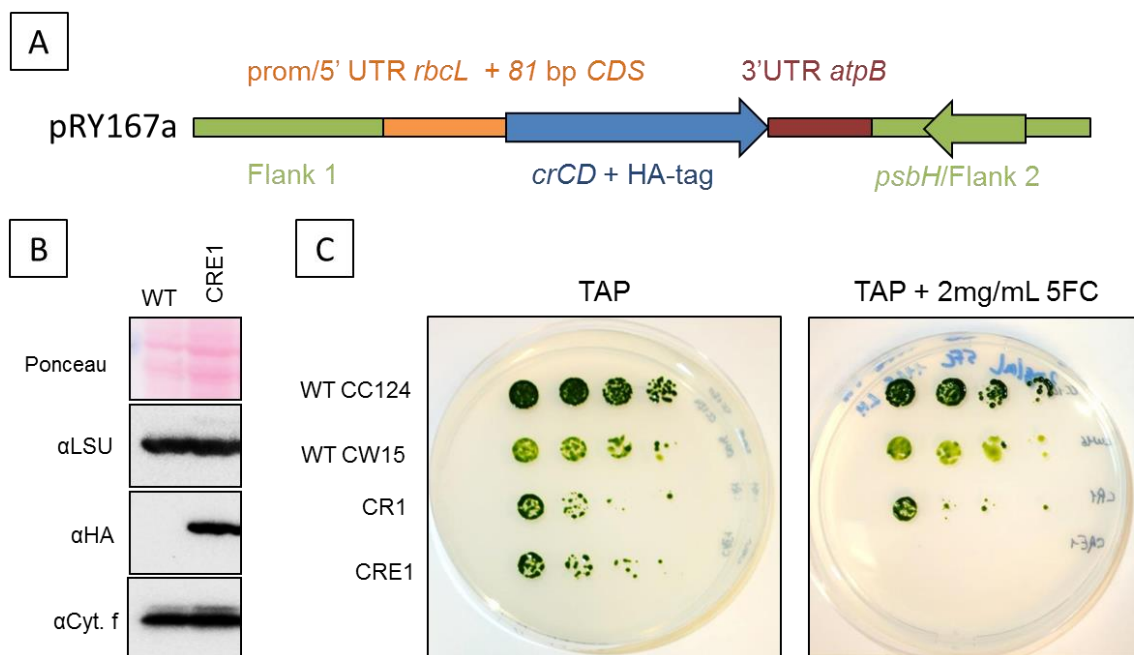


Figure 4.2: CRE1 expresses cytosine deaminase and is sensitive to 5-fluorouracil (5FC)

A - Schematic representation of pRY167a plasmid fragment used for transformation of TN72 strain to generate CRE1 line. Flank 1 and *psbH*/Flank2 – flanking sequence for recombination; prom/5'UTR*rbcL*+81bpCDS – regulatory sequence of *rbcL* gene; *crCD*+HA-tag – Chlamydomonas chloroplast optimized CDS of cytosine deaminase gene with additional C-terminal HA-tag; 3'UTR *atpB* – 3'UTR and termination sequence of *atpB* gene. Bars not to scale. Copyright: Rosie Young.

B – Immunoblot of total proteins from wild type CW15 (WT) and CRE1 strain using antibodies against LSU, HA-tag (detecting the HA tagged-Cytosine deaminase protein) and cyt. *f* as a loading control.

C - Growth test of WT (CC124), Cell-wall less control strain (CW15), and two transformants bearing cytosine deaminase gene (CD) under the control of *rbcL* 5'UTR (CR1) and *rbcL* 5'UTR and first 81 bp of *rbcL* coding sequence (CRE1). All strains were grown on TAP (left) and TAP supplemented with 5-fluorouracil (right) to test its toxicity for the cells. Cell drops were from four dilutions: 10^6 , 10^5 , 10^4 , 10^3 cells per mL (from left to right).

Mutagenesis of CRE1

The CRE1 strain, being sensitive to 5FC, is a suitable tool to search for unknown factors interacting with *rbcL* 5'UTR. Therefore, we conducted a mutagenesis campaign by subjecting CRE1 to different mutagen agents. First, an insertional mutagenesis, using the pBC1 cassette containing the *aphVIII* paromomycin resistance gene was conducted (as previously done in the lab²⁴⁷) (Xenie Johnson, Laura Houilles: thesis). In parallel

experiments, CRE1 was treated with fluorodeoxyuridine (FUDr, 1mM) and ultra violet radiation (UV, $3.6 \text{ erg} \times \text{mm}^{-2} \times \text{s}^{-1}$ for 30 seconds) for random chemical and physical mutagenesis respectively. Approximately 800 clones resistant to 5FC were recovered and tested either by fluorescence or by growth test on MIN medium to screen for photosynthesis deficiency. From the strains tested, 5 were non- and 2 partially-photosynthetic (4 were obtained by insertional mutagenesis: I1-I4 and 3 by UV mutagenesis: UV1-UV3). The selected strains were further analyzed by western blotting (Fig. 4.3).

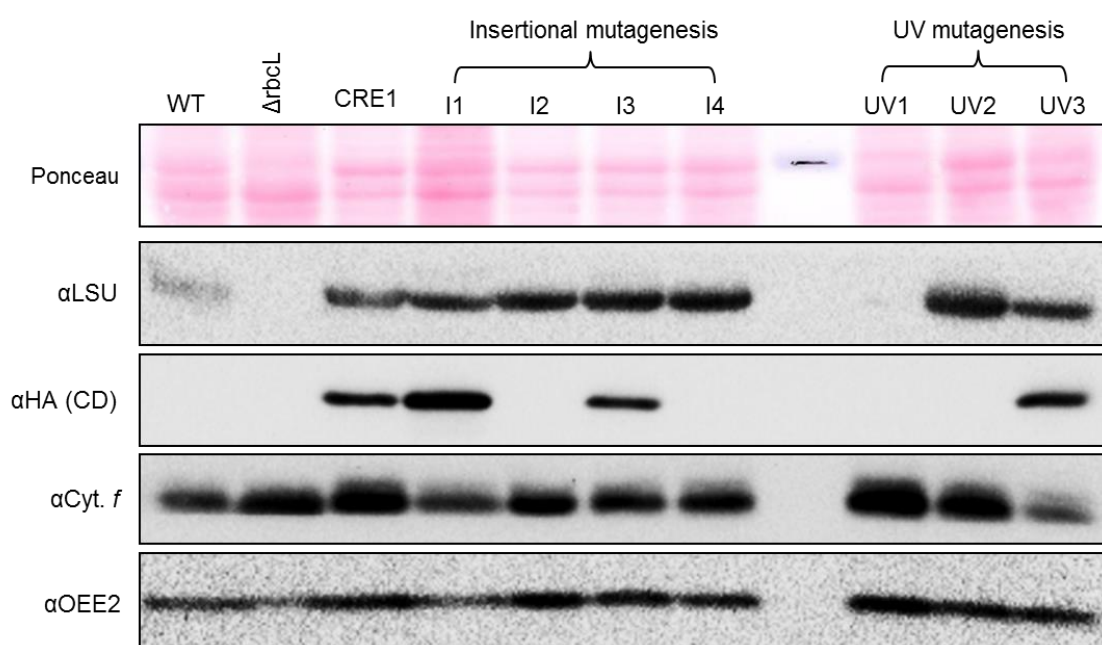


Figure 4.3: Accumulation of Rubisco and CD in non-photosynthetic clones recovered from CRE1 mutagenesis

Immunoblot of total proteins from wild type CW15 (WT), $\Delta rbcL$ (Rubisco negative control) CRE1 (CD positive control) and strains picked after mutagenesis (4 from insertional and 3 from UV mutagenesis). Antibodies against LSU, HA-tag (cytosine deaminase) and Cyt. *f* and OEE2 as loading controls.

Out of 7 strains, 3 (I1, I3 and UV3) did not have any defect in CD or LSU protein accumulation, This phenotype suggests that two mutations occurred in these strains; a first one either in *codA* CDS affecting its activity but not its stability, or a mutation that influence 5FC metabolism (see Discussion and perspectives), combined with a second site mutation affecting photosynthetic gene expression (I1 and UV3 came to be affected in PSI and cyt. *f* accumulation respectively: not shown). I2 and UV2 did not accumulate CD but were not affected in LSU accumulation. Here again, this is probably caused both

by a cis-mutation or cpDNA rearrangement caused by the mutagenic treatment (transformants of this kind were also observed previously (Rosie Young: personal communication), coupled to a second site mutation inducing a photosynthesis defect (we did not follow the origin of photosynthetic deficiency). Finally, in the UV1 strain, neither LSU nor CD did accumulate making it the sole potential strain of interest. To test genetic linkage between UV1 and MRL1 locus, a functional complementation analysis of UV1 was performed with a plasmid containing a WT version of *MRL1* (pMRL1CdomHA). In parallel, genetic linkage was assessed by a cross between UV1 and *mrl1* (*mrl1.5* BC132) mutant strain. If the mutations lie in two different genes, photosynthetic progenies (whose proportion compared to non-PS cells would depend on the loci linkage) should be obtained. UV1 was successfully complemented with pBC1 and no photosynthetic cells could be found in the progenies of the cross, suggesting that both strains carry a defect in the same gene, leading us to conclude that UV1 is another *mrl1* allele.

4.2.2. Biochemical approach: Co-immunoprecipitation of MRL-HA

MRL1 was found in HMWC in the chloroplast stroma that most probably contain also other elements, among others *rbcL* mRNA¹. We hypothesized that additional proteins e.g. new *rbcL* T-factor could be found within the complex and tried to use a biochemical approach to identify these factors.

Generation of a tagged MRL1 strain

Because of the lack of appropriate antibody against MRL1, we generated a HA-tagged version of the protein to allow its detection and possibly its precipitation. Two MRL1-HA versions were created: in the first one, the tag lies at the very C-terminus of the protein (MRL1-HA Cter) and in the second one within an exposed loop localized at the beginning of the so called MRL1-C domain of the protein (MRL1-HA Cdom)(see¹)(Fig 4.4). The reason behind was that modular proteins of the TPR and PPR families could undergo posttranslational proteolytic cleavage (Stefania Viola, Olivier Vallon: personal communication), which could stay unnoticed with a tag at the very C-terminus. The resulting constructions were used to complement *mrl1* knock out strain (*mrl1.5* BC132-). The MRL1-HA protein was detected in the transformants at around 150 kDa – a size which is compatible with the predicted size of 149 kDa (after removal of the 20 aa of predicted cTP, HA-tag included). It is however in contrast to what was observed (~116 kDa) in the initial MRL1 paper using an MRL1-raised antibody. In our hands however, this antibody did not give any reliable signal. Interestingly, in MRL1-HA Cdom transformants, MRL1 indeed seemed to be fragmented an observation that was made as

well for several other OTAFs (Stefania Viola, unpublished) (Fig. 4.4). It remains unknown whether all the forms are truly MRL1 related, or merely HA cross-contaminants. If they turn out to be MRL1 products, whether they all are active or not, and what the physiological relevance of this increased proteolytic susceptibility might be, remains to be tested. The absence of a full-length product in transformant 5 was initially puzzling. However as the C-terminal tail of MRL1 is dispensable for its activity, the lower molecular mass band we observed may then represent an incomplete MRL1 that is still able to complement the mutation and that was obtained by incomplete transgene insertion¹. Contrary to the situation observed for MR1L-HA Cter transformants, MRL1 HA Cdom protein accumulation was not always consistent with the accumulation of LSU. In those MRL1-HA Cdom transformants the HA-tag insertion may have also had an impact on the stability and/or activity of MRL1. We decided to continue to work only with MRL1-HA Cter strains.

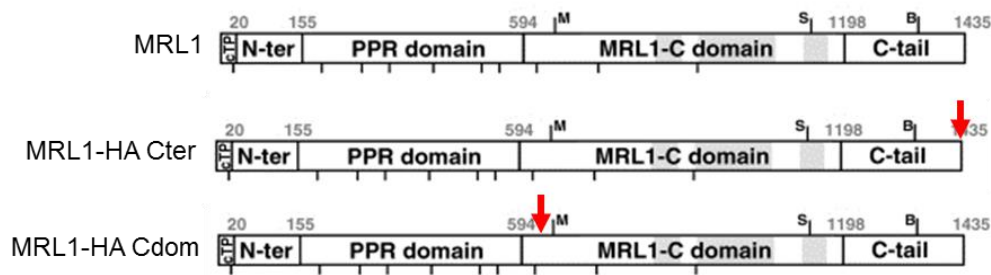


Fig 4.4: MRL1 protein

Schematic representation of MRL1 protein sequence and two HA-tag containing versions. Red arrows point to the location of the HA insertion. Modified from¹

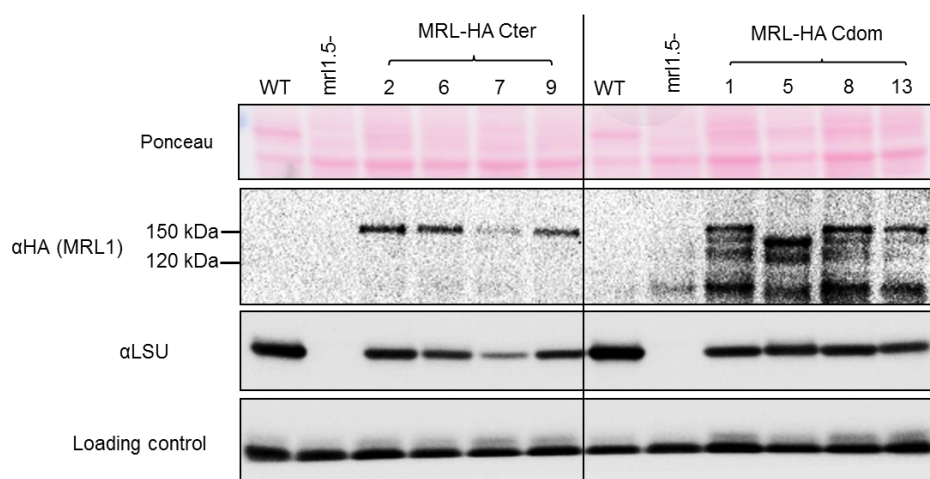


Figure 4.5: MRL1-HA strains

Immunoblot of total proteins from wild type T222+ (WT), MRL1 knock out strain (mrl1.5 BC132-), selected MRL-HA Cter B and MRL-HA Cdom transformants strain. Antibodies against LSU, HA-tag (MRL1) and Cyt. f as a loading control were used.

ImmunoPrecipitation and Mass-spectrometry

For further characterization, we used a complemented strain displaying the highest levels of Rubisco accumulation (MRL1-HA Cter A#2) in a co-immunoprecipitation experiment in a search for MRL1 partners. French press lysates from MRL1-HA Cter A#2 and a *mrl1* knock down strain complemented with an untagged version of MRL1 (cMrl1.5)(used as a negative control) were ultracentrifuged, the supernatant was concentrated using centrifugal concentrating units with a 30 kDa cut-off and further incubated with anti-HA decorated magnetic beads in a batch precipitation. After multiple washes, bound- proteins were removed from the beads by boiling and loaded on SDS-PAGE gel (Fig. 4.6).

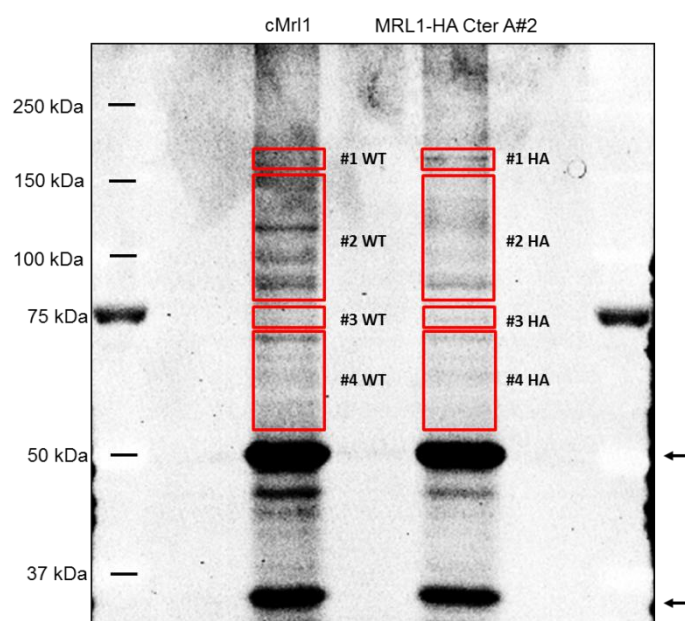


Figure 4.6: MRL1-HA Co-immunoprecipitation

SyproRuby stain of antiHA-interacting proteins in the Control strain extract (complemented *mrl1.5* (cMrl1)) and MRL1-HA Cter A#2 soluble extract. Proteins were removed from the anti-HA beads by boiling and separated on a SDS-PAGE. Molecular weights are indicated on the sides. Red boxes show the gel fragments that were cut for further mass-spec analysis. Arrows point migration of Heavy (50 kDa) and Light (30 kDa) chains of the antibody detached from the beads.

Each gel lane was cut into 4 bands, from which peptides were extracted and analyzed by Orbitrap-Mass-spectrometry at the IBPC facility with the kind help of Christophe Marchand (UMR8226). Low stringency detection thresholds were applied. A customized Augustus v5.3 of *Chlamydomonas* genome was used for protein annotation

assignment^{298,299}. In particular, the 32-OPR cluster from chromosome 15 corrected annotations was manually added to the genome list (with the kind help of Yves Choquet). The experiment was done on two biological replicates each injected twice. 1074 proteins were identified in both MRL1-HA Cter A#2 and in a control strain (cMRL1.5). Out of those, 490 proteins were found in both samples, and as such represent unspecific binding to the resin, and 335 were specific for the MRL1-HA extract (249 specific to the control) (Table 4.2). Out of them, the records with the highest MS score are presented in the Table 4.3 (Full list of the proteins in the Appendix 1), as well as their putative localization based on TargetP software prediction. We could precipitate MRL1 efficiently, as revealed both by the SyproRuby stain where MRL1 full-length protein is detected in band 1 (see Fig. 4.6), and by MS identification. Yet no obvious candidate for an interaction with it could be identified, neither by visual inspection of the SyproRuby-stained gel, nor after mass spec identification.

On the other hand, many, most probably irrelevant (e. g. flagellar proteins, pherophorins) proteins were co-precipitated. This high background level is most probably due to the high sample complexity (Fig. 4.6), resulting in unspecific protein binding to the beads or to MRL1 itself because of protein stickiness (see below).

Table 4.2: Number of proteins found in the extracts used in the Co-IP against HA antibody

| Strain | Precipitated | Specific | Unspecific | Total |
|-------------------|--------------|----------|------------|-------|
| MRL1-HA | 739 | 335 | 490 | 1074 |
| Control (cMrl1.5) | 825 | 249 | | |

TOP Scores

| Nr. | Accession | Description | Band | MS score | Target. |
|-----|--------------------|---|---------|----------|---------|
| 1 | Cre06.g298300.t1.1 | Pentatricopeptide repeat protein, stabilizes rbcL mRNA | 1,2,3,4 | 962,65 | C |
| 2 | Cre12.g531500.t1.2 | Flagellar Associated Protein, FAP134 | 2,3 | 285,94 | O |
| 3 | Cre12.g528000.t1.2 | Predicted protein; 5'-AMP-activated protein kinase, gamma subunit | 2,3 | 283,06 | C |
| 4 | Cre02.g077750.t1.2 | Flagellar Associated Protein, FAP211 | 2,3 | 272,50 | O |
| 5 | Cre05.g238687.t1.1 | Predicted protein; Pherophorin | 1,2,3 | 271,47 | O |
| 6 | | Photosystem I P700 chlorophyll a apoprotein A2 GN=psaB | 3 | 222,90 | Encoded |
| 7 | Cre03.g172950.t1.2 | Centromere/microtubule binding protein, CBF5 | 1,2 | 219,05 | Cyt. |
| 8 | Cre02.g118300.t1.2 | DEAD box ATP-dependent RNA helicase | 1,2 | 202,66 | O |
| 9 | Cre17.g726750.t1.2 | 3-deoxy-D-arabino-heptulosonate 7-phosphate synthetase | 1,3 | 201,80 | C |
| 10 | Cre02.g081700.t1.1 | predicted protein; Pherophorin | 3 | 201,27 | O |

Others, RNA-related

| Nr. | Accession | Description | Band | MS score | Target. |
|-----|--------------------|--|------|----------|---------|
| 1 | Cre17.g708750.t1.2 | No information | 3 | 196,15 | O |
| 2 | Cre10.g441200.t1.2 | Predicted protein; LUPUS LA PROTEIN-RELATED | 1,2 | 141,48 | O |
| 3 | Cre01.g028200.t2.1 | DEAD-box RNA helicase | 3 | 138,51 | O |
| 4 | Cre06.g307850.t1.2 | No information | 1,2 | 122,60 | O |
| 5 | Cre16.g662902.t1.1 | Predicted protein; TRANSLATION INIT. FACTOR EIF-2B SUB. BETA | 1,2 | 113,54 | O |
| 6 | Cre07.g314900.t1.1 | Predicted protein; ATP-dependent RNA helicase pitchoune | 1,2 | 108,81 | O |
| 7 | Cre10.g437150.t1.2 | Cyclin-related pentatricopeptide repeat protein PPR6 | 1 | 97,34 | O |
| 8 | Cre08.g376200.t1.2 | Predicted protein; TPR repeat-containing protein | 1 | 69,22 | M/O |
| 9 | Cre13.g607900.t1.1 | Predicted protein; TPR protein CPLD68 | 2 | 59,76 | M |
| 10 | Cre15.g638303.t1.1 | Predicted protein; RAP domain containing protein | 1 | 55,45 | M |

Table 4.3: MRL1-HA Co-immunoprecipitated proteins

The table presents the ten best scoring proteins found specifically in the MRL1-HA sample. In the bottom panel, specific, but low-scoring candidates predicted to interact with RNA. Band – number of the band in which protein was detected. Target: C – chloroplast; M – mitochondrion; Cyt. – cytosol; O – other than chloroplast/mitochondrion; Encoded – chloroplast encoded. Samples from two biological replicates were analyzed twice each; in orange records found in all assays.

4.3. Part II - Is MRL1 alone?

Fruitless search of the MRL1 partner/T-factor led us to consider that MRL1 may serve as a lone OTAF in *rbcL* expression and play a dual role in its regulation by stabilizing the transcript (M-factor) and initiating translation (T-factor). We then decided to test this hypothesis using standard biochemical techniques.

4.3.1. MRL1 level limits *rbcL* mRNA accumulation and is required for LSU accumulation.

rbcL transcript does not accumulate in the *mrl1* knock out mutants. Complementation with tagged MRL1 yielded several strains expressing different amounts of MRL1. This is an expected result as the level of transgene expression should depend on the integration site of the insert and on the number of the insertions. Using independent transformants that contain variable amounts of MRL1-HA we determined how MRL1 levels affect LSU transcript and protein accumulation. Fig. 4.7 shows the MRL1-HA content in these transformants and the resulting restoration of *rbcL* mRNA (A), and LSU (B) accumulation profiles.

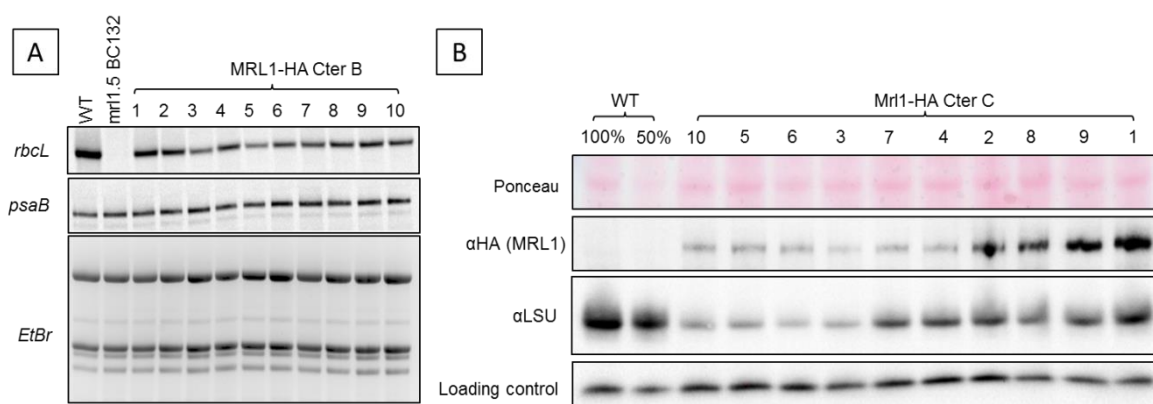


Fig 4.7: MRL1-HA accumulation correlates with *rbcL* mRNA and LSU levels

A - RNA blot and B – Immunoblot showing different levels of *rbcL* mRNA/LSU accumulation in a series of MRL1-HA Cter transformants. *psaB* probe used as a RNA loading control; Appropriate antibodies were used to detect LSU, HA-tag (MRL1) and Cyt. *f* (as a loading control).

While characterization of additional transformants would be required to better evaluate these relationships, the linear fit observed between *rbcL* transcript accumulation and MRL1 amounts still indicates that MRL1 is limiting, as would be expected from the

inferred action of this PPR, by a direct interaction with the mRNA target. LSU accumulation on the other hand is less directly dependent on MRL1, as shown by the asymptotic fit observed. We could not reach WT levels neither of *rbcL* transcript nor of LSU protein accumulation. Strain with the highest *rbcL* mRNA accumulation had only 39% compared to WT transcript's level, while reaching 84% of WT LSU accumulation level. In general, we could observe a twofold higher accumulation of LSU than its transcript suggesting additional level of expression regulation (Fig. 4.8) (see Introduction: Nucleus and chloroplast crosstalk).

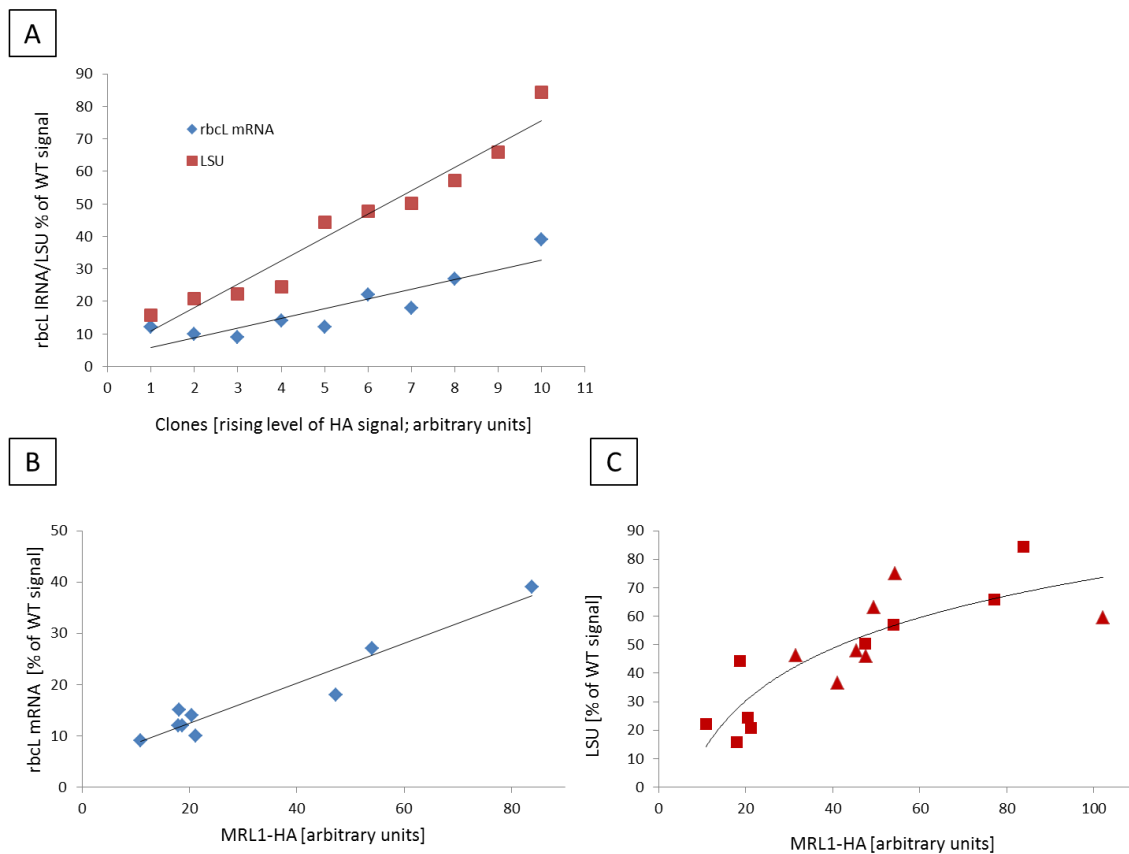


Figure 4.8: MRL1-HA is necessary for *rbcL* mRNA and LSU accumulation

A - Plot showing the accumulation levels of *rbcL* transcript and LSU protein relative to WT in different MRL1-HA Cter C transformants.

B – Correlation between MRL1-HA abundance (relative units) and *rbcL* transcript level;

C – Relationship between MRL1 accumulation (relative units) and LSU accumulation in strains from B. Red triangles represent strains from a first round of transformation (MRL1-HA Cter A) that were added for protein quantification.

4.3.2. *MRL1 is a stable protein*

As in our transformants MRL1 levels corresponded with Rubisco accumulation, MRL1 might serve a regulatory function *in vivo* merely through a control of its accumulation. One of the prerequisites for that is the short lifetime of a trans-factor, which is necessary for dynamic changes in its abundance. We selected the MRL1-HA Cter transformant displaying the highest level of Rubisco to test the stability of MRL1 together with LSU. We conducted an immunochase experiment treating the exponentially growing cells with inhibitors of cytoplasmic and/or chloroplast translation (Cycloheximide and Lincomycin) over 6 hours. As shown on the Fig. 4.9, MRL1 is relatively stable as no visible decrease of its accumulation was observed during the time of the experiment. Similarly, Rubisco levels did not change after 6h of the treatment, which was expected as the assembled enzyme is very stable. MRL1 stable behavior thus parallels what has been seen for TAA1¹⁸¹, MDH1 and TDA1 (Stefania Viola: unpublished results) bearing long half-lives, and is in contrast to the short half-lived MCA1 factor.

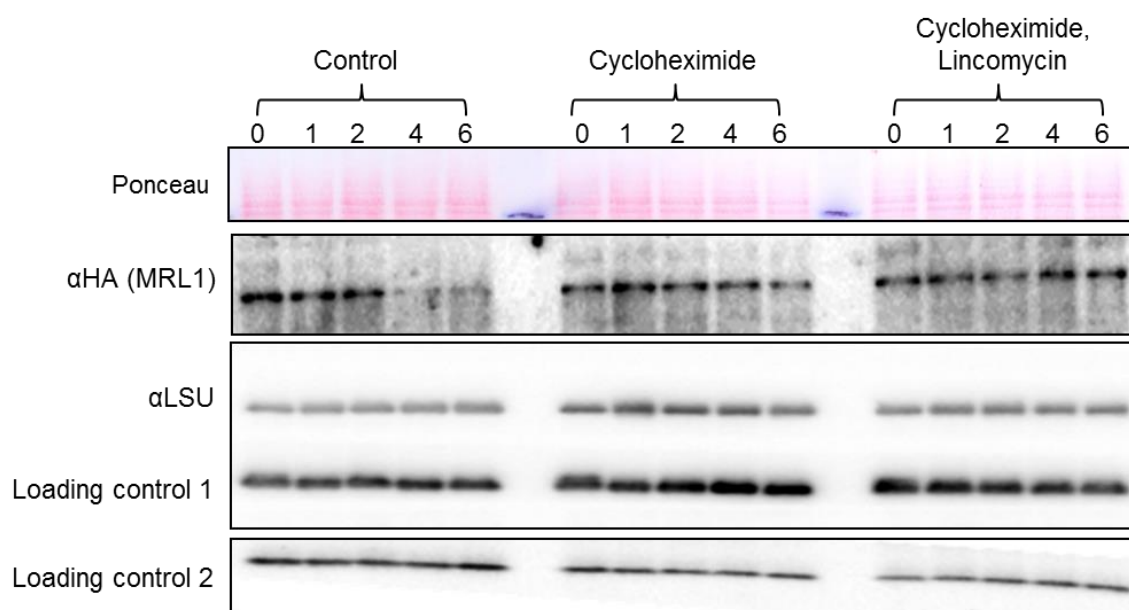


Figure 4.9: MRL1 stability

Immunoblot after an inhibitor-chase experiment of total proteins from MRL1-HA (A#2) Cter strain subjected to cycloheximide or cycloheximide and lincomycin treatment as compared to the no-treatment control. Timepoints: 0, 1, 2, 3, 4 and 6 hours are indicated. Antibodies against LSU, HA-tag (MRL1), cyt. *f* (Loading control 1) and OEE2 (Loading control 2) were used. Lower HA signal in 4h and 6h timepoints in the control results from a technical problem.

4.3.3. *MRL1* is required for *rbcL* translation

If MRL1 is truly the sole nucleus-encoded factor specifically required for *rbcL* expression, we wondered whether it would not have a dual function both in transcript stabilization and translation. To this end, we used poly(G) tracts to stabilize *rbcL* transcript to test whether MRL1 is required for *rbcL* translation. Poly(G) sequences form stable secondary structure that prevent the 5' to 3' exonucleolytic cleavage of the mRNA rendering *rbcL* mRNA a priori independent from MRL1 stabilizing effect. We introduced the 18 G residues at *rbcL* transcription start site by biolistic transformation of the $\Delta rbcL$ (ΔR T1.3+) and $\Delta rbcL$;mrl1 (ΔR T1.3+;mrl1.5+) strains at the *rbcL* locus. As shown in Fig. 4.10, the poly(G)*rbcL* constructs allow *rbcL* mRNA accumulation in *mrl1*-mutated background to about 65 % of the polyG-*rbcL* mRNA levels observed in WT.

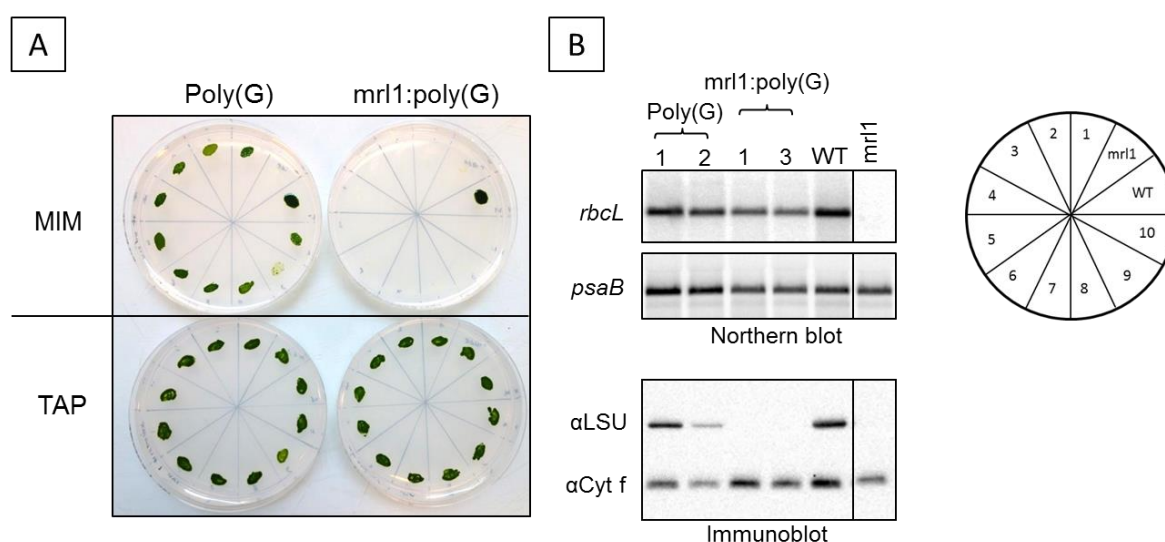


Fig. 10: Poly(G) effect on *rbcL* and MRL1 dependence

A – Growth test on MIN and TAP medias of poly(G) transformants in WT (PolyG) and *mrl1* (*mrl1*:poly(G)) background. On the side - scheme of the strains positioning on the plates.

B – northern blot (top) and immunoblot (bottom) of selected strains from A. Northern were probed with *rbcL* and *psaB* used as a control. Antibodies against LSU and cyt. *f* (loading control) were used. Line on the blots marks removed lanes that were irrelevant for the figure.

Surprisingly, polyG-*rbcL* transcripts accumulate to much lesser extent than the *rbcL* mRNA in an otherwise WT context and reaches on average 37% of WT level. The reason for this decrease induced by the polyG insertion is not known. It was unexpected, as a polyG inserted at the same position to drive the expression of an *rbcL* 5'UTR-*uidA*

reporter gene did not show a similar effect on mRNA accumulation²⁴⁶. Most importantly, the polyG-*rbcL* transcript is translatable in WT background, yielding LSU to accumulate, however this is no longer the case in absence of the MRL1 factor. While the polyG-*rbcL* construct bypass MRL1 role as an M-factor, it reveals that MRL1 is required for *rbcL* translation. As *rbcL* 5'UTR is MRL1's targets, MRL1 role in translation occurs at the initiation step.

4.4. Discussion and perspectives

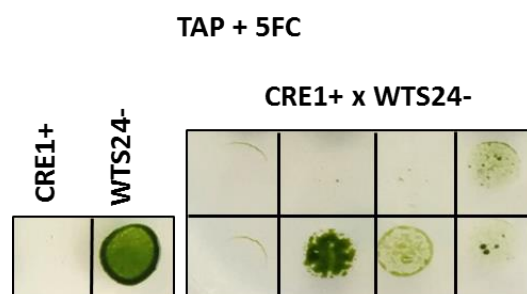
4.4.1. In a search of *rbcL* T-factor

Many nucleus-encoded, gene specific factors responsible for chloroplast gene expression have been described (see Introduction: Table 1.2). Organelle Trans-acting Factors (OTAFs) most notably constitute proteins necessary for chloroplast transcript stabilization and maturation (M-factors) and translation (T-factors). MRL1, M-factor of *rbcL* was found in ~800 kDa HMWC in the chloroplast stroma. It most probably contains also LSU and/or *rbcL* mRNA, as in $\Delta rbcL$ strain and after RNase treatment the complex' weight shifted towards ~600 kDa and 650 kDa respectively¹. The high mass suggests that additional proteins might be associated with MRL1 within the complex. We used genetic (mutagenesis) and biochemical (Co-IP) approaches in a throughout effort to find a putative T-factor of *rbcL* or any other trans-acting protein responsible for LSU expression - all in all however, without success.

First, this might be due to technical difficulties and limitations of co-immunoprecipitation coupled to Mass-spec experiments. MRL1 is a low-abundant protein which means that high sample concentration/big volumes need to be used for the precipitation. Low levels of MRL1 are even more pronounced in the complemented strains due to high silencing of the inserted tagged version of the protein. We have observed up to an 8 fold decrease in the expression levels of MRL1 in a 6 months period. Use of high sample concentration/volume for precipitation increases the risk of unspecific, contaminating protein binding to the beads. The soluble protein sample which was used for the co-immunoprecipitation experiment contains also thousands of proteins, most of them irrelevant to us that can be the source of false positive records detected. We have detected over a thousand proteins in both positive (HA-tagged MRL1) and control samples which indicates very high un-specificity of the column. Further experiments require a vast optimization, most notably decreasing the sample complexity to limit the noise. One of the possibilities would be to use isolated chloroplast stroma as an input for

Co-IP. To this end, we envision to cross the MRL1-HA strain with the cell wall deficient strain (to obtain cell-wall less MRL1-HA) to improve the harvest yield of intact chloroplast necessary for efficient isolation. Another way would be to change the tag at MRL1, which might decrease the unspecific binding. FLAG or Strep tags were successfully used in the laboratory therefore they are a reasonable alternative to be considered. Finally, increasing the accumulation level of MRL1 by overexpression is theoretically possible. However, considering the high silencing rate and the fact that we did not recover any complemented mutants with LSU amounts higher than the WT, probably not feasible.

Second, the insertional mutagenesis conducted by Xenie Johnson and Laura Houilles in the laboratory and our trials using negative selection based on cytosine deaminase resulted only in the discovery of new *mrl1* alleles. Our approach, focused on 5'UTR of *rbcL*, was especially meant to reveal specifically factors required for *rbcL* mRNA stability and translation. The fact that no potential candidate was found might be again caused by the limitation of the techniques used. On one hand, if the candidate gene encode for an essential protein (e.g. ribosomal subunit) it would not be selected during the screen. At the same time, the cytotoxicity based screen has its own technical limitations. A first and simple reason might be that not enough clones could have been screened to find the mutant. Our approach increases the specificity of the mutagenesis but does not attenuate completely e.g. the insertional hot-spot effects. In addition, the screen suffers from false positive clones, such as 5FC insensitive transformants that had unaffected *codA* expression. As authors of the screen mention it²⁹⁷, we also have observed that the cell wall acts as a partial barrier for 5FC and thus mutations leading to cell wall restoration could also be selected: crossing CRE1 to a walled WT strain (WT24-) resulted in restoring the 5FC resistance to half of progenies of the cross despite uniparental heritage of *codA* gene (Sup Fig 4.1).



Sup. Fig. 4.1:

Sensitivity to 5-fluorouracil (5FC) is altered by the cell wall presence

Drop test on TAP plates containing selective amounts of 5FC of CRE1, a WT strain (WTS24-) and an octad obtained from their cross.

4.4.2. *M-factor's dual function?*

Increasing number of evidence suggest that mRNA stabilization and translation initiation are not separated but linked together by proteins of dual function. OTAFs' division of labor onto M and T-factors is not discrete, but rather M-factors may play a more pronounced role during gene translation than initially thought. Poly(G) tracts inserted before *petD*, *psbB* and *psbD* genes in the absence of their stabilization factors restore accumulation of those transcripts but not their translation^{64,164,186,187}. The recent characterization of the TAA1 protein, whose absence has a dramatic effect on *psaA* mRNA accumulation levels and whose role in translation is revealed by the *polyG-psaA* construct even led the authors to call it a T factor, rather than an M one¹⁶⁹. This was also observed for the MCA1 protein: the *petA* M-factor is also a prerequisite for efficient translation of *petA* transcript. This is apparent as the artificially stabilized *petA* transcript was translated 10 fold less in the *mca1* mutant than with the presence of the MCA1 (MCA1 present) background¹⁶⁴. Demonstration that MCA1 could be found in HMW fractions comigrating with *petA*-associated T-factor TCA1 suggests that it may be required for TCA1 recruitment to the mRNA²¹⁰. The observation that contrary to MCA1, no sRNA footprint could be detected for TCA1 suggests that TCA1 does not bind directly to RNA and substantiates this hypothesis⁶. Whether MRL1 role in translation consists in recruiting a trans-activator or the initiation complex is still not known. Yet as mentioned above, our failure in finding a potential new *rbcL* trans-factor might have been due to technical limitations or not sufficient tenacity on our side but it is also possible that MRL1's binding is sufficient to promote directly the ribosome recruitment.

4.4.3. *Insight in MRL1 mode of action?*

We discovered that MRL1 accumulation did not change drastically after 6 hours of cytoplasmic and/or chloroplast translation arrest. Its half-life is therefore difficult to assess with the immunochase experiments due to the longer-term cytotoxicity of the inhibitors, which would affect the observations. We have launched collaboration with Aleix Gorsh and Alison Smith (Cambridge, UK) to create a tagged MRL1 under the control of a riboswitch, whose expression can be turned off and on depending on the availability of vitamins. This would allow easy tracking of the protein level over a longer time range leading to a better estimation of its half-time. Nonetheless, MRL1 is stable in standard growth conditions and its accumulation (degradation) is not influenced by the translation of the *rbcL* transcript.

Because of its long half-life, MRL1 abundance cannot change rapidly to adapt to the environmental and physiological cues. *rbcL* transcript accumulation is linearly related to the MRL1 levels. Most probably then, all MRL1 pool in the stroma is bound to the untranslated (at a given moment) *rbcL* transcripts. That MRL1 amounts are limiting is further exemplified by the reduced accumulation of transcripts driven by the same *rbcL* 5'UTR (like the endogenous *rbcL* and an *rbcL* 5'UTR-driven reporter gene, see chapter 2) when they are in several copies versus the one endogenous copy in WT. At the same time the correlation between MRL1 and LSU levels is not linear but rather follows a logarithmic curve. Note also the two-fold discrepancies between *rbcL* mRNA and LSU levels in the transformants (Fig. 4.8). Although we lack a true measure of LSU synthesis depending on MRL1 amounts, this may suggest that LSU expression is regulated on translational level, and that MRL1 is part of this process. This would be consistent with its inferred role in translation as demonstrated by the polyG-*rbcL* transformant analysis. Indeed, other M factors like the recently described MAC1 factor¹⁶⁸ show a dose-dependence for their target mRNA accumulation (in this case, *psaC*) but not for the translated product. This kind of regulation would require interaction with an additional factor or a modification of MRL1 activity (as its abundance does not change). Such a trans-factor could be a difficult to co-precipitate ribosomal subunit. Noteworthy, a *Chlamydomonas* mutant of one of those subunits: plastid ribosomal subunit 6 (*prps6*) was isolated in the laboratory. The mutation alters translation of several transcripts, among them *rbcL*. This phenotype might be linked to differences in translation initiation between chloroplast genes. The recently resolved structure of higher plant ribosome shows distinctive differences compared to the bacterial 70S ribosome, most notably, at the entry site. It was suggested that gene specific factors could interact with mRNA and/or ribosome to modulate the translation initiation. To test the possible recruitment by MRL1 of the ribosomal 30S complex, we would like to characterize this possible interaction using tag-bearing MRL1-HA strains and a strep-tagged ribosomal subunit Rps12 (to facilitate ribosome purification). We have also created double mutants of those tagged proteins in *prps6* mutant background to test how this putative interaction is challenged in the *prps6* mutant. Together these experiments may help shed light on MRL1 requirement for *rbcL* translation initiation.

4.4.4. MRL1, a regulator of Rubisco accumulation?

Changes in MRL1's lifetime and abundance may be triggered by environmental signals to control Rubisco accumulation. Three possible ways could be envisioned to regulate MRL1 activity on a longer time scale; i) transcriptional control, ii) control of its

activity through post-translational modifications, iii) and/or control of degradation rates, induced by post-translational modifications.^{168,300}

It would be interesting to test how MRL1 itself is regulated in situations where Rubisco expression is either increased or decreased. Rubisco expression is known to positively respond to variations in CO₂ and light intensity. This could be conditions to test if *rbcL* increased translation is preceded by changes in MRL1 transcript or protein accumulation. Conversely, because of its high abundance, Rubisco is highly affected by nitrogen and sulfur starvation, conditions leading to its degradation³⁰¹. It would be worth to follow MRL1 accumulation during *rbcL* translation arrest and degradation, as was already demonstrated for multiple regulatory factors, such as MCA1, TAA1 and MAC1 in different starvation conditions^{164,168,169}.

Changes in MRL1 activity could also be caused by post-translational modifications. MAC1, a factor necessary for *psaC* transcript accumulation, can be phosphorylated and its phosphorylation state varies depending on iron availability and cell redox changes (cues altering its target: a PSI subunit)¹⁶⁸. Most interestingly, MRL1 has been identified as a phosphoprotein in a large-scale proteomics study. MRL1 sequence has six possible phosphorylation sites³⁰⁰ (compared to 2 in MAC1). Therefore it would be worth exploring whether this modification could modulate MRL1's function in a dynamic way.

Acknowledgements

Our thanks go to collaborators: Rosie Young and Saul Purton (UCL, London, UK), Aleix Gorsh and Alison Smith (Cambridge, Cambridge, UK) and students that participated in the work: Carine Borges, Alexandra Zakieva and Loic Helio.

5. Chapter II: Rubisco LSU synthesis depends on its oligomerization state in *Chlamydomonas reinhardtii*

This chapter treats about Rubisco CES process in *Chlamydomonas* highlighting its occurrence as well as presenting genetic and biochemical data supporting the identification of an assembly intermediate acting as the translational inhibitor of *rbcL* gene. The results are presented in the manuscript that follows.

Rubisco LSU synthesis depends on its oligomerization state in *Chlamydomonas reinhardtii*

Wojciech Wietrzynski^a, Sonia Fieulaine^b, Francis-Andre Wollman^a, Katia Wostrickoff^a.

^a Centre National de la Recherche Scientifique, Unité Mixte de Recherche 7141/Université Pierre et Marie Curie, Institut de Biologie Physico-Chimique, Paris 75005, France

^b Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198 Gif-sur-Yvette cedex, France

Abstract

Ribulose 1,5-biphosphate Carboxylase/Oxygenase (Rubisco) appears as one of the key enzymes of the photosynthesis-driven life on Earth. In the first step of the Calvin Benson Basham Cycle, it catalyzes the fixation of the atmospheric CO₂ into biologically available, organic carbon playing a pivotal role in the global carbon cycle. Due to its environmental and economic importance it became the object of extensive, bioengineering research. Nonetheless, one of the difficulties standing against the effort of Rubisco improvement is the yet only partially unraveled mechanism of its assembly.

Rubisco biogenesis requires the expression of its two subunits. In eukaryotic organisms those two subunits are encoded by spatially separated genomes of different ploidy. Large subunit (LSU) is encoded by a single gene (*rbcL*) in the chloroplast, whereas small subunit (SSU) is produced from a family of nuclear genes (*RBCS*). Both assemble in the chloroplast stroma in a 1:1 ratio to form a hexadecameric holoenzyme. Owing to the dual genetic origin of its two subunits, the stoichiometric formation of Rubisco in the chloroplast needs to be finely tuned. It has been previously demonstrated that accumulation of LSU and SSU is a coordinated process – SSU is being degraded in the absence of LSU, while LSU translation is being hampered if SSU is not present. The latter regulatory process, linking the translation of a chloroplast subunit to its assembly state, has been coined the CES process (for Control by Epistasy of Synthesis). Interestingly, it has been shown to operate throughout the green lineage in both higher plants and green alga in case of Rubisco LSU.

Here we unravel the CES-underlying mechanism in the model alga *Chlamydomonas reinhardtii*. Using genetic and biochemical approaches we show that the process results from an autoregulation by unassembled LSU. We show in vivo the presence of Rubisco assembly intermediates accumulating in absence of assembly. Furthermore, analysis of

Rubisco LSU oligomerization mutants leads us to propose a model where CES translational inhibition is triggered by the accumulation of a specific LSU oligomer.

Introduction

Ribulose biphosphate carboxylase/oxygenase (Rubisco) is the key enzyme in the light-driven carbon assimilation pathway. Emerging around 3.5 billion years ago, even before the beginning of the oxygen-evolving photosynthesis, it is now one of the most abundant proteins on Earth ^{14,223}. In the first step of Calvin-Benson-Basham cycle, Rubisco catalysis the fixation reactions of the atmospheric CO₂ into biologically available, organic carbon. Throughout the time, different forms of Rubisco evolved and are now present in all photosynthetic organisms ^{9,223}. The most widespread clade: Form I, is present in cyanobacteria, green algae and vascular plants. It consists of two: Large (LSU) (~52 kDa) and Small (SSU) (~16 kDa) subunits. In most eukaryotic organisms those two subunits are encoded by spatially separated genomes of different ploidy. Large subunit is encoded by a single gene (*rbcL*) in the chloroplast, whereas small subunit is a product of a family of nuclear genes (*RBCS*). In chloroplast stroma both subunits assemble in a 1:1 ratio to create a hexadecameric holoenzyme⁹.

Recently, tremendous progresses have been made in the comprehension of the mechanisms leading to Rubisco biogenesis³. SSU expression is mostly regulated transcriptionally: *RBCS* genes are light induced, some of their isoforms may be organ-specific³⁰². *RBCS* genes were early characterized as being part of the PhANG genes³⁰³ (reviewed¹⁷²), a set of genes undergoing retrograde signaling in response to chloroplast translation and redox status. SSU will thereafter be imported to the chloroplast via Tic/Toc import system ³⁰⁴ where it undergoes transit peptide cleavage and post translational Met1 modification ²⁸⁸. It was also recently reported that Raf2 (Rubisco accumulation factor-2) chaperone could interact with it, possibly to stabilize the protein before its proper assembly with LSU ²⁸⁵. More is known about the biogenesis of the large subunit whose expression, although taking place in the chloroplast – an organelle of the bacterial origin, relies on a multitude of nucleus-encoded factors^{86,91,107}. One of those trans-factors is a pentatricopeptide repeat protein - MRL1 which is necessary for the stabilization of *rbcL* transcript in *Chlamydomonas* and of the processed form in *Arabidopsis* ¹. Another factor shown to play a role in *rbcL* expression is BSD2 (Bundle sheath defective-2) protein, first identified in maize ³⁰⁵. In the green alga *Chlamydomonas reinhardtii*, the BSD2 orthologue co-migrates with *rbcL* transcripts on polysomes and was proposed to interact co-translationally with *de novo* synthesized LSU polypeptide as its first chaperone ²⁷². LSU undergoes also significant post-translational modifications but their precise localization in the biosynthesis pathway is still

an open question.²³⁹ It was suggested that nascent LSU is recruited by the chloroplast folding machinery: DnaK/DnaJ/GrpE chaperones and later by CPN60/CPN23/CPN10 chaperonin complex. Furthermore LSU oligomerizes in a step-wise manner to create an octameric core of the enzyme. Because of the hydrophobicity of LSU surface making it aggregation-prone, this process requires the assistance of assembly-chaperones. Three such proteins have been described in cyanobacteria, green algae and plants: RBCX²⁷⁶, RAF1²⁶⁹ and the previously mentioned RAF2²⁸⁵ (that *in vitro* interacts to some extent with LSU). RBCX and RAF1 both are believed to stabilize the LSU during dimerization and stabilization until the binding of SSU, thereby leading to a displacement of the chaperones as demonstrated *in vitro*²⁷³. While RBCX and RAF1 indisputably can lead to folding and assembly of the L₈ core *in vitro*, their role *in vivo* and possible functional redundancy is still under debate. RBCX isn't an absolute requirement *in vivo*- at least for β -cyanobacterial species where this gene does not cluster within the Rubisco operon such as in *Synechococcus elongatus* PCC7942²⁷⁷. Whether this dispensability still holds true for cyanobacterial species presenting a Rubisco LXS operon awaits further confirmation^{276,277,306}. To date, there is no evidence of its requirement in algae and plants, where two RBCX isoforms, which would form homodimers, are found³⁰⁷. The requirement for RAF1 also seems to differ in the green lineage: RAF1 knockout is lethal for maize seedlings and results in Rubisco deficiency²⁶⁹, however in *Synechocystis* sp. PCC 6803 its absence has no evident phenotype in normal growth conditions²⁸¹.

Moreover, given the significant energetic input needed to create a necessary amount of Rubisco, the distinctiveness of its subunits' origins and their assembly, all points to the existence of a precise regulatory mechanism orchestrating its synthesis. It has been already proposed that other multimeric photosynthetic complexes: Photosystems I, II (PSI, PSII respectively), Cytochrome *b₆f* and ATP-synthase undergo a regulation process depending on their assembly state in *C. reinhardtii*^{197,203, 204, 207, 205}. Translation of certain chloroplast-encoded subunits of those complexes is epistatically regulated through the presence of their assembly partners – a process called CES (for Control by Epistasy of Synthesis). Observations that LSU synthesis diminishes in *Chlamydomonas* *RBCS* knockout mutants²⁰⁶ and tobacco *RBCS* knock-down lines²¹² suggested a similar mechanism for Rubisco. Indeed, it was proven that in maize and tobacco chloroplasts LSU translation was negatively autoregulated in an assembly-dependent manner^{213 214}. The conservation across photosynthetic eukaryotes of the CES process makes LSU an interesting case to study the evolution of the underlying mechanism. As a first step towards this goal, we aimed to undertake a molecular characterization of its actors in the genetically tractable microalgae - *Chlamydomonas reinhardtii*.

Here we highlight the mechanism of LSU translation autoinhibition in the absence of the SSU in *Chlamydomonas*. Using different Rubisco assembly mutants we were able to determine the possible *in vivo* intermediates of Rubisco formation and LSU-chaperone(s) complexes showing that *rbcL* translation is not only regulated by unassembled LSU but is dependent on its oligomerization state.

Results

***rbcL* downregulation of synthesis in *Chlamydomonas* *RBCS* mutant.** Previous work showed that deletion of the *RBCS* genes in *Chlamydomonas reinhardtii* resulted in the significant diminishment of large subunit translation²⁰⁶. We repeated this observation in an independent mutant, hereafter called $\Delta RBCS$ strain. This strain was readily isolated from backcrosses of the Cal.005.013 strain²⁹³ presenting a large deletion encompassing the two *RBCS* linked genes- to our laboratory reference strain. Contrary to the other *RBCS* mutant described-T60.3²⁰⁶, this strain is fertile allowing further genetic analysis as presented below. In absence of SSU, LSU accumulated to ~1% of wild type level (WT), revealing the concerted accumulation of Rubisco subunits. (Fig. 1A). Moreover, as expected from Khrebtukova and Spreitzer²⁰⁶, *rbcL* exhibits a lower translation in $\Delta RBCS$ strain compared to WT (Fig. 1B) as shown on the ¹⁴C labeling experiment where chloroplast synthesis is monitored throughout a 7 min radioactive pulse. This is a posttranslational process as *rbcL* mRNA level is not affected in the $\Delta RBCS$ strain as compared to WT (Fig. 1C).

LSU initiation of translation is impaired in absence of SSU. Previous reports about genes coding for proteins of photosynthetic complexes that undergo CES translation regulation show that their 5'UTR bear all cis-acting elements necessary for the process to occur^{5,207} indicating that in all cases studied so far, regulation of translation of CES proteins occurs at the initiation step. To test whether the native 5'*rbcL* regulatory sequence is required as well for *rbcL* inhibition of translation, we replaced the *rbcL* endogenous locus by a 5'UTR *psaA*-driven *rbcL* gene, using biolistic transformation of the $\Delta rbcL$ strain (ΔR T1.3+) (Supplementary Materials and Methods). The resulting 5'UTR*psaA:rbcL* transformants were fully phototrophic and accumulated WT-levels of LSU demonstrating that *psaA* 5'UTR is able to efficiently drive *rbcL* expression (Fig. 2A). We further crossed a representative 5'UTR *psaA:rbcL* transformant (mt +) with the $\Delta RBCS$ strain (Cal.13.1B; mt -) and obtained progenies with uniparental inheritance of 5'UTR*psaA:rbcL* gene and 2:2 distribution of the $\Delta RBCS$ mutation (data not shown). Progenies from distinctive genotypes (5'UTR*psaA:rbcL* and $\Delta RBCS$;5'UTR*psaA:rbcL*)

were used in pulse labeling experiment to monitor *rbcL* translation rates in nonnative 5'UTR context and test whether LSU synthesis is still affected by the absence of SSU. As shown in Fig. 2B, *rbcL* translation rate was shown to be similar between sister strains and comparable to the WT, which demonstrates that *rbcL* 5'UTR replacement by another, unrelated endogenous regulatory sequences allows *rbcL* to escape the CES regulation.

To test whether *rbcL* 5'UTR is sufficient to confer Rubisco's CES regulation to an unrelated gene, we aimed to analyze the expression of a fusion between the *petA* gene encoding cytochrome *f*, a core protein of the Cytochrome *b₆f* complex, and *rbcL* 5' regulatory region (5'UTR*rbcL*:*petA*; hereinafter referred to as the reporter) in presence or absence of SSU. We used the previously described plasmid pRFFFiK^{1,203}, to transform both the WT strain and the Δ RBCS strain (Cal.13.5A+), yielding respectively the (5'UTR*rbcL*:*petA*) and (Δ RBCS;5'UTR*rbcL*:*petA*) transformants. In normal growth conditions cytochrome *f* is a stable protein, whose accumulation level mirrors its rate of translation, and as such makes it a faithful reporter (proxy) of LSU regulation. In the experiment shown on Fig. 3 we compared the accumulation of the cytochrome *f* reporter protein in representative transformants in presence (5'UTR*rbcL*:*petA*) or absence (Δ RBCS;5'UTR*rbcL*:*petA*) of Rubisco small subunit. Cytochrome *f* accumulation driven by *rbcL* 5'UTR was shown to follow the behavior of LSU and was almost totally impaired in absence of SSU. Altogether, this indicates that both *rbcL* and the reporter gene translation initiation are controlled by Rubisco assembly state.

Translation initiation is inhibited by unassembled LSU. Two possible mechanisms could account for the observed translational repression of both *rbcL* and the reporter gene (*petA*) in absence of SSU. In the first case, the small subunit could be necessary for direct or indirect trans-activation of large subunit expression. Alternatively, in the absence of its partner, un-sequestered LSU might inhibit its own translation via an auto-regulatory feedback. Distinguishing between the two hypotheses is possible by following the expression of the reporter gene in a context where both Rubisco subunits are absent (detailed in ⁵). A trans-activation model predicts that the absence of SSU and LSU should yield a low accumulation of the reporter, similarly to what is observed in the simple Δ RBCS mutant as the reporter gene expression should be solely dependent on SSU. On the other hand, in the autoregulation model, LSU deletion would cause the absence of the inhibitor and lead to a high accumulation of the reporter, irrespective of the presence or absence of SSU. To prevent native LSU accumulation, we generated strains bearing a truncation of 116 aa in the *rbcL* gene. The central *rbcL* part is deleted, but a short, truncated polypeptide of 14 kDa composed of the N-terminal part (107 aa)

fused in frame to the C-terminus (9 aa) is predicted to be expressed. Biolistic transformation of the $\Delta rbcL$ (ΔR T1.3+) and $\Delta RBCS$ (Cal.13.5A+) strains with the pLStr plasmid carrying this truncation (Supplementary Materials and Methods) yielded strains where truncated LSU is expressed in presence of SSU (LSU_{tr}) or not ($LSU_{tr};\Delta RBCS$). We investigated LSU synthesis and accumulation rates in LSU_{tr} strains as compared to $\Delta RBCS;5'UTRrbcL:petA$ (wild-type *rbcL* context) . *In vivo* pulse labeling experiment revealed that the truncated LSU is robustly synthesized in the (LSU_{tr}) transformants (Fig 4A). Furthermore, its synthesis rate was not altered in the absence of SSU ($LSU_{tr};\Delta RBCS$). Yet, *rbcL* truncation led to the complete absence of large subunit accumulation, which could not be detected even as trace amounts in the expected 14 kDa size range (data not shown). In addition, strains with truncated LSU in the reporter gene background were generated by biolistic transformation of a $5'UTRrbcL:petA$ strain (RJ24) and $\Delta RBCS;5'UTRrbcL:petA$ strain (RCal $\Delta K5$), which had undergone beforehand excision of the selectable *aadA* marker. In the resulting transformants ($5'UTRrbcL:petA$ and $\Delta RBCS;5'UTRrbcL:petA$), *cyt. f* accumulated to WT levels, irrespective of SSU presence as seen in Fig 4B), in sharp contrast with the $\Delta RBCS;5'UTRrbcL:petA$ strain in which *cyt. f* did not accumulate (Fig 3). We concluded that full length LSU accumulation is required for translation inhibition to occur, indicating that in *Chlamydomonas*, unassembled LSU exhibits an auto-regulatory inhibition of its own translation as it was proposed for tobacco²¹³.

Assembly intermediates accumulate in absence of SSU. Rubisco assembly pathway is supposed to go through several LSU oligomerization steps followed by SSU binding^{308,309}. To investigate in which oligomerization state the unassembled, repressor-competent LSU would be, we performed a Native PAGE analysis of soluble extracts from the $\Delta RBCS$ strain (Fig. 5A). Immunoblot against LSU readily detects native Rubisco holoenzyme in a diluted WT extract (2%). Note that no other assembly intermediates were detected even after prolonged membrane exposure (data not shown) consistent with the idea that Rubisco assembly is a fast and dynamic process. Use of the $\Delta rbcL$ extracts revealed that the anti-Rubisco antibody cross-reacts with two LSU-unrelated bands, marked by an asterisk on the figure. These two bands are also found in the $\Delta RBCS$ extracts indicating further that they are not related to SSU either. In the absence of SSU, the residual unassembled LSU (corresponding to about ~1% of WT level, Fig. 1) partitions into three LSU-reactive distinctive complexes (Fig. 5A, $\Delta RBCS$ lane). Using 2D electrophoresis and immunoblotting (Fig. 5B), we identified a band migrating above 720 kDa, which we attributed from previous work on pea extracts (from Roy H. and coworkers, as early as³¹⁰) and maize²⁶⁹ to CPN60 chaperonin-bound LSU. A similar

observation was made for CPN60 bacterial homolog GroEL in *in vitro* reconstitutions^{273,309}. This attribution was confirmed with the use of a CPN60 α / β 1 directed antibody, which revealed two reactive bands. The upper one is found to co-migrate with the 720 kDa LSU complex, whereas the lower one could correspond to free CPN60 monomers. We suspected the two other LSU-associated complexes to be LSU oligomers bound to assembly chaperones. Indeed, Rubisco's LSU₈ core oligomerization is known to require assembly chaperones such as RBCX or RAF1 factor identified in maize (12) (Rubisco accumulation factor 1). Most interestingly, reconstitution experiments performed *in vitro* using denatured cyanobacterial LSU from *S. elongatus* sp. PCC7942 (Syn7942) and RAF1 in absence of SSU³⁰⁹ showed the presence of two LSU-RAF1 complexes of around 159 kDa and 764 kDa, as estimated by native gels and confirmed by SEC-MALS. The sizes of the complexes we detected here *in vivo* are somewhat different but still, appear similarly on the Native PAGE (below the holoenzyme and below LSU-GroEL/ES complex respectively), suggesting a possible association of LSU to RAF1 in *Chlamydomonas*. This prompted us to raise an antibody directed against *Chlamydomonas* RAF1 (Cre06.g308450, Sup. Fig.1). After two dimensional electrophoresis and immunoblotting of native soluble Δ RBCS extracts (Fig. 5B), we were able to detect a RAF1 signal co-migrating with LSU in these two complexes below the 720 kDa and 480 kDa markers. We note that most of the signal is found in the lower molecular LSU complex (hereafter called LMW-LSU), whose apparent size could be compatible with an inferred interaction of RAF1 as a dimer with an LSU dimer. On the other hand, in the Δ rbcL extract the HMW-RAF1 signal disappears while LMW-RAF1 signal undergoes a significant shift in position (Sup. Fig 2), suggesting a true interaction with LSU rather than a simple co-migration. These observations strongly suggest that in *Chlamydomonas* RAF1 plays a role in LSU stabilization as seen *in vitro* for cyanobacterial LSU, RAF1 most probably interacts with LSU dimer, and stays to some part associated with higher LSU oligomers *in vivo* as well.

CES regulation no longer occurs in Rubisco oligomerization mutants. To determine whether a specific LSU assembly intermediate is linked to Rubisco CES repression we designed two sets of mutations in the *rbcL* sequence aiming to alter its oligomerization and compare their effect on the CES regulatory process. We first introduced two substitutions within the *rbcL* gene (E109A and R253A) by biolistic transformation of the Δ rbcL (Δ R T1.3+) strain with the pLS₂mut plasmid. This set of mutations aimed to destabilize the formation of two salt bridges between adjacent large subunits and thus to prevent the formation of a LSU dimer (The resulting LSU₂mut (*rbcL*_{E109A-R253A}) transformants were subsequently crossed to Δ RBCS (Cal13.1B-) to generate

Δ RBCS;LSU₂mut progenies. Those strains are non-photosynthetic and accumulate soluble LSU to a lower level than those of Δ RBCS strain (Fig. 6A). To monitor the rate of LSU synthesis in the dimerization mutants we performed *in vivo* pulse-labelling experiment with ¹⁴C acetate as shown in (Fig. 6B) for two representative progenies of the cross. In both LSU₂mut and Δ RBCS;LSU₂mut genotypes, LSU translation rate is even higher than WT levels, thereby indicating that translation inhibition in absence of SSU is prevented in this mutant LSU form. Further characterization by Native PAGE and immunoblotting demonstrates that the introduced mutations prevent accumulation of any LSU intermediates except the LSU-CPN60 complex which in both genotypes is more abundant than in Δ RBCS (Fig. 6C). No monomeric LSU could be detected, suggesting that LSU dimer is the first stable form of the post-chaperonin pathway. Our data do not allow determining whether the mutated LSU monomer cannot interact with RAF1 or other chaperones or whether the resulting complex is unstable *in vivo*. The data presented reasonably rule out the CPN60-LSU complex to be able to mediate the translation repression, as the CES translational regulation no longer occurs even though the LSU-Chaperonin complex is present. A stable, non CPN60-bound LSU is necessary for the regulation of its own translation.

Next, we created a mutated LSU which we hypothesized, should be able to dimerize but would not oligomerize further due to sterical impairment caused by the mutations. To this end, we introduced a triple ("ARD": A143W-R215A-D216A) substitution in LSU sequence by transformation of the Δ *rbcL* (Δ R T1.3+) and Δ RBCS (Cal13.5A+) strains with the pLS ARD plasmid (Supplementary Materials and Methods). These residues were chosen as to create a sterical clash between the introduced tryptophan residues from two adjacent LSU dimers (Sup. Fig. 3) and to prevent the formation of inter-dimer stabilizing salt bridges. LSU₈mut (*rbcL*_{A143W-R215A-D216A}) transformants were obtained in RBCS WT and mutant context (Δ RBCS;LSU₈mut). As expected and shown for representative strains from the different genotypes, the LSU₈ mutations resulted in a complete loss of phototrophy. Soluble LSU accumulated to levels comparable to the Δ RBCS strain, irrespective of the presence or absence of SSU (Fig. 7A). We analyzed LSU translation rate of in LSU₈mut and Δ RBCS;LSU₈mut by ¹⁴C pulse-radiolabeling (Fig 7B). In both mutants, LSU was shown to be synthesized even at a higher rate than WT (compare LSU₈mut lane and Δ RBCS; LSU₈mut lane), irrespective of the presence of SSU. This indicates that the introduced substitutions prevented LSU negative autoregulation to occur.

To further substantiate this conclusion, we combined the same LSU₈ substitutions to the presence of the 5'UTR*rbcL-petA* reporter by transformation of a representative

$\Delta RBCS$;5'UTRrbcL:*petA* transformant, who had undergone *aadA* marker removal (RC Δ K) using the pLS ARD plasmid. The resulting transformants were crossed to the WT strain (WTS24 mt-) to segregate the $\Delta RBCS$ mutation and isolate progenies bearing the LSU₈ mutations combined to the 5'UTRrbcL:*petA* reporter gene in presence (LSU₈mut;5'UTRrbcL:*petA*) or absence ($\Delta RBCS$; LSU₈mut;5'UTRrbcL:*petA*) of SSU. We monitored the *petA* reporter gene translation by analyzing *cyt f* accumulation. Fig 7C shows the results obtained for representative progenies, as compared to $\Delta RBCS$;5'UTRrbcL:*petA* strain in which as shown in Fig 3, *cyt f* reporter accumulation is strongly diminished due to CES regulation. In sharp contrast, even though LSU levels are found to be similar in the LSU₈mut mutants (LSU₈mut;5'UTRrbcL:*petA* and $\Delta RBCS$; LSU₈mut;5'UTRrbcL:*petA*) compared to $\Delta RBCS$ (with a slight overaccumulation, reaching 2%), cytochrome *f* on the other hand is expressed and not affected by the presence or absence of SSU. Noteworthy, cytochrome *f* accumulation is even found to accumulate to higher extent compared to the strain exhibiting the native LSU in presence of the reporter gene (5'UTRrbcL:*petA*), thereby exhibiting a similar trend as seen in LSU synthesis. Altogether, this indicates that both LSU and cytochrome *f* undergo a high sustained synthesis in the ARD mutants, irrespective of Rubisco assembly state.

Rubisco assembly intermediates in the LSU₈mut oligomerization mutant reveal the LSU CES repressor. The LSU₈ mutant strains, with or without SSU, were analyzed by Native PAGE in comparison with the $\Delta RBCS$ strain to characterize the pattern of accumulation of LSU intermediates (Fig. 7D). Detection with the anti-Rubisco antibody yields an identical pattern for LSU₈mut and $\Delta RBCS$; LSU₈mut extracts. The high molecular weight LSU-CPN60 complex observed above 720 kDa is present and more abundant than in $\Delta RBCS$ strain. The Rubisco-specific band that we attributed to LSU dimer bound to RAF1 is still present, but is slightly less abundant. Last, a new Rubisco specific band of about 100 kDa, of low abundance and diffuse appearance, can be observed in both LSU₈mut mutant strains (Fig 7D, dashed box). Native Rubisco form II holoenzyme from *Rhodospirillum rubrum* soluble extracts separated on similar CN PAGE was detected at the same position indicating that this band most likely represents an LSU dimer (data not shown). Most interestingly, the high molecular weight LSU-RAF1 complex present in $\Delta RBCS$ is absent in LSU₈mut mutants. We conclude that the ARD mutations indeed prevent the further oligomerization of LSU dimers thereby preventing the formation of the HWM-LSU complex, which we attribute to an LSU₈-RAF1 species. Most interestingly, we suggest that this HMW-LSU complex is likely to be the inhibitor of

rbcL translation in Δ RBCS strain, as its disappearance caused by ARD mutations is concurrent with the escape from CES regulation (Fig. 7C).

Discussion

Coordinated expression of very abundant proteins constituting photosynthetic complexes and originating from different cellular compartments is crucial for cell's energetics. It was previously demonstrated that in *Chlamydomonas* a number of chloroplast genes encoding core subunits from photosynthetic complexes undergo regulatory loops depending on their assembly state with the remaining complex's subunits. This feedback, called the CES process (Control by epistasy of synthesis), occurs at the level of translation initiation. Its importance may be reflected by its prevalence, as CES subunits have been identified in all photosynthetic complexes so far: Photosystems I and II, Cyt b6f, ATP synthase and Rubisco (²⁰⁶ and this work) allowing for fine tuning of their expression by respect to the presence of their assembly partners. Moreover Rubisco is an especially interesting case, as a CES behavior for LSU has also been observed in higher plants, providing the first example of the conservation of this regulatory process to higher multicellular eukaryotes. To shed more light on the mechanisms of this phenomenon we used *Chlamydomonas* as a convenient model for genetic approaches to demonstrate that *rbcL* expression is controlled by LSU assembly state. We could further demonstrate that this control depends on the oligomerization state of LSU giving an insight on the Rubisco biosynthesis pathway.

LSU CES process results from a negative autoregulation on translation initiation.

We confirmed that LSU is a CES subunit in *Chlamydomonas*, as an inhibition of *rbcL* translation occurs in the absence of the small subunit (Fig 1). Through swapping of *rbcL* 5'UTR we were able to show that this regulatory sequence contains the cis-elements responsible for the assembly-dependent regulation (Fig. 2), as its absence prevents the CES regulation to occur (5'UTR_{psaA}-*rbcL* construct; Fig. 2). Moreover, *rbcL* 5'UTR is sufficient to confer the CES regulation to an unrelated gene, indicating that *rbcL* coding sequence is not required (5'UTR_{rbcL}-*petA* construct; Fig 3A). This demonstrates that the regulated step occurs at the level of LSU initiation of translation, and does rely neither on a regulation of translational elongation, whose inhibition has been suggested to be released upon exposure from dark to light for several chloroplast proteins including LSU²⁵² nor an early co-translational degradation. LSU translation was found to be inhibited under oxidative stress as well³¹¹. This inhibition was further suggested to result

from an autoregulation, induced by a structural conformational modification in oxidized LSU leading to the exposure of the otherwise buried LS N-terminal domain. This domain adopts a ferredoxin fold structure, similar to an RNA Binding domain (RBD), and was found to have RNA binding capacity^{216,256}, although unspecific. In this model, LSU N-terminal domain would bind *rbcL* mRNA when exposed to an oxidative stress, resulting in the translation inhibition. However, contrary to LSU CES process, the regulated step would not be translation initiation but rather elongation²⁵⁶.

We used cytochrome *f* accumulation as a proxy to monitor the efficiency of translation inhibition conferred by *rbcL* 5'UTR to the *petA* reporter gene. Surprisingly, cytochrome *f* accumulation, which mirrors its synthesis rate, drops below the detection limit (under 3% of WT level, fig 3) in absence of Rubisco assembly. The basis for this seemingly more efficient translational inhibition compared to the endogenous *rbcL* gene translation inhibition (10%) is currently not known. We note however that the 5'UTR*rbcL*-*petA* reporter fusion is not expressed as efficiently as the endogenous *petA* gene, but accumulate to only about 50% in an otherwise WT genetic background (compare 5'UTR*rbcL*-*petA* to WT dilution series, Fig. 3). This lower expression could be expected from the titration of limiting factors required for LSU expression, which in this genetic context, would partition on the two copies of *rbcL* 5'UTR. Such a limiting factor could be the MRL1 factor, required for *rbcL* mRNA stabilization¹. Indeed, mRNA levels of both *petA* and *rbcL* are decreased by half in this *rbcL* 5'UTR-*petA* strain (data not shown). Yet, this titration is without effect on LSU expression, while apparently limiting for cytochrome *f* accumulation. This reflects probably a non- optimal translation, whose effect could be further negatively enhanced in a repressing context, thereby accounting for the reporter gene's strong repression in absence of SSU. Nonetheless, it still proved an effective tool to decipher LSU CES process.

In particular, use of the reporter gene expression allowed us to determine that LSU undergoes an autoregulation. Indeed, preventing native LSU accumulation by altering LSU sequence or structure in the truncation (LSU_{tr}, Fig. 4A), dimerization (LSU₂mut, Fig. 6B) or oligomerization mutants (LSU₈mut; Fig. 7B) invariably results in a high synthesis rate, which is no longer negatively regulated in absence of SSU, as revealed either by pulse labeling experiments or cytochrome *f* reporter gene accumulation. Note that the enhanced cytochrome *f* accumulation in these strains can truly be compared to our ΔRBCS; 5'UTR *rbcL*-*petA* standard as all the strains bear two *rbcL* 5'UTR copies, and should be similarly titrated by *rbcL* 5'UTR specific limiting factors.

Fate of unassembled LSU: As mentioned previously, the CES process results in an inhibition of synthesis. We noted that in absence of Rubisco assembly, LSU translation is inhibited but not completely impaired as newly-synthesized LSU is readily detected in pulse-labeling experiments (Fig 1). Yet LSU accumulation level drops to about 1%, as estimated by comparing the LSU signal in Δ RBCS to a WT dilution series. This unassembled LSU was shown to partition between 3 complex species (Fig. 5C and discussion below), that all in all appear fairly stable as shown by immunochase over more than two hours (data not shown). The discrepancy between synthesis and accumulation rates thereby suggests a bottleneck in LSU assembly, indicating that a factor, most probably one of the assembly chaperons, is limiting. This is further exemplified in the Δ RBCS;5'UTRpsaA:*rbcL* strain, exhibiting high LSU synthesis rate but again much lower Rubisco accumulation levels (up to 8% of WT levels). This suggests that the maximal amount of unassembled LSU that can be stabilized by the chaperonin and chaperones reaches about 8% of the WT level. Therefore even in conditions where synthesis rates are decreased, part of the neo-synthesized LSU most likely undergoes proteolysis as well, when not stabilized by other factors. We further note that the LSU₈mut oligomerization mutant and LSU₂ dimerization mutant have comparably high LSU synthesis rate, but lesser accumulation levels compared to Δ RBCS. We infer this to the fact that the stability of LSU in these mutants is partially compromised ($t_{1/2}$ of 30 minutes in LSU₈mut mutants, data not shown). Beside protease-driven disposal of excess LSU, alternative hypotheses such as LSU-insoluble aggregates formation as proposed by ³¹² under oxidative stress, or by ³¹³ in the same Δ RBCS strain could occur. We could not reveal such aggregates in the Δ RBCS strain, but did observe triton-insoluble LSU in both oligomerization mutants (data not shown). Most probably these would form prior LSU loading on the CPN60/CPN23/CPN10 complex as these strains exhibiting high translation rate of *rbcL* transcript show only a slight increase in the abundance of CPN60-LSU form compared to Δ RBCS (Fig. 7D). The insoluble fractions would have in a large part being precipitated by our protein extract preparation method and gone unnoticed. Whether true inclusion bodies form in the oligomerization mutants remains to be determined. All in all, these considerations do not affect our conclusions, as soluble LSU is stable enough to accumulate in the LSU₈ and LSU₂ oligomerization mutants.

Insights into Rubisco assembly pathway. Our characterization of the LSU assembly intermediates accumulating in the Δ RBCS strain revealed the existence of two LSU containing complexes, beside the CPN60-LSU complex identified previously in plants by *in organello* translation(see ³¹⁰). We suggest these to be LSU oligomers associated to

the RAF1 chaperone (Fig. 5B) based on the following criteria: i) RAF1 is known to interact with LSU from cross-linking experiments in maize²⁶⁹ and forms stable LSU-RAF1 complexes in reconstitution of cyanobacterial LSU with purified RAF1³⁰⁹; ii) Chlamydomonas RAF1 and LSU are found to co-migrate in the Δ RBCS strain, but not in the Δ RBCL strain; iii) the presence of these complexes vary in the oligomerization mutants used.

Both oligomerization mutants depicted the expected phenotypes suggested from structure prediction. Notably, all affected residues are conserved throughout the organisms bearing Rubisco Form IB (from cyanobacteria to higher plants). Similar mutations introduced in recombinant LSU from *Synechococcus* sp. PCC6301 yielded similar results in *in vitro* reconstitution experiments²⁸⁰. Introduction of the R212S replacement (equivalent to R215 in Chlamydomonas) led to LSU-dimers formation, but not to higher oligomeric forms, whereas the E106Q mutation (Syn. numbering, equivalent to E109 in Cr) led to compromised dimerization but not to free LSU accumulation. In our cases, we expected even more stringent phenotypes as multiple mutations were combined at once.

Altogether, this suggests that the LMW-LSU complex migrating below the 480 kDa native marker may represent the LSU₂-RAF1 intermediate, whose estimated size in Chlamydomonas would be around 200 kDa (2×52 kDa LSU + 2×51 kDa RAF1 = 206 kDa). Accordingly, this complex is not present in the dimerization mutant. The observation that in the LSU₈ oligomerization mutant, LSU dimers without RAF1 are formed (Fig 7D) could either suggest that RAF1 amounts are limiting or that the LSU mutations prevent efficient RAF1 binding. The second HMW-LSU-RAF1, migrating above the holoenzyme complex, is present in Δ RBCS, but no longer in the LSU₈ oligomerization mutant ARD. As such it most probably represents an octamer of large subunits complexed to at least one RAF1. *In vitro* work revealed the presence of a 760 kDa LSU₈-RAF1₈ complex. However, the signal ratio of antibodies against LSU and RAF1 between those bands (Fig 5) observed *in vivo* suggests that the proportion of LSU to RAF1 in the HMWC is not the same as in the lower band, indicating either a more labile interaction of RAF1 with the octameric LSU core, or a different composition with at yet unknown interactants. That points to LSU₂-RAF1 being most probably the stable form of RAF1 intermediate as mentioned previously *in vitro*³⁰⁹.

Interestingly, in-organello pulse experiment on pea chloroplasts (which would be in essence limited by the availability in unassembled SSU, or chaperones) performed either at room temperature or at 4°C to slow the assembly process, identified respectively a 7S

LSU associated complex attributed to LSU dimer³¹⁴ and an LSU₈-like species called Z³¹⁵ whose sizes could be compatible respectively with our LMW- and HMW-LSU oligomer³¹⁵. Notably, we and others^{269,314,315} did not detect LSU monomer in any of our mutants, and in particular not in the LSU₂mut dimerization mutant. This indicates that free LSU monomer is not a stable form, but rather a very transient and fragile state further supporting the current model for Rubisco biogenesis³: newly-synthesized LSU would need to be kept unfolded, maybe with the help of a dedicated chaperone such as BSD2^{271,272}, until its loading on the chaperonin. Its release would be followed by a quick dimerization involving either/or both RAF1 and RBCX chaperones, leading to stabilization of the LSU₂-RAF1 intermediate, before further oligomerization up to an LSU₈ core, still chaperone associated, followed by SSU binding and chaperones displacement. Whether or not these assembly intermediates highlighted in absence of SSU represents the real Rubisco assembly pathway cannot be definitely inferred from our data. However, the fact that the assembly intermediates in the ARD mutant show the same pattern with or without SSU, would not favor the hypothesis of an interaction of SSU with the LSU dimer. Using the same reasoning, nor would SSU interact with LSU monomer as the assembly intermediates do not change in the dimerization mutant, irrespective of SSU availability.

Tentative identification of the repressor form. As the CES process results from an autoregulation by LSU, it must be mediated by one of the three stable assembly intermediate found to accumulate in absence of SSU. CPN60 bound LSU, representing the first step of LSU folding, is frequently observed in pulse-chase experiments in wild-type situation and is thus unlikely to serve as a regulator of expression. This is further supported by our observations that all oligomerization mutants we tested accumulate this intermediate even though translation inhibition of *rbcL* does not occur (Fig 6 and Fig. 7D). (Higher amounts of CPN60-bound LSU in the un-repressed mutants reflect probably higher rate of LSU translation.) Rather, the effector of the translation inhibition is therefore likely to be attributed to the HMW-LSU complex, whose absence is correlated to the escape of the CES process in both the dimerization and LSU₈mut mutants (*rbcL*_{E109A-R253A} and *rbcL*_{A143W-R215A-D216A}). Interestingly, our data rule out the possibility that the repressor would be constituted by free monomeric LSU or by the LSU dimer. At the same time, in absence of the assembly partner it would seem reasonable to proceed with the assembly up to the step where the assembly partner binding reaction would occur: assembly could then resume rapidly whenever the partner is available again. The suggested LSU₈-RAF1 oligomer would fulfill this criteria, as it is been proposed to be the last oligomerization step before SSU binding. Unfortunately, our data do not allow to

precise the exact stoichiometry of RAF1 in the repressor form. It is thus possible that the HMW-LSU-RAF1 complex contains other proteins that would mediate the translational repression. Interestingly, the RNA binding activity of Rubisco RRM could suggest that the *rbcL* mRNA could be in this complex, thereby rendering it unavailable for translation. In such case, the repression would result from a direct interaction between LSU- and its transcript. Structural analysis of the LSU octamer complexed to RBCX in cyanobacteria indeed shows that SSU binding induces structural changes²⁸⁰. Whether the RRM domain is affected, cannot reasonably be questioned before a better definition of the full composition of the repressor. Alternatively, the CES translation inhibition could be mediated by an additional trans-acting factor, as has been suggested for *cyt. f* regulation. In this model, the MCA1 protein that is responsible for *petA* mRNA stabilization is being targeted for degradation via the presence of *cyt f* unassembled repressor motif²¹⁰, resulting in translation inhibition. The MRL1 factor, interacting with *rbcL* mRNA and promoting its stabilization, would be therefore a likely candidate¹.

Evolutionary conservation of Rubisco CES process. The present work highlights *Chlamydomonas* LSU as an autoregulated CES subunit, as was previously demonstrated for LSU from higher plants, thereby indicating a possible conservation of the underlying mechanism. Interestingly, the main actors both in the assembly pathway (such as RAF1 and RBCX) and in *rbcL* biogenesis (such as MRL1) are conserved in the green eukaryotes. All, but MRL1, are conserved in cyanobacteria as well. So far, the CES process has been viewed as a regulatory process linked to the endosymbiosis event to allow fine-tuning of the assembly process, as a mean to adapt to the different synthesis levels of subunits encoded in different genetic compartments. Whether the CES process is indeed an organellar specificity in the case of LSU is currently been investigated. Special attention will be given to the role of the MRL1 organellar trans-factor in this process. Yet, as MRL1 is long-lived (data not shown), the process would differ from the model suggested for CES control of cytochrome *f* translation, where MCA1 induced degradation in absence of cytochrome *f* assembly would result in the observed inhibition of translation. Altogether our work reveals that LSU CES regulatory mechanism is conserved in photosynthetic eukaryotes, and sets the bases for further exploration of the underlying factors at work.

Materials and Methods

Cultures and strains

If not stated otherwise, wild type and mutant strains of *C. reinhardtii* were grown on solid Tris-acetate-phosphate medium (TAP)³¹⁶ supplemented with agar and in liquid cultures under continuous, dim light ($7 \mu\text{M photons} \times \text{m}^{-2} \times \text{s}^{-1}$) on an orbital shaker (120 rpm) at 25°C. Cells from exponentially growing cultures ($2 \times 10^6 \text{ cells mL}^{-1}$) were used for all experiments.

Chlamydomonas genetics methods

Chloroplast transformation was done as described in¹⁹⁷ using an in house- built helium-driven particle gun. Chlamydomonas mating and progeny isolation were done as in Harris; 2009³¹⁶.

Nucleic acids manipulations

If not stated otherwise DNA manipulations were done following standard protocols as in Sambrook et al. 1989³¹⁷. RNA extractions and blotting was performed as in Drapier et al., 2002²⁴⁸.

Plasmids and strains preparation

Plasmids carrying mutations aimed to introduce a truncation (pLStr) or a triple ARD substitution in LSU sequence (pLS ARD), to prevent LSU dimerization (pL2mut), or carrying the *psaA*-driven *rbcL* gene (paAR) are described in supplementary materials and methods. All plasmids contain the 5'*psaA-aadA-atpB* 3' selection marker conferring resistance to spectinomycin, flanked by direct repeats³¹⁸ allowing the cassette removal in absence of selection pressure, at neutral positions either at *rbcL* 5' (BseRI site) or 3'end (AflII site). Plasmid pRFFFiK aimed to express the *petA* gene from *rbcL* 5' regulatory regions was described in¹. Biolistic transformation was used to transform appropriate recipient strains, which were as specified in the text the WT T222+ strain, the *rbcL* deletion strain ΔR T1.3+, the RBCS mutant Cal.13.5A+ strain (back-crossed progeny of the CAL005.01.13 strain described in²⁹³), and RCal ΔK strain (Cal13.5A+ transformed with pRFFFiK (containing the *petA* sequence where the endogenous *petA* 5'UTR was swapped by the *rbcL* promoter and 5'UTR, described in¹) and subjected to cassette removal³¹⁸). Transformants were brought to homoplasmy by 6 rounds of subcloning on selective media (TAP supplemented with 500ug/mL Spectinomycin), after which homoplasmy was confirmed by PCR analysis.

Pulse experiment

Chlamydomonas ^{14}C -acetate pulse experiment was done as described in Drapier et al., 2007²⁰⁵. 5×10^6 cells were washed once in MIN-Tris medium then resuspended in 5 ml fresh MIN-Tris medium and incubated for 1 h at RT with vigorous shaking to remove the acetate from the medium at a light intensity of $30 \mu\text{mol photons} \times \text{s}^{-1} \times \text{m}^{-2}$. Subsequently, 5 μl of $1 \text{mg} \times \text{ml}^{-1}$ cycloheximide and 50 μCi of ^{14}C -acetate were added simultaneously to the cells. After 7 min of vigorous shaking cells were mixed with 35 ml of cold TAP medium supplemented with 40mM acetate and immediately spun down. Cell pellets were afterwards washed in 5 mM Hepes buffer supplemented with EDTA-free protease inhibitors mix (Roche), resuspended in 0.1M DTT, 0.2M Na_2CO_3 and flash-frozen. Prior to denaturation, samples were suspended 1:1 in 5% SDS, 20% sucrose solution, boiled for 1 minute, then spun down at 12 000g for 15 minutes. The supernatant was subsequently loaded on urea 12-18% polyacrylamide gradient gels using in house-built gel system. Afterwards, gels were Coomassie-stained, dried and exposed to an autoradiography screens for at least one month. Phosphorescence signal was measured using a Typhoon FLA 9500 phosphorimager (GE Healthcare).

Protein analysis

Protein electrophoresis in denaturing conditions was performed according to the modified Laemmli protocol³¹⁹. For protein loading of “whole cell” samples, chlorophyll fluorescence was used for quantification according to¹⁹⁷. Samples were suspended 1:1 in 5% SDS, 20% sucrose solution and boiled for 1 minute, then spun down at 12 000g for 15 minutes to remove insoluble material and subsequently analyzed using 12% SDS-polyacrylamide gels.

Colorless Native electrophoresis (CN-PAGE) was done according to a modified Shägger protocol³²⁰. Cell pellets from 200 mL of *Chlamydomonas* culture were resuspended in an extraction buffer (20 mM HEPES pH= 7.5, 20 mM KCl, 10% glycerol, 2x EDTA-free protease inhibitor mix (ROCHE)), and broken using a French press apparatus (at 6000 psi). The soluble fraction was collected after centrifugation at 267000 rcf at 4°C for 25 min and concentrated using Amicon Ultra 30 kDa cutoff centrifugation units (Millipore). Protein concentrations were measured colorimetrically by Bradford assay³²¹ using QuickStart Bradford Dye reagent (Bio-Rad). 70 μg of protein was loaded on MiniProtean 4-16% gradient gels (Invitrogen). Migration was undertaken at 4°C at constant voltage of 60 V for 1h than 120 V. Native gels used for immunoblotting were first incubated 1 hour in 2% SDS, 0.67% β -mercaptoethanol prior to their transfer on nitrocellulose membranes (2h30 at $1 \text{mA} \times \text{cm}^{-2}$).

For immunoblot analysis, proteins were transferred onto nitrocellulose membranes, and the membranes were blocked with 5% (w/v) skim milk in a phosphate-buffered saline (PBS) solution plus 0.1% (w/v) Tween 20 (Sigma-Aldrich). The target proteins were immunodecorated with primary antibodies and then incubated with horseradish peroxidase-conjugated anti-rabbit IgG antibodies (Catalogue number: W4011, Promega) at a dilution ratio of 1:20,000. Primary antibodies used were directed against Rubisco whole holoenzyme (kindly provided by Dr. Spencer Whitney, used at a dilution of 1:80000), Cpn60 α/β 1 (kind gift of Michael Schroda, used at a dilution of 1:2000). Antibodies directed against the PSI subunit PsaD and α -tubulin were purchased from Agrisera (catalogue number: AS09 461 and AS10 680) and used respectively at a dilution of 1:40,000 and 1:10,000. Antibodies against cytochrome *f* (PetA; used at a dilution ratio of 1:100,000) were prepared using purified cytochrome *f* injected to rabbits. For RAF1 antibody production, recombinant *C. reinhardtii* RAF1 (Cre06.g308450) was expressed in *E. coli* using a codon-adapted synthetic cDNA (Genscript, Piscataway, NJ, USA). The protein was purified using GST-tag affinity and used directly as an antigen in rabbits (Genscript, Piscataway, NJ, USA). The resulting antiserum was used at a dilution of 1:30000. Immuno-reactive proteins were detected with Clarity One detection reagent (Biorad) and visualized using the ChemiDoc XRS+ System (Bio-Rad).

Acknowledgements

We thank Michael Schroda (University of Kaiserslautern, Germany) and Spencer M. Whitney (ANU Canberra, Australia) for kindly providing the anti-CPN60 and anti-Rubisco antibodies respectively.

We acknowledge basic support from Centre National de la Recherche Scientifique (CNRS) and Université Pierre et Marie Curie (UPMC) to UMR7141, and competitive funding from Labex Dynamo (ANR-11-LABX-0011-01). W.W. benefited from a doctoral support from ED515, Complexité du Vivant and Labex Dynamo (ANR-11-LABX-0011-01). We thank David Stern (BTI, Ithaca, NY) and PICS 5462 for support to K.W. during the initial stage of this project.

Figure legends

Fig. 1 LSU synthesis rate and accumulation in absence of its assembly partner

Characterisation of the Δ RBCS mutant. (A) Immunoblot showing protein level accumulation of Rubisco subunits in the Δ RBCS strain, using an antibody directed against whole Rubisco holoenzyme.

(B) ^{14}C pulse experiment showing the translation rate of LSU in the Δ RBCS as compared to the WT in upper panel, and mRNA accumulation in the lower panel, as probed by an *rbcL* and *psaB* probe, used as a loading control. In both panels, large subunit deletion strain Δ *rbcL* is used as a negative control.

Fig. 2 Swapping of *rbcL* regulatory sequences impairs the CES regulation.

(A Upper) photosynthetic phenotypes of 5'UTR*psaA*:*rbcL* transformants and corresponding Rubisco subunits' accumulation tested by a western blot analysis (A lower panel). In Δ RBCS;5'UTR*psaA*:*rbcL* LSU is accumulating to higher levels than in Δ RBCS (~8% WT; quantification not shown). TAP stands for Tris-Acetate-Phosphate medium, MIN is acetate-free, phototrophy-selective medium.

(B) ^{14}C pulse labeling experiment showing *rbcL* translation rate in 5'UTR*psaA*:*rbcL* background with and without small subunit compared with wild-type, Δ *rbcL* and Δ RBCS strains. Dashed line marks two removed lanes that were irrelevant to the figure.

Fig. 3 Expression of cytochrome *f* is inhibited in the absence of Rubisco small subunit.

Immunoblot using antibodies directed against the proteins depicted at the right, showing Rubisco and cytochrome *f* accumulation levels in the wild-type, Δ RBCS, Δ *rbcL* and 5'UTR*rbcL*:*petA* strains with and without SSU. *PsaD* accumulation is used as a loading control.

Fig. 4 CES regulation no longer occurs in the absence of LSU accumulation.

(A) ^{14}C labeling experiment showing chloroplast proteins synthesis rate of WT (T222+), $\Delta rbcL$, $\Delta RBCS$, LSU_{tr} , (transformants 1-3) and $\Delta RBCS;LSU_{tr}$ (1 and 4) strains. Dashed line marks two removed lanes that were irrelevant to the figure.

(B) Immunoblot depicting LSU and *cyt f* accumulation in representative transformants bearing a truncation within the *rbcL* gene, associated or not to the $\Delta RBCS$ mutation, and compared to the wild-type and $\Delta RBCS;5'UTRrbcL:petA$ strains.

Fig. 5 LSU assembly intermediates accumulate in the SSU-lacking strain

(A) Immunoblot with the antibody directed against LSU after Native PAGE analysis of soluble protein extracts from WT (diluted to 2% as not to obscure the gel), $\Delta rbcL$ and $\Delta RBCS$ strains. The migration of native molecular weight markers is indicated on the right. The position of Rubisco holoenzyme, as deduced from the WT signal, is indicated as well.

(B) Analysis of the second dimension on SDS-PAGE gel by immunodetection of proteins putatively involved in the complexes with LSU in $\Delta RBCS$ strain, using anti-LSU, anti-CPN60 and anti-Raf1 antibodies. Dashed lines drawn to help with the alignment. Red asterisks mark cross-contaminating signals of the anti-LSU antibody.

Fig. 6 LSU_2 mutations alter LSU accumulation and CES regulation

(A) Impairment in Rubisco accumulation is revealed by the absence of phototrophic growth in the LSU_2mut and $\Delta RBCS;LSU_2mut$ strains as probed by spot tests on acetate-free minimal media (MIN). Growth on TAP is shown as a control. The corresponding soluble LSU accumulation detected by immunoblot is shown.

(B) *rbcL* translation rate in LSU_2mut and $\Delta RBCS;LSU_2mut$ measured by short ^{14}C pulse labeling experiment and compared to $\Delta RBCS$ and WT.

(C) Immunoblot with the Rubisco antibody after Colorless Native PAGE analysis of soluble protein fractions of WT (note the dilution), $\Delta rbcL$, $\Delta RBCS$, LSU_2mut and $\Delta RBCS;LSU_2mut$ strains. Dashed line marks two removed lanes that are irrelevant to the figure.

Fig. 7 Disruption of LSU oligomerization alters LSU CES regulation

(A) Impairment in Rubisco accumulation is revealed by the absence of phototrophic growth in the LSU₈mut and ΔRBCS;LSU₈mut strains as probed by spot tests on acetate-free minimal media (MIN). Growth on TAP is provided as a control. The corresponding soluble LSU accumulation detected by immunoblot is shown.

(B) *rbcL* translation rate in LSU₈mut and ΔRBCS;LSU₈mut measured by short ¹⁴C pulse labeling experiment and compared to ΔRBCS and WT. Dashed line marks two removed lanes that were irrelevant to the figure.

(C) Immunoblot showing LSU and cytochrome *f* accumulation levels in the wild-type (WT), ΔRBCS, ΔrbcL, LSU₈mut and ΔRBCS; LSU₈mut strains and in those latter three genetic contexts combined with the 5'UTR*rbcL*;*petA* reporter gene background. PsaD accumulation is provided as a loading control.

(D) Immunoblot with the Rubisco antibody after Colorless Native PAGE analysis of soluble protein fractions of WT (note the dilution), ΔrbcL, ΔRBCS;LSU₈mut and ΔRBCS;LSU₈mut strains. The dashed box marks the diffused band attributed to LS dimer.

Sup. Fig. 1 Anti-RAF1 antibody

Immunoblot showing reactivity of raised αRAF1 antibody. Dilutions of GST-tag purified RAF1 protein (85% purity) were compared to dilutions of whole cell extracts of WT (T222+) strain. The band migrating at around 75 kDa in the purified protein sample is RAF1-GST (50 kDa RAF1 + 25 kDa GST-tag). Right panel shows antibody reactivity after blotting of 2D native PAGE gel, where 70 μg ΔRBCS soluble extract were separated. Arrows indicate the position of RAF1full-length product.

Sup. Fig. 2 RAF1 accumulates in the absence of LSU

Immunoblot of a first dimension Native-PAGE of soluble extracts from ΔrbcL and ΔRBCS probed with anti-RAF1 antibody. It shows that in the absence of LSU, RAF1 accumulates in oligomeric form.

Sup. Fig. 3 Close-up of the mutated residues in LSU structure

Residues mutated in (A) LSU₂mut (E109A and R253A) and (B) LSU₈mut (A143W-R215A-D216A) are highlighted in *Chlamydomonas reinhardtii* LSU structure (PDB: 1R8).

Bibliography

- 1 Ellis, R. J. The most abundant protein in the world. *Trends in biochemical sciences* **4**, 241-244, (1979).
- 2 Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E. & Scott, S. S. Distinct form I, II, III, and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco evolution and structure/function relationships. *Journal of experimental botany* **59**, 1515-1524, (2008).
- 3 Andersson, I. & Backlund, A. Structure and function of Rubisco. *Plant physiology and biochemistry : PPB / Societe francaise de physiologie vegetale* **46**, 275-291, (2008).
- 4 Bracher, A., Whitney, S. M., Hartl, F. U. & Hayer-Hartl, M. Biogenesis and Metabolic Maintenance of Rubisco. *Annual review of plant biology*, (2017).
- 5 Morita, K., Hatanaka, T., Misoo, S. & Fukayama, H. Unusual Small Subunit That Is Not Expressed in Photosynthetic Cells Alters the Catalytic Properties of Rubisco in Rice. *Plant Physiology* **164**, 69-79, (2014).
- 6 Börner, T. The discovery of plastid-to-nucleus retrograde signaling—a personal perspective. *Protoplasma*, (2017).
- 7 Chan, K. X., Phua, S. Y., Crisp, P., McQuinn, R. & Pogson, B. J. Learning the Languages of the Chloroplast: Retrograde Signaling and Beyond. *Annual Review of Plant Biology* **67**, 25-53, (2016).
- 8 Jarvis, P. Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytologist* **179**, 257-285, (2008).
- 9 Grimm, R. *et al.* Postimport methylation of the small subunit of ribulose-1,5-bisphosphate carboxylase in chloroplasts. *FEBS Letters* **408**, 350-354, (1997).
- 10 Feiz, L. *et al.* A protein with an inactive pterin-4a-carbinolamine dehydratase domain is required for Rubisco biogenesis in plants. *The Plant journal : for cell and molecular biology*, (2014).
- 11 Woodson, J. D. & Chory, J. Coordination of gene expression between organellar and nuclear genomes. *Nature reviews. Genetics* **9**, 383-395, (2008).
- 12 Barkan, A. & Small, I. Pentatricopeptide repeat proteins in plants. *Annual review of plant biology* **65**, 415-442, (2014).
- 13 Hammani, K. *et al.* Helical repeats modular proteins are major players for organelle gene expression. *Biochimie* **100**, 141-150, (2014).
- 14 Johnson, X. *et al.* MRL1, a conserved Pentatricopeptide repeat protein, is required for stabilization of *rbcl* mRNA in *Chlamydomonas* and *Arabidopsis*. *Plant Cell* **22**, 234-248, (2010).
- 15 Brutnell, T. P., Sawers Rj Fau - Mant, A., Mant A Fau - Langdale, J. A. & Langdale, J. A. BUNDLE SHEATH DEFECTIVE2, a novel protein required for post-translational regulation of the *rbcl* gene of maize. (1999).
- 16 Doron, L., Segal, N., Gibori, H. & Shapira, M. The BSD2 ortholog in *Chlamydomonas reinhardtii* is a polysome-associated chaperone that co-migrates on sucrose gradients with the *rbcl* transcript encoding the Rubisco large subunit. *The Plant journal : for cell and molecular biology*, (2014).
- 17 Houtz, R. L., Magnani, R., Nayak, N. R. & Dirk, L. M. Co- and post-translational modifications in Rubisco: unanswered questions. *Journal of experimental botany* **59**, 1635-1645, (2008).
- 18 Onizuka, T. *et al.* The *rbcl* gene product promotes the production and assembly of ribulose-1,5-bisphosphate carboxylase/oxygenase of *Synechococcus* sp. PCC7002 in *Escherichia coli*. *Plant & cell physiology* **45**, 1390-1395, (2004).
- 19 Feiz, L. *et al.* Ribulose-1,5-bis-phosphate carboxylase/oxygenase accumulation factor1 is required for holoenzyme assembly in maize. *Plant Cell* **24**, 3435-3446, (2012).

- 20 Saschenbrecker, S. *et al.* Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. *Cell* **129**, 1189-1200, (2007).
- 21 Emlyn-Jones, D., Woodger, F. J., Price, G. D. & Whitney, S. M. RbcX can function as a rubisco chaperonin, but is non-essential in *Synechococcus* PCC7942. *Plant & cell physiology* **47**, 1630-1640, (2006).
- 22 Tarnawski, M., Gubernator, B., Kolesinski, P. & Szczepaniak, A. Heterologous expression and initial characterization of recombinant RbcX protein from *Thermosynechococcus elongatus* BP-1 and the role of RbcX in RuBisCO assembly. *Acta biochimica Polonica* **55**, 777-785, (2008).
- 23 Kolesinski, P. *et al.* Insights into eukaryotic Rubisco assembly - crystal structures of RbcX chaperones from *Arabidopsis thaliana*. *Biochimica et biophysica acta* **1830**, 2899-2906, (2013).
- 24 Kolesinski, P., Belusiak, I., Czarnocki-Cieciura, M. & Szczepaniak, A. Rubisco Accumulation Factor 1 from *Thermosynechococcus elongatus* participates in the final stages of ribulose-1,5-bisphosphate carboxylase/oxygenase assembly in *Escherichia coli* cells and in vitro. *The FEBS journal*, (2014).
- 25 Wostrikoff, K., Girard-Bascou, J., Wollman, F. A. & Choquet, Y. Biogenesis of PSI involves a cascade of translational autoregulation in the chloroplast of *Chlamydomonas*. *The EMBO Journal* **23**, 2696, (2004).
- 26 Kuras, R. & Wollman, F. A. The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga *Chlamydomonas reinhardtii*. (1994).
- 27 Minai, L., Wostrikoff, K., Wollman, F. A. & Choquet, Y. Chloroplast biogenesis of photosystem II cores involves a series of assembly-controlled steps that regulate translation. *Plant Cell* **18**, 159-175, (2006).
- 28 Choquet, Y. *et al.* Translation of cytochrome f is autoregulated through the 5' untranslated region of petA mRNA in *Chlamydomonas* chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 4380-4385, (1998).
- 29 Drapier, D., Rimbault, B., Vallon, O., Wollman, F.-A. & Choquet, Y. Intertwined translational regulations set uneven stoichiometry of chloroplast ATP synthase subunits. *The EMBO Journal* **26**, 3581-3591, (2007).
- 30 Khrebtukova, I. & Spreitzer, R. J. Elimination of the *Chlamydomonas* gene family that encodes the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 13689-13693, (1996).
- 31 Rodermel, S., Haley, J., Jiang, C. Z., Tsai, C. H. & Bogorad, L. A mechanism for intergenomic integration: abundance of ribulose bisphosphate carboxylase small-subunit protein influences the translation of the large-subunit mRNA. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 3881-3885, (1996).
- 32 Wostrikoff, K. & Stern, D. Rubisco large-subunit translation is autoregulated in response to its assembly state in tobacco chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 6466-6471, (2007).
- 33 Wostrikoff, K., Clark, A., Sato, S., Clemente, T. & Stern, D. Ectopic Expression of Rubisco Subunits in Maize Mesophyll Cells Does Not Overcome Barriers to Cell Type-Specific Accumulation. *Plant Physiology* **160**, 419-432, (2012).
- 34 Dent, R. M., Haglund, C. M., Chin, B. L., Kobayashi, M. C. & Niyogi, K. K. Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of *Chlamydomonas reinhardtii*. *Plant Physiol* **137**, 545-556, (2005).
- 35 Choquet, Y. & FA, W. *Chlamydomonas* sourcebook - The CES process. *Chlamydomonas sourcebook; Chapter 29; Volume 2*, (2007).
- 36 Liu, C. *et al.* Coupled chaperone action in folding and assembly of hexadecameric Rubisco. *Nature* **463**, 197-202, (2010).

- 37 Hauser, T. *et al.* Structure and mechanism of the Rubisco-assembly chaperone Raf1. *Nature structural & molecular biology*, (2015).
- 38 Roy, H., Bloom, M., Milos, P. & Monroe, M. Studies on the assembly of large subunits of ribulose biphosphate carboxylase in isolated pea chloroplasts. *The Journal of cell biology* **94**, 20-27, (1982).
- 39 Mühlbauer, S. K. & Eichacker, L. A. Light-dependent Formation of the Photosynthetic Proton Gradient Regulates Translation Elongation in Chloroplasts. *Journal of Biological Chemistry* **273**, 20935-20940, (1998).
- 40 Shapira, M. *et al.* Differential regulation of chloroplast gene expression in *Chlamydomonas reinhardtii* during photoacclimation: light stress transiently suppresses synthesis of the Rubisco LSU protein while enhancing synthesis of the PS II D1 protein. *Plant molecular biology* **33**, 1001-1001, (1997).
- 41 Yosef, I. *et al.* RNA binding activity of the ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit from *Chlamydomonas reinhardtii*. *The Journal of biological chemistry* **279**, 10148-10156, (2004).
- 42 Cohen, I., Knopf, J. A., Irihimovitch, V. & Shapira, M. A Proposed Mechanism for the Inhibitory Effects of Oxidative Stress on Rubisco Assembly and Its Subunit Expression. *Plant Physiology* **137**, 738-746, (2005).
- 43 Knopf, J. A. & Shapira, M. Degradation of Rubisco SSU during oxidative stress triggers aggregation of Rubisco particles in *Chlamydomonas reinhardtii*. (2005).
- 44 Zhan, Y. *et al.* Localized control of oxidized RNA. *Journal of cell science* **128**, 4210-4219, (2015).
- 45 Bracher, A., Starling-Windhof, A., Hartl, F. U. & Hayer-Hartl, M. Crystal structure of a chaperone-bound assembly intermediate of form I Rubisco. *Nature structural & molecular biology* **18**, 875-880, (2011).
- 46 Hubbs, A. & Roy, H. Synthesis and Assembly of Large Subunits into Ribulose Bisphosphate Carboxylase/Oxygenase in Chloroplast Extracts. *Plant Physiology* **100**, 272-281, (1992).
- 47 Hubbs, A. E. & Roy, H. Assembly of in vitro synthesized large subunits into ribulose-bisphosphate carboxylase/oxygenase. Formation and discharge of an L8-like species. *Journal of Biological Chemistry* **18**, 6, (1993).
- 48 Brutnell, T. P., Sawers, R. J., Mant, A. & Langdale, J. A. BUNDLE SHEATH DEFECTIVE2, a novel protein required for post-translational regulation of the *rbcl* gene of maize. *The Plant Cell* **11**, 849-864, (1999).
- 49 Boulouis, A. *et al.* The Nucleus-Encoded trans-Acting Factor MCA1 Plays a Critical Role in the Regulation of Cytochrome f Synthesis in *Chlamydomonas* Chloroplasts. *The Plant Cell* **23**, 333-349, (2011).
- 50 Harris, E. H. 2000 (Academic Press, 2009).
- 51 Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual*. (Cold spring harbor laboratory press, 1989).
- 52 Drapier, D., Girard-Bascou, J., Stern, D. B. & Wollman, F.-A. A dominant nuclear mutation in *Chlamydomonas* identifies a factor controlling chloroplast mRNA stability by acting on the coding region of the *atpA* transcript. (2002).
- 53 Fischer, N., Stampacchia, O., Redding, K. & Rochaix, J.-D. Selectable marker recycling in the chloroplast. *Molecular and General Genetics MGG* **251**, 373-380, (1996).
- 54 Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. (1970).
- 55 Wittig, I. & Schagger, H. Advantages and limitations of clear-native PAGE. *Proteomics* **5**, 4338-4346, (2005).
- 56 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. (1976).

Figure List

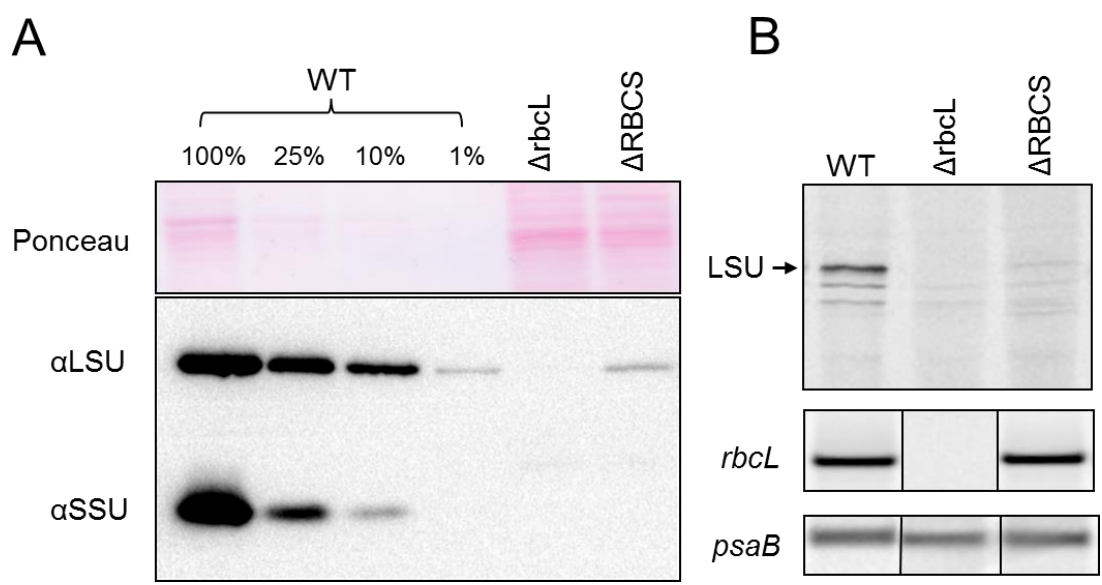


Fig. 1: LSU synthesis rate and accumulation in absence of its assembly partner

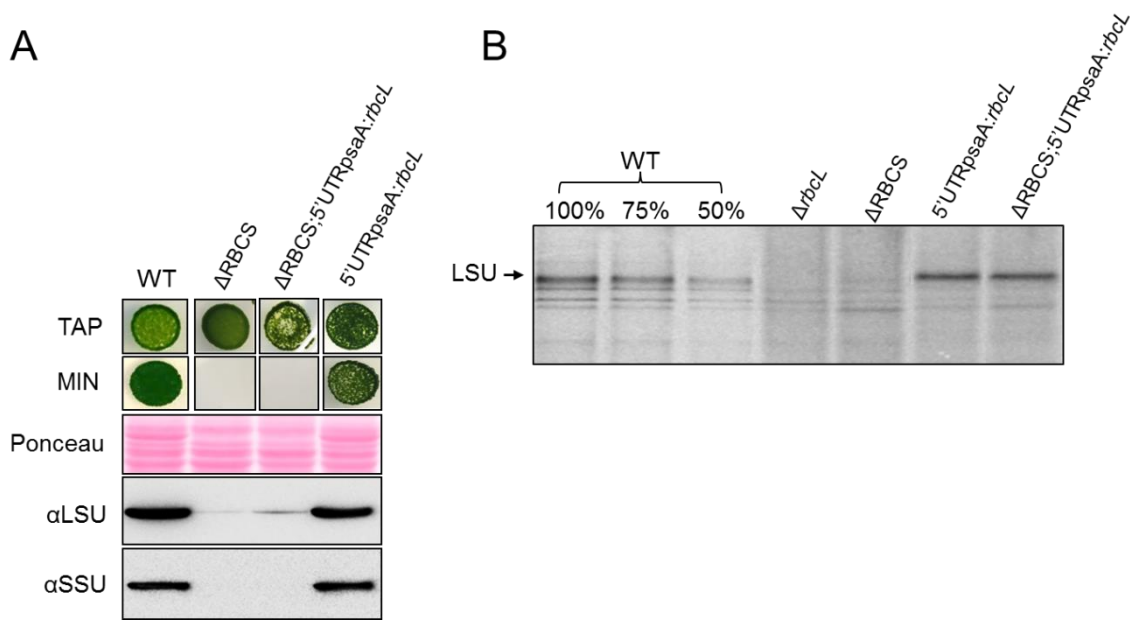


Fig. 2: Swapping of *rbcL* regulatory sequences impairs the CES regulation.

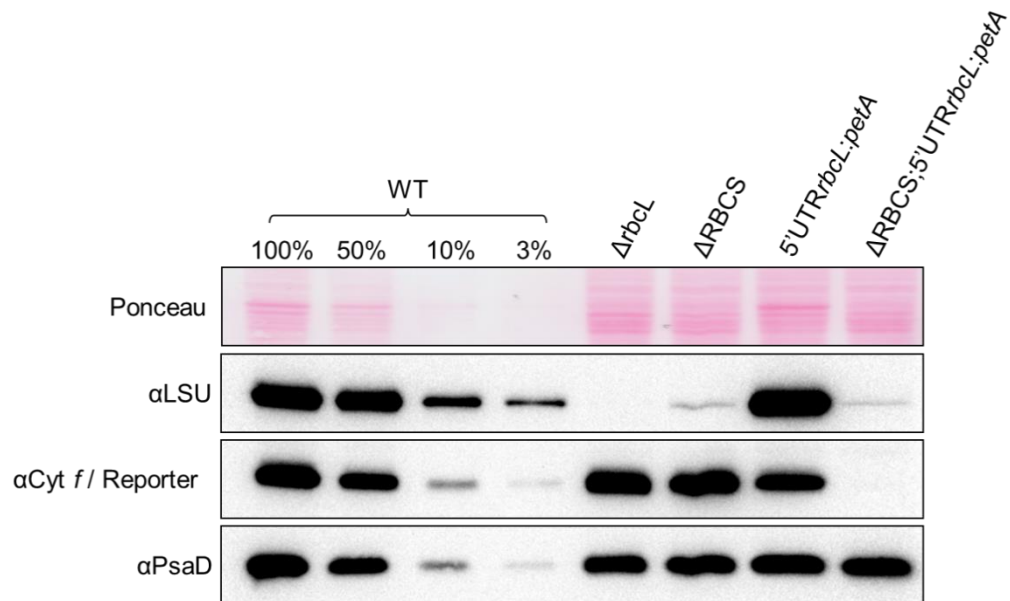


Fig. 3: Expression of cytochrome *f* is inhibited in the absence of Rubisco small

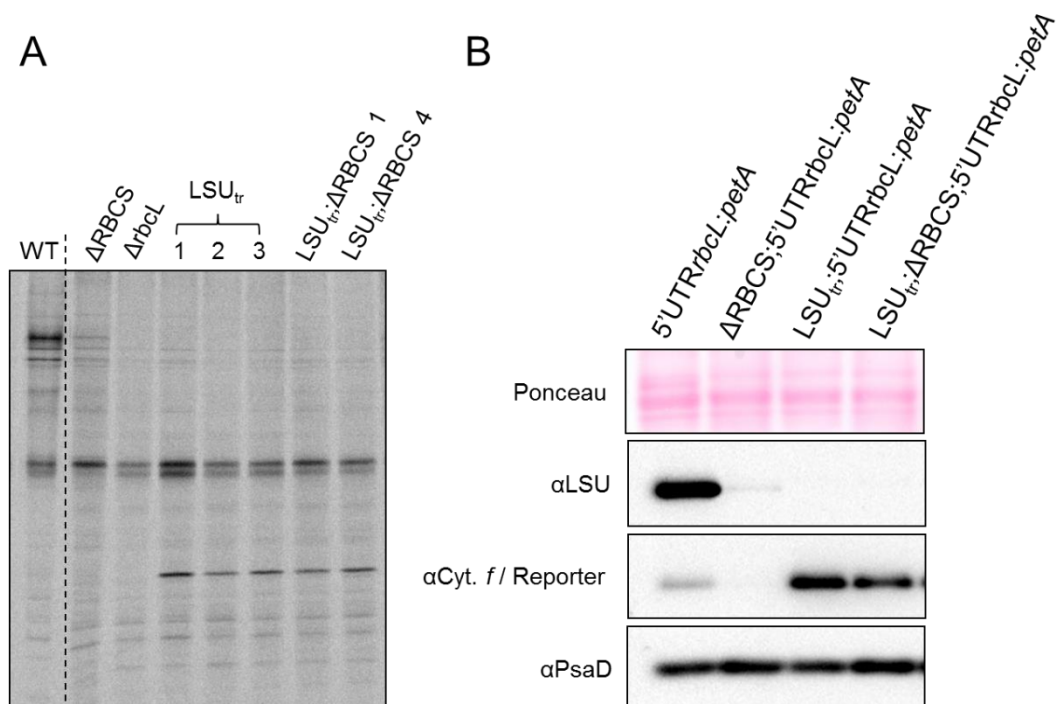


Fig. 4: CES regulation no longer occurs in the absence of LSU accumulation subunit.

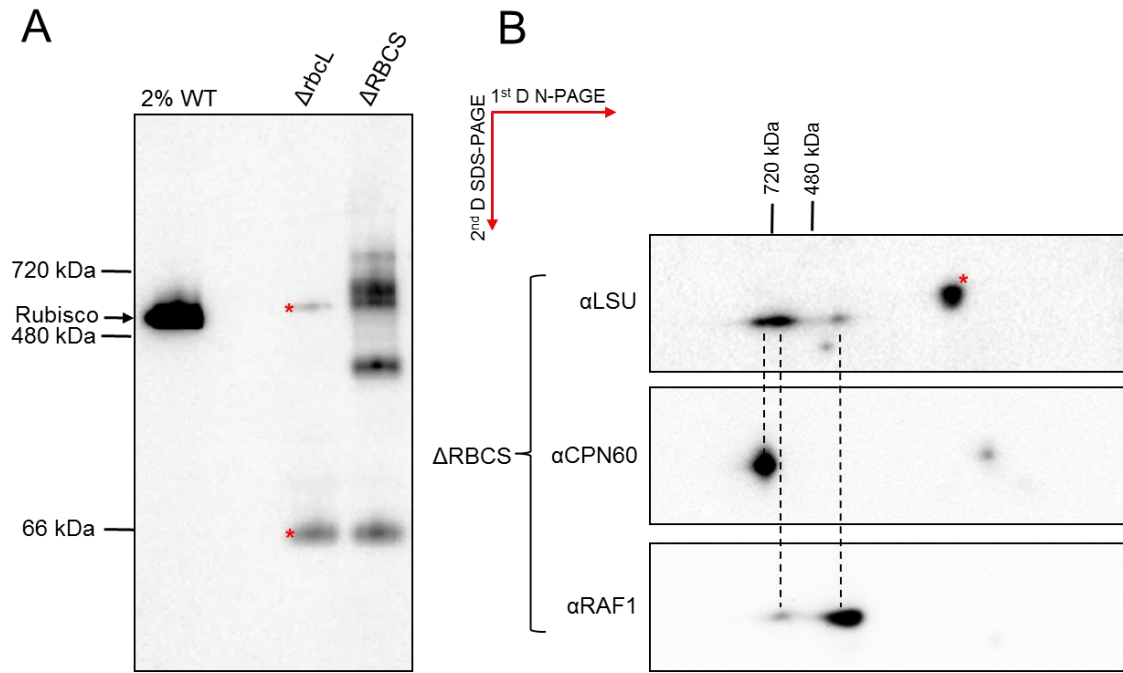


Fig. 5: LSU assembly intermediates accumulate in the SSU-lacking strain

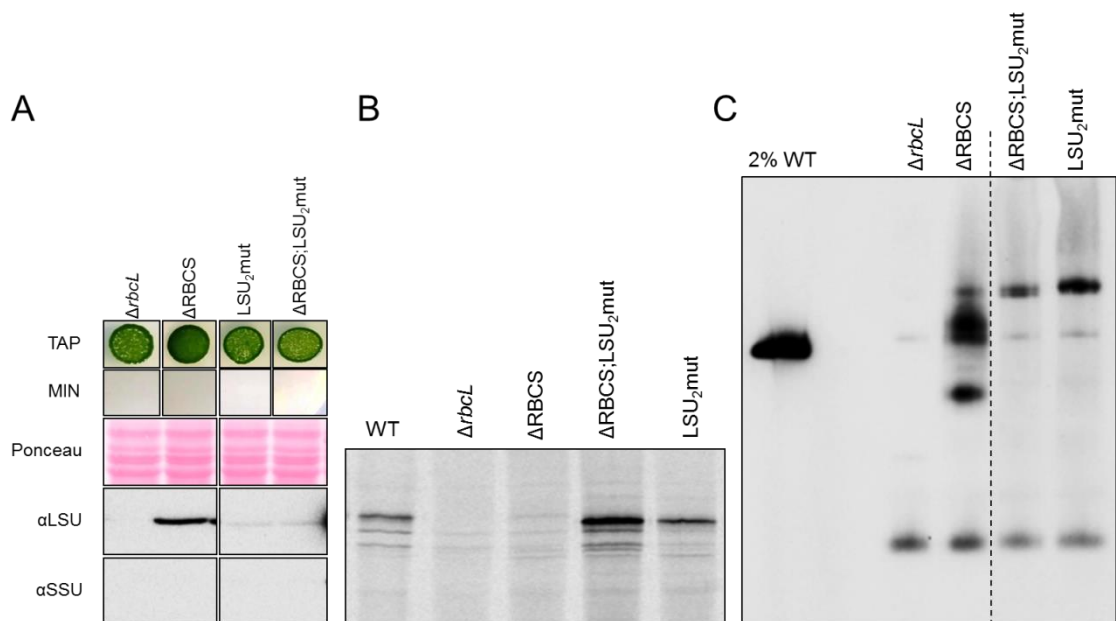


Fig. 6: LSU_2 mutations alter LSU accumulation and CES regulation

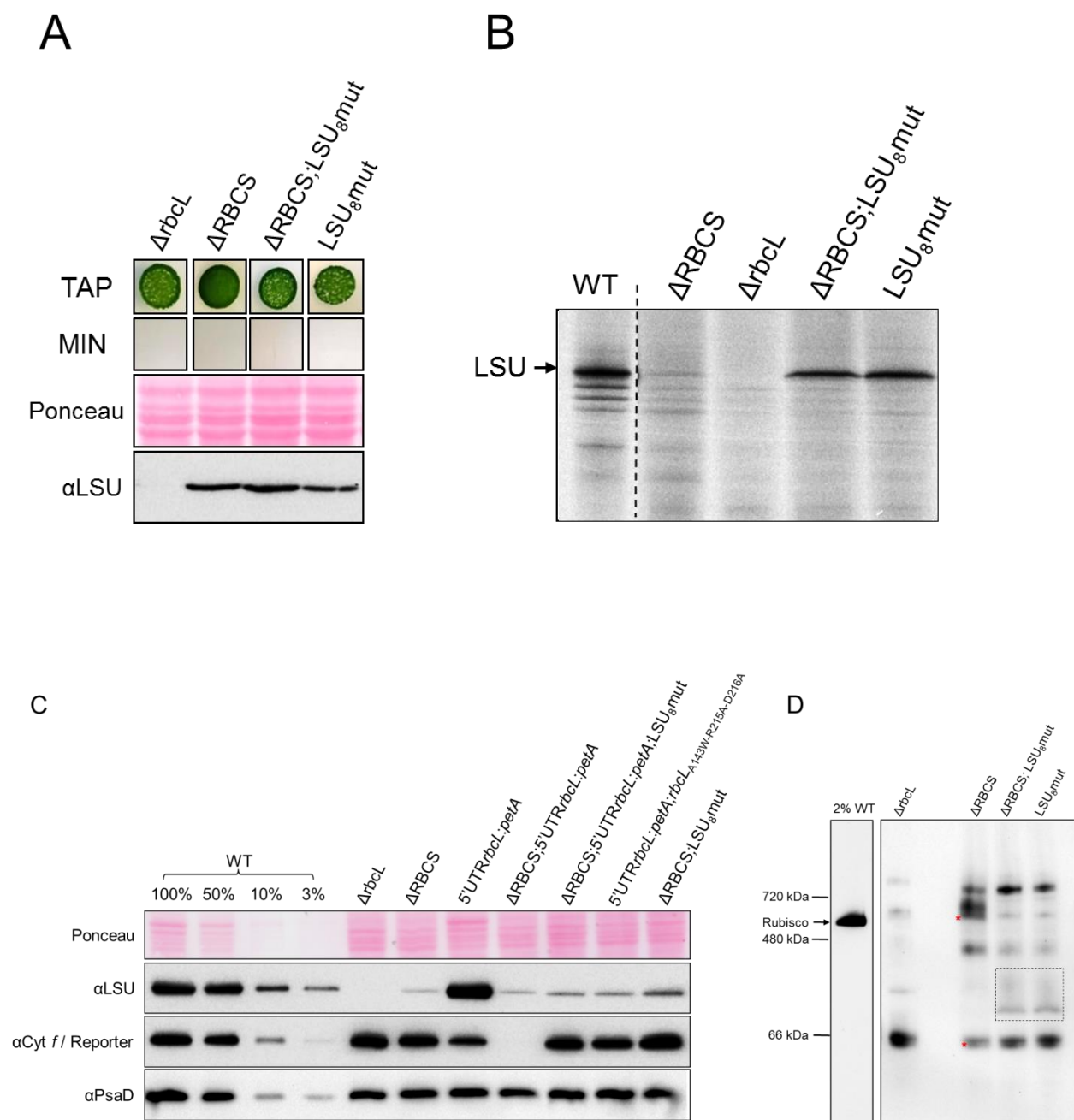
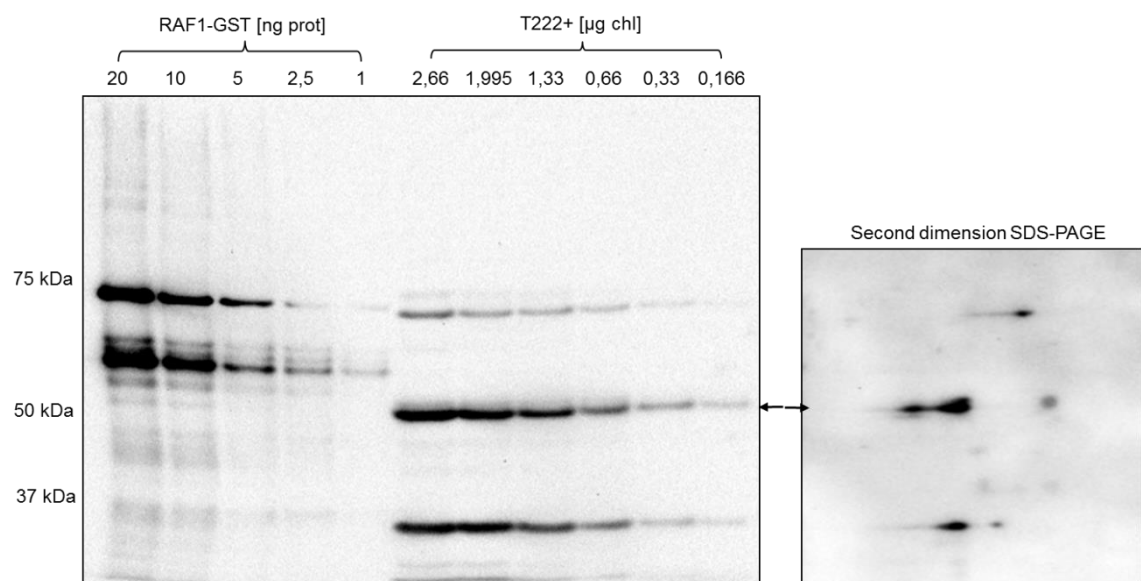
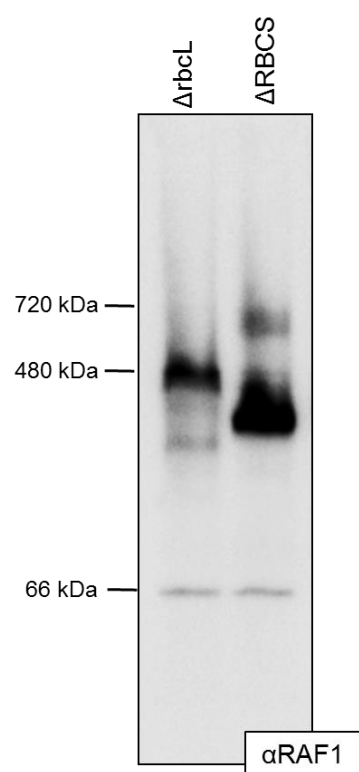


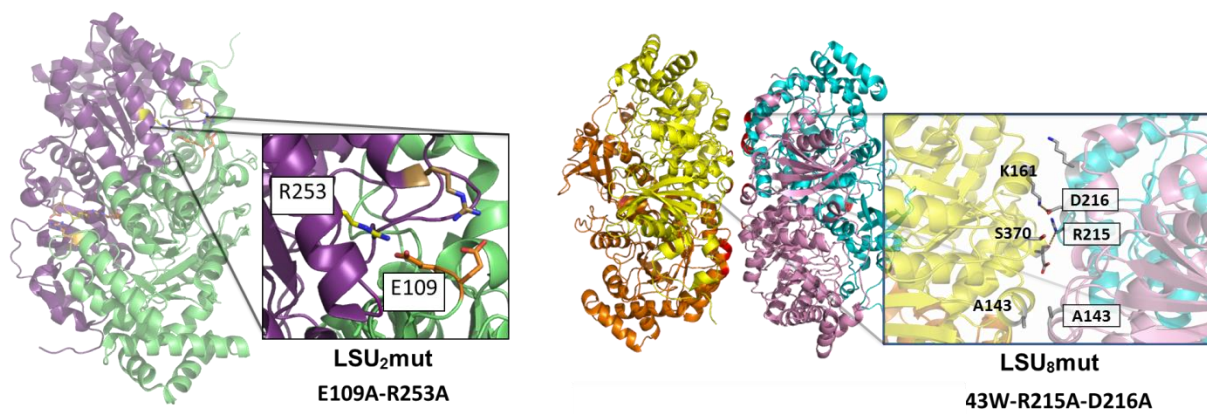
Fig. 7: Disruption of LSU oligomerization alters LSU CES regulation



Sup. Fig. 1: Anti-RAF1 antibody



Sup Fig. 2: RAF1 accumulates in the absence of LSU



Sup. Fig. 3 Close-up of the mutated residues in LSU structure

6. Chapter III: Further exploration of limitations in Rubisco biosynthesis – Supplementary results and discussion

In this part, further experiments designed to answer questions raised by the observations of CES-deregulated strains will be presented. They consist of pilot trials to create tools to get a closer look at the mechanism of LSU regulation.

6.1. Additional results for: “Rubisco LSU synthesis depends on its oligomerization state in *Chlamydomonas reinhardtii*”

6.1.1. Generation of $\Delta RBCS$;5'UTR $psaA$: $rbcL$ strain and its additional phenotype.

The low levels of LSU intermediates in the $\Delta RBCS$ strain at first led us to consider that their accumulation was limited because of the low translation rate of LSU. To generate a strain with high translation rates and lacking SSU we crossed 5'UTR $psaA$: $rbcL$ (aAR3.6+) strain with $\Delta RBCS$ (Cal13.1B-) with the aim to select SSU deficient progenies. Due to the inefficiency of the cross we did not obtain full tetrads but still were able to recover progenies with and without $RBCS$ expression. $\Delta RBCS$;5'UTR $psaA$: $rbcL$ cells were non-photosynthetic due to SSU absence (see Fig 2 in previous chapter). They did accumulate LSU to a higher level than $\Delta RBCS$ (~10% of the WT) (Fig. 6.1) most probably due to higher translation rate of LSU caused by the 5'UTR swapped to the one of $psaA$ (see Fig 2 in previous chapter).

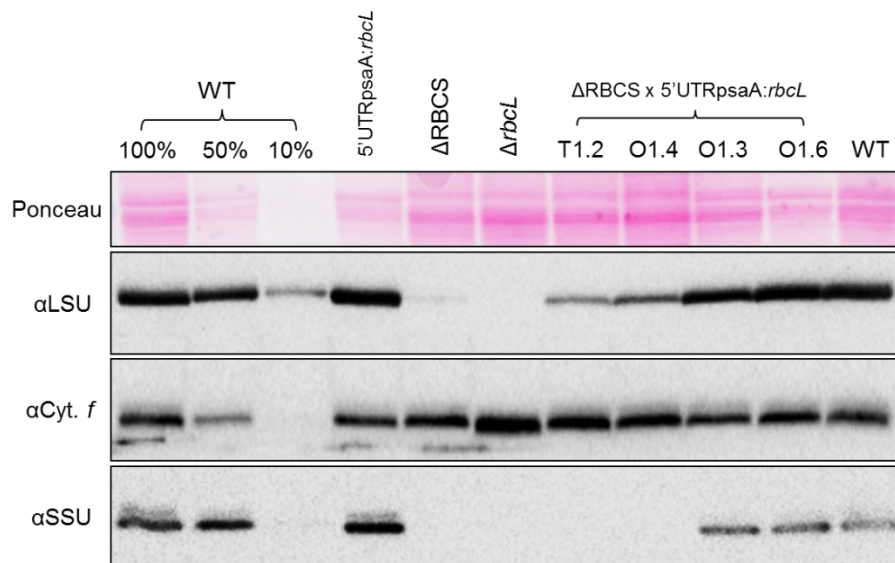


Fig. 6.1: LSU accumulation in the $\Delta RBCS$ x 5'UTR $psaA$: $rbcL$ cross

Immunoblot showing LSU and SSU accumulation in the progenies of the cross between 5'UTR $psaA$: $rbcL$ (aAR3.6+) with $\Delta RBCS$ (Cal13.5A-). Antibodies against LSU, SSU, and *cyt. f* as a loading control were used. Two independent progenies are shown for each $\Delta RBCS$ or WT genotypes.

In this context, Δ SSU-induced autoinhibition (CES) does not take place and LSU is translated at almost WT rates (Fig 2B in previous chapter). We reasoned that the increase in unassembled LSU may be reflected by a higher amount of the HMW-LSU-RAF1 complex, which we proposed to be the CES repressor. Such a strain over-accumulating the CES repressor would be very useful to approach its further characterization. We therefore tested the distribution of LSU between oligomerization intermediates, in Δ RBCS and Δ RBCS;5'UTRpsaA:*rbcL* as presented in Figure 6.2. Δ RBCS;5'UTRpsaA:*rbcL* accumulated the same intermediates as Δ RBCS. LMW- and HMW-LSU-RAF1 complexes accumulated to the same level as in Δ RBCS strain. CPN60-bound LSU level on the other hand were approximately 10 times higher in Δ RBCS;5'UTRpsaA:*rbcL*. We concluded that: i) higher translation rate of *rbcL* does not result in higher accumulation of the RAF1 intermediates ii) RAF1 is limiting for their accumulation and iii) CPN60 can bound significant amounts of free LSU.

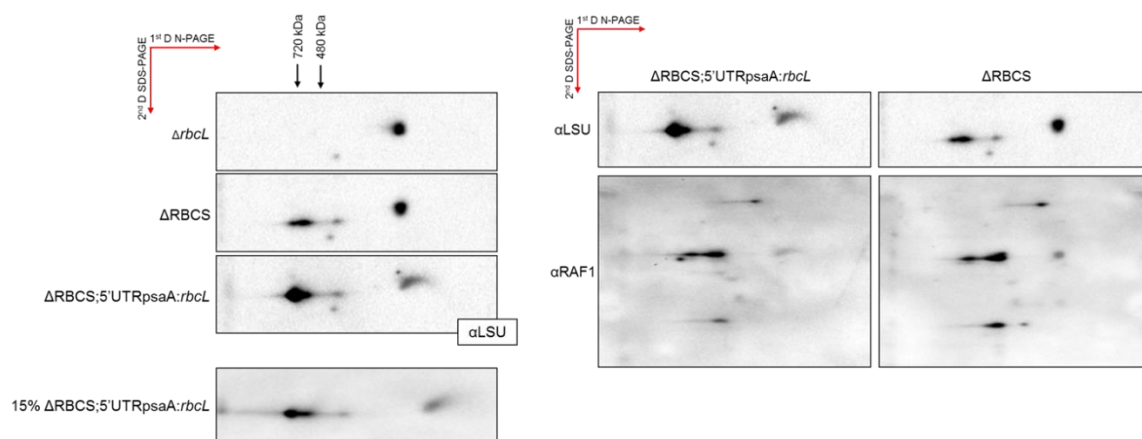


Fig. 6.2: LSU oligomerization states in Δ RBCS;5'UTRpsaA:*rbcL*

Immunoblot showing LSU and RAF1 accumulation in Δ RBCS;5'UTRpsaA:*rbcL* (Cal13aAR3.6) and Δ RBCS (Cal13.5A-) strains. Antibodies against LSU and RAF1 were used. The whole membrane is shown for RAF1 detection on the right panel.

6.1.2. Reporter gene accumulation is altered in Δ RBCS;5'UTRpsaA:*rbcL*.

Next, we wanted to verify if the increased synthesis of LSU has an effect on the reporter gene inhibition. To obtain Δ RBCS;5'UTRpsaA:*rbcL*;5'UTR*rbcL*:*petA* (Cal13aARRF) strains, we transformed the Δ RBCS;5'UTRpsaA:*rbcL* strain (aARCal13 1.6+) with the (pRFFFiK) plasmid. We expected the same, inhibition of the reporter gene as in Δ RBCS;5'UTR*rbcL*:*petA* background, resulting in no cytochrome *f* accumulation.

To our surprise, the reporter gene levels were found to be much higher in the transformants and reached around 50% of cytochrome *f* accumulation level observed in 5'UTR*rbcL*:*petA* strain. This could mean that *petA* translation is much less inhibited than in Δ RBCS;5'UTR*rbcL*:*petA*. We could not explain this phenotype as i) the amount of the “inhibitory” oligomer available for the regulation should be at least the same as in Δ RBCS and Δ RBCS;5'UTR*rbcL*:*petA* (Fig. 6.2) and ii) the number of *rbcL* 5'UTRs shared between reporter and LSU that could titrate the trans-factors (see Introduction: CES process) in the transformants should be the same as in Δ RBCS. Therefore this increase in reporter gene accumulation could not be due merely due to a different availability of limiting *rbcL* gene specific OTAFs. A question remains: does this relatively high accumulation of cytochrome *f* reporter still results from translation repression, or does this level result from impairment in the CES regulatory process? We decided then to directly compare this reporter accumulation to 5'UTR*psaA*:*rbcL*;5'UTR*rbcL*:*petA* progenies exhibiting RBCS expression in a larger number of strains. To this end, we back-crossed selected Δ RBCS;5'UTR*psaA*:*rbcL*;5'UTR*rbcL*:*petA* transformant (#6) with WT (WTS24-).

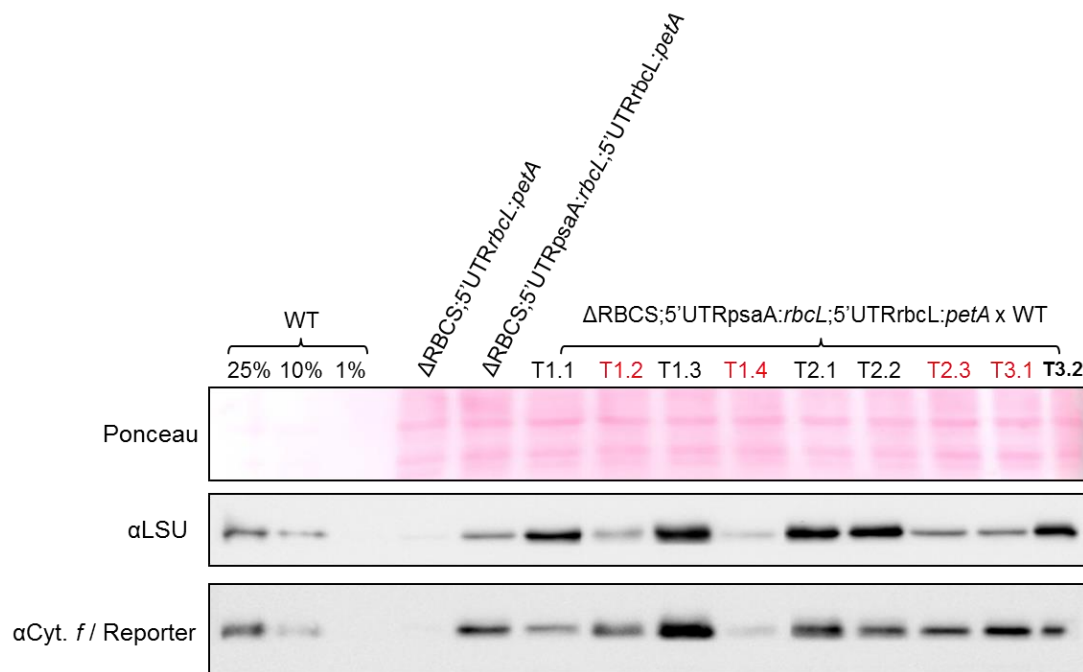


Fig. 6.3: Δ RBCS;5'UTR*psaA*:*rbcL*;5'UTR*rbcL*:*petA* cross

Immunoblot showing LSU and cyt. *f* accumulation in the progenies of the cross between Δ RBCS;5'UTR*psaA*:*rbcL*;5'UTR*rbcL*:*petA* and the WT (WTS 24-). Antibodies against LSU and cyt. *f* were used. Ponceau stain provided as a loading control. Nomenclature: T1.2 – second progeny of tetrad 1. Strains with the Δ RBCS genotype are shown in red.

We obtained one complete tetrad (plus additional progenies) and were surprised again to see that accumulation of cyt. *f* could be observed in all progenies and varied significantly between strains from almost WT levels to around 10% of WT (Fig. 6.3). We cannot explain this phenomenon differently than by the delicate balance between the amounts of inhibitor and its target that varies between the strains. The amount of *rbcL* and *petA* transcripts as well as their translation rates in 5'UTR ψ saA:*rbcL* and 5'UTR*rbcL*:*petA* contexts must be different from WT due to non-native trans-factors, non-native CDS (that can influence translation efficiency and folding of the 5'UTR) and even different translation modes and translation localization (different translation initiation modes, cyt. *f* is co-translationally inserted into the membrane etc.). Slight differences in the availability and ratios of LSU and SSU can influence the CES regulation and the accumulation of Rubisco and reporter proteins. The situation could be especially dynamic in the 5'UTR ψ saA:*rbcL* context where *rbcL* and RBCS expression do not respond to the same environmental cues, this way destabilizing the system. Clearly, a more throughout-out analysis is needed to elucidate the regulatory crosstalk in this complicated background.

6.1.3. What happens with LSU when Rubisco assembly is perturbed?

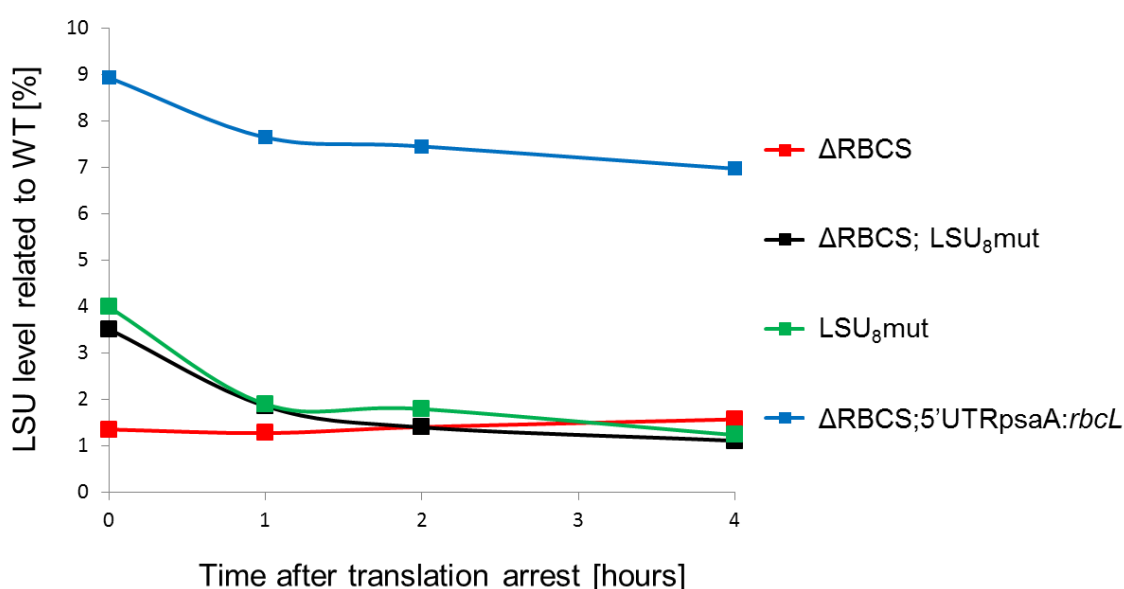
Δ RBCS, LSU₈mut as well as Δ RBCS;5'UTR ψ saA:*rbcL* strains were shown to accumulate approximately 1% and 9% of WT LSU. It is obvious (also from the results) that the mutations we have introduced must have had an effect on the balance between rates of translation, assembly and degradation of LSU. Knowing that, we decided to test the fate of LSU in the situation where its synthesis and oligomerization are perturbed.

6.1.4. Stability of LSU

In a first assay we used the Δ RBCS, LSU₈mut, Δ RBCS; LSU₈mut, and Δ RBCS;5'UTR ψ saA:*rbcL* to test if the half-life ($t_{1/2}$) of LSU is influenced by the mutations and deregulation of the CES process. We conducted an immunochase experiment treating the exponentially growing cells with inhibitors of chloroplast translation (Lincomycin) and monitoring LSU accumulation over 4 hours. As shown on the Fig. 6.4, LSU in Δ RBCS is rather stable. On the other hand, in LSU₈mut, Δ RBCS; LSU₈mut, and Δ RBCS;5'UTR ψ saA:*rbcL* it undergoes partial degradation in the first hour and is more stable in the remaining time of the experiment. Despite the degradation process taking place, LSU half-life is twice higher in Δ RBCS;5'UTR ψ saA:*rbcL* than in LSU₈mut background (Fig 6.4 Bottom).

6.1.5. Fractionation.

Several studies suggest that LSU could be found in different cell locations. Indeed, it is known that in normal growth conditions (ambient CO₂), LSU is mainly present in the pyrenoid fraction. Neosynthesized LSU has also been found associated either to thylakoids, or to a biogenenic membranes, close to the pyrenoid⁸². We wondered in here if LSU may be differentially distributed in the cell depending on the CES context. At the same time, LSU was suggested to participate in formation of stress granules (SG) that aggregate RNA to protect it from oxidative stress, those granules would be dense and resistant bodies whose presence was suggested in the *Chlamydomonas* strains lacking SSU³¹³.



LSU degradation in the first 1h

ΔRBCS; 5'UTRpsaA:rbcL = 16,5%; Half-life = 90,5 min

ΔRBCS; LSU₈mut = LSU₈mut = 54,9%; Half-life = 28,7min

Fig. 6.4: LSU degradation in the assembly mutants

Plot showing changes of the relative level of LSU accumulation (% of WT) after addition of an inhibitor of chloroplastic translation in LSU₈mut, ΔRBCS; LSU₈mut, and ΔRBCS; 5'UTRpsaA:rbcL mutants as compared to ΔRBCS. LSU half-life during the first hour of the experiment is indicated in the box. The experiment was done on two biological replicates for each strain.

We decided to tackle those two problems in the same line of experiments. We used a modified protocol from³¹³ (see Methods) to discriminate LSU localization between chloroplast fractions. French press lysates were ultra-centrifuged and separated into:

supernatant (S) and pellet (P). Half of the resuspended pellet was treated with 2% Triton 100 for 15 minutes and centrifuged to produce Triton soluble (supernatant: TS) and Triton insoluble (pellet: TI) fractions. Fractions obtained for WT (T222+), Δ RBCS, LSU₈mut, Δ RBCS; LSU₈mut and Δ RBCS;5'UTRpsaA:*rbcL* strains were quantified and adjusted to represent the same cell number and dilutions between soluble and pellet fractions were made to ensure proper loading. They were separated on SDS-PAGE gels and immunoblotted. Fig 6.5 shows the immunoblot probing LSU localization as compared to fraction markers: CPN60 α / β 1, Tubulin used as markers of soluble proteins and cyt.f, LHCII as markers of membrane proteins. We observe a relatively small fraction of Rubisco found in the ultracentrifugation pellet fraction (P and TS) in WT, Δ RBCS and Δ RBCS;5'UTRpsaA:*rbcL*. We attribute this signal to a contamination from soluble Rubisco, which would result from two reasons: a high abundance of Rubisco in WT and secondly, the formation of membrane vesicles entrapping stromal fraction during the preparation (as it is mirrored by the distribution of other soluble proteins such as CPN60 and tubulin). In the WT, most of LSU and SSU are bound in the soluble holoenzyme. We observe that this contaminating Rubisco in the pellet can be fully solubilized by 2% Triton 100 as no signal of LSU or SSU was detected in TI fraction. In the Δ RBCS mutant, where LSU accumulated to 1-2% of WT level, a similar pattern of LSU distribution could be observed, with most of the signal coming from soluble LSU. No SSU signal was detected, thus the soluble LSU in Δ RBCS does not result from traces of remaining holoenzyme and does not represent the same population as in WT (Note the exposition differences between WT and the mutants to account for differences in LSU abundancies). The same picture is seen for Δ RBCS; 5'UTRpsaA:*rbcL* strain that, as expected, does not accumulate SSU and for whom most of (10% of WT) LSU is found in the soluble fraction. Contaminating, pellet-found LSU is present in all the strains pointing towards some methodological artefact. On the other hand, ~80% and 40% of total LSU could be found in the pellet fraction of LSU₈mut and Δ RBCS; LSU₈mut strains respectively. Interestingly, a major part of this population was insensitive to Triton 100 solubilization. Moreover, almost all residual SSU in the LSU₈mut strain was associated with this form of LSU. Altogether this suggests the massive presence of LSU aggregates in this strain.

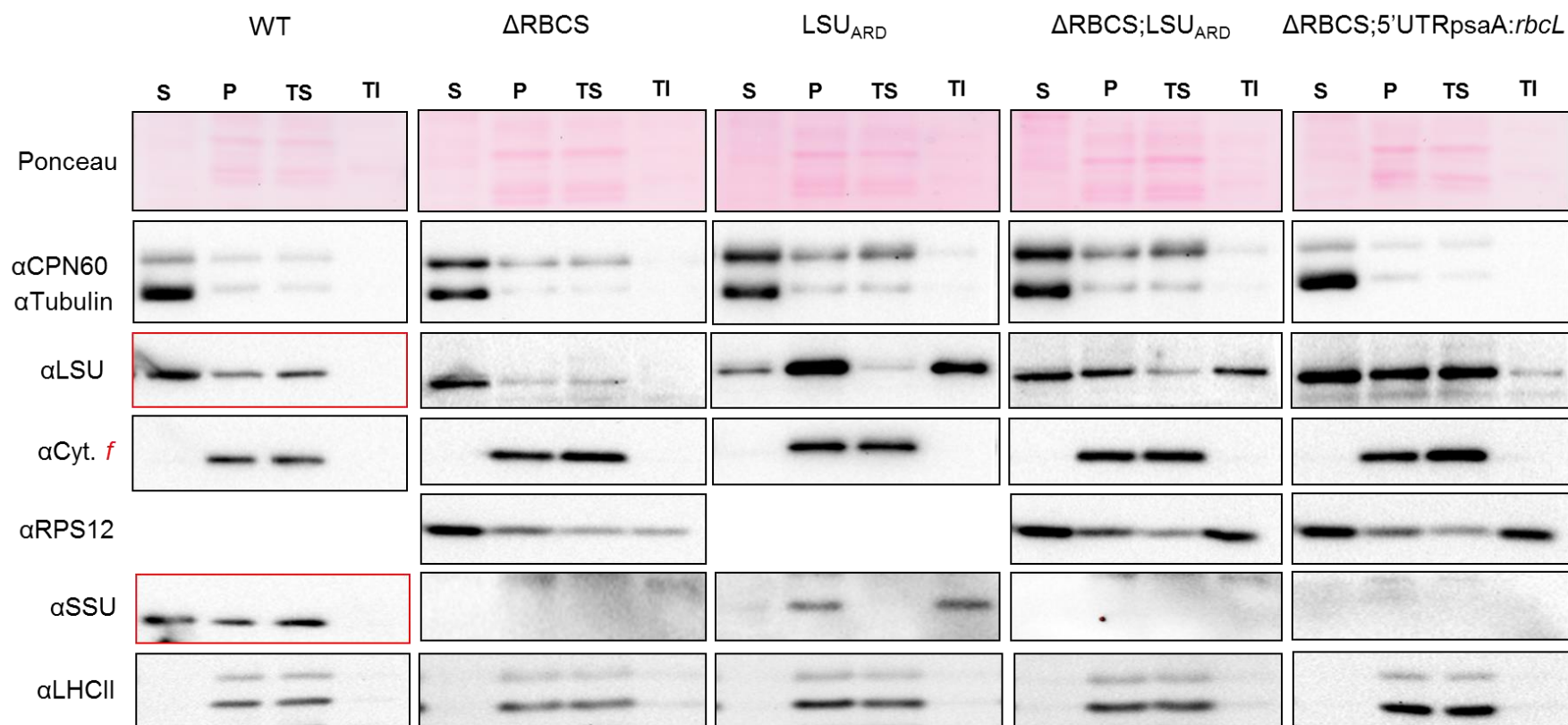


Fig. 6.5: LSU distribution between the cell fractions

Immunoblot showing accumulation of Rubisco subunits LSU and LSU in the soluble (S), pellet (P), Triton 100-soluble pellet (TS) and Triton 100-insoluble pellet (TI) fractions of WT (T222+), ΔRBCS (Cal13.1B-), LSU₈mut, ΔRBCS; LSU₈mut and ΔRBCS;5'UTR_{psaA}:*rbcL* (Cal13aAR3.6) strains. Antibodies against LSU, SSU, CPN60α and Tubulin (as soluble protein controls) cyt. *f* and LHCII (as membrane proteins control) and RPS12 were used. The exposition of WT LSU and SSU blots (red contour) was shorter than for the other strains. The experiment was done on two biological replicates for each strain.

6.1.6. EPYC1 accumulation is decreased in Rubisco deficient strains.

To follow this observation we verified if the over-accumulation of LSU in the pellet fraction is not linked to a mal-formation of the pyrenoid caused by premature interactions with LSU and EPYC1 protein. EPYC1 (before LCI5) has been reported to be essential for the pyrenoid formation in *Chlamydomonas* where it scaffolds dense Rubisco lattice which is the major part of this structure³²². Thanks to a kind gift of Martin Jonikas (Princeton, USA) we could use the α EPYC1 antibody to verify its behavior in LSU-altered mutants. Sup. Fig. 6.6 shows a test immunoblot against LSU and EPYC1 of whole cell extracts of two Δ RBCS strains (CAL13.1B-, Cal13.5A+), Δ RBCS;5'UTRpsaA:*rbcL*, LSU₈mut, LSU₂mut (*rbcL*_{E109A-R253A}) and two Δ *rbcL* mutants (Δ RT1.3, Δ RbcL1.7.5) compared to WT (T222+) and *epyc1*. In all the mutants EPYC1 accumulation is strongly diminished.

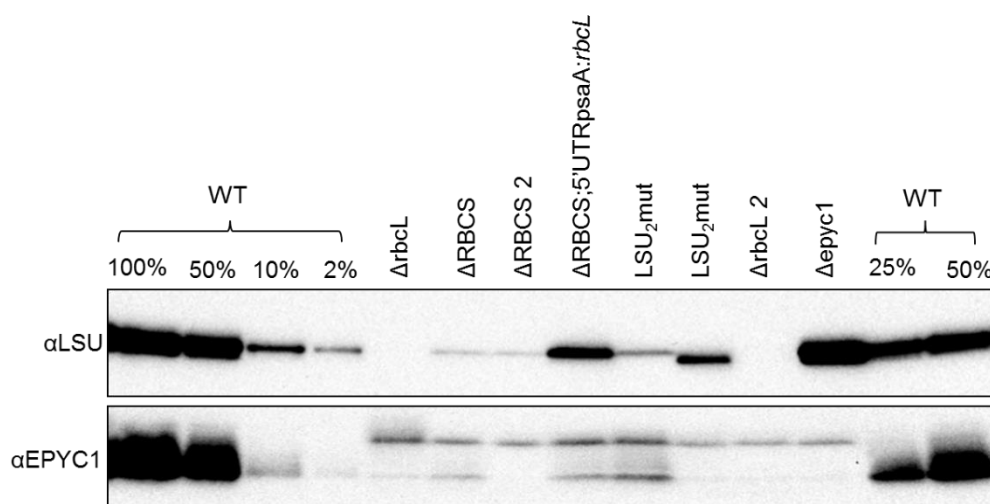


Fig. 6.6: Accumulation of EPYC1 in the mutants altered in LSU biosynthesis

Immunoblot of whole cell extracts of WT (T222+), Δ epyc1 and Rubisco biosynthesis mutants (details in the text). Antibodies against LSU and EPYC1 were used. Red asterisk marks the cross-reacting signal of α EPYC1 antibody.

Quantification of LSU and RAF1 levels. We have used purified RAF1-GST and Rubisco LSU standard (Agrisera) to calculate the abundance of those two proteins in the cell. Dilutions of the standards served to obtain the standard curve of antibody reactivity. It was used to quantify LSU and RAF1 accumulation levels in the dilutions of WT whole cell extract sample. Fig. 6.7 and 6.8 show immunoblot of the respective samples and the calculated values for both proteins. For the quantification a constant number of 1,4 μ g of chlorophyll per 1 million cells was used. Additionally, Fig 6.7 gives the approximate

numbers for the amount of LSU oligomers in Δ RBCS strain, based on the relative proportions of the three LSU-associated complexes, and their respective contributions in term of LSU units. Indeed, from densitometric analysis of native blots, LSU-CPN60 population would contribute to 8% of total LSU signal and involve one LSU unit, the HMW-LSU complex would represent 80% of the signal and contribute to 8 LSU units, and last, the remaining LMW-LSU complex was estimated to account for 12% of the signal and involve 2 LSU units. The molecular weights used for the calculations are of 52kDa for LSU and 50 kDa for RAF1.

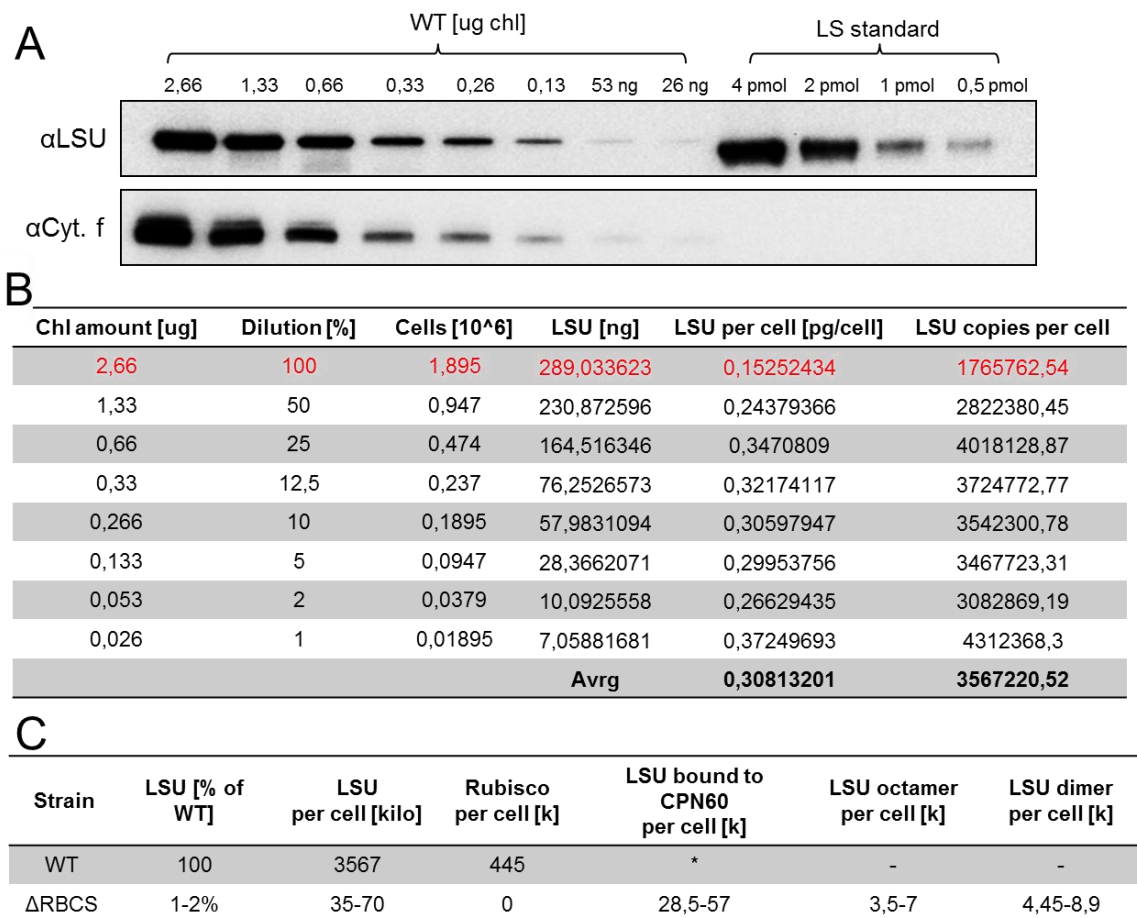


Fig. 6.7: LSU quantification

A - Immunoblot showing reactivity of α Rubisco antibody. Dilutions of whole cell extracts of WT (T222+) strain were compared to LSU standard (Agrisera) dilutions.

B - LSU quantity in WT dilutions calculated according to the standard curve done on the standard protein.

C - Approximate amounts of LSU oligomers in Δ RBCS strain. The approximation that 1,4 μ g chl corresponds to 10^6 cells was used for the calculations. LSU molecular mass used: 52 kDa. Results are rounded up. See description in the text.

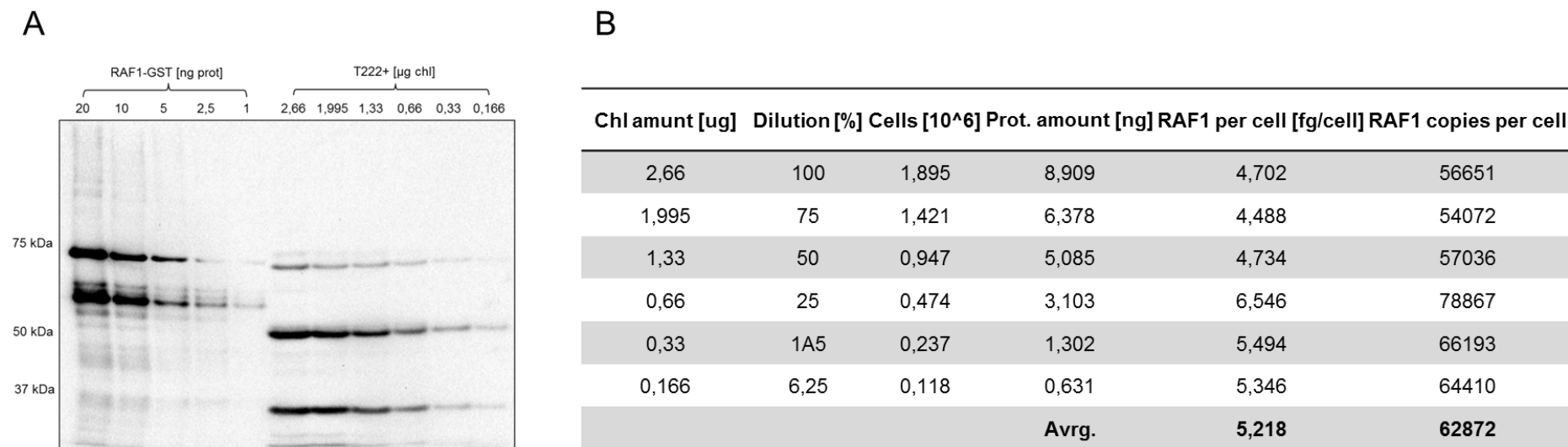


Fig. 6.8: RAF1 quantification

Immunoblot showing reactivity of raised α RAF1 antibody. Dilutions of GST-tag purified RAF1 protein (85% purity) were compared to dilutions of whole cell extracts of WT (T222+) strain. Band migrating at around 75 kDa in the purified protein sample is RAF1-GST (50 kDa RAF1 + 25 kDa GST-tag).

B - RAF1 quantity in WT dilutions calculated according to the standard curve done on purified protein. The approximation that 1,4 μ g chl corresponds to 10^6 cells was used for the calculations. RAF1 molecular mass used: 50 kDa. Results are rounded up. See description in the text.

6.2. Discussion

6.2.1. LSU is relatively stable when chaperone-bound

LSU immunochase demonstrated that in the CES-inhibited situation (Δ RBCS) LSU is stable during 4h after translation arrest. This was expected from LSU CES behavior, where unassembled LSU mediates the translation repression. As such, the inhibitory LSU intermediate should theoretically be stable. The same should be true for LMW-LSU₂-RAF1 complex due to stabilizing effect of the chaperone. On the other hand, Δ RBCS;5'UTRpsaA:*rbcL*, and LSU₈mut background showed a faster degradation of LSU in the first hour of the treatment ~90 min and ~30 min respectively (which is still much longer than the half-life of unassembled SSU: ~15 min). The significantly faster (3x) degradation rate of LSU₈mut compared to Δ RBCS;5'UTRpsaA:*rbcL*, although both display similar translation rates, could be attributed to the destabilizing effect of the ARD mutations that prevent oligomerization of LSU. From the native gel analysis presented in chapter 2, it is likely that this difference corresponds mainly to the fast degradation of RAF1-unbound LSU dimers and inability to form an octameric LSU complex. On the other hand, the different stability of unassembled LSU in Δ RBCS;5'UTRpsaA:*rbcL* was more unexpected. As native LSU is produced, it seems unlikely that the HMW-LSU₈ complex or the LMW-LSU dimer exhibit different stability rates in Δ RBCS;5'UTRpsaA:*rbcL* versus Δ RBCS, unless their composition (associated proteins or RNA) is not the same. Therefore we may hypothesize that LSU leaks out from the CPN60-LSU and gets degraded because the chaperones' pool is exhausted. Nevertheless, this does not exclude that the CPN60 *per se* has a stabilizing effect, which probably accounts for the relatively high residual steady state of LSU in Δ RBCS;5'UTRpsaA:*rbcL* (see below). At the same time one could think that in Δ RBCS;5'UTRpsaA:*rbcL* the HMW-LSU inhibitor- because of the lack of its putative target (no 5'UTR of *rbcL*) has a different stability than in Δ RBCS. More experiments need to be done to exclude measurements errors caused by low starting level of the LSU in the tested strains. Chase analysis in native conditions should allow to better understand the differences in stability of respective LSU oligomers between the strains (e.g. between Δ RBCS and Δ RBCS;5'UTRpsaA:*rbcL*).

6.2.2. Limiting steps of LSU assembly

In the Δ RBCS;5'UTRpsaA:*rbcL* strain, *rbcL* is translated at rates similar to WT (slightly lower) due to its swapped 5'UTR that prevents CES regulation triggered by SSU

absence. Because of that LSU accumulates to higher amounts than in $\Delta RBCS$. Yet, visibly, the level of free LSU cannot exceed 10% of WT LSU accumulation level. This increase in LSU (compared to $\Delta RBCS$) is found to be associated with the CPN60 complex (Fig. 6.1). On the other hand, RAF1-associated intermediates do not change in abundance, nor does RAF1 accumulation (Fig. 6.2B). All in all, the amount of soluble LSU accumulating in $\Delta RBCS;5'UTRpsaA:rbcL$ does not correspond to its rate of translation (Fig. 2 in previous chapter) which indicates increased proteolysis of unbound LSU or early aggregation of unfolded LSU. At the same time, we have shown that a significant part of LSU is stable in this background during 4 hours after chloroplast translation arrest (Fig. 6.4), which suggest that it can be stabilized by CPN60 and RAF1. All in all, our observations suggest that: i) accumulation of unassembled LSU is limited by the auxiliary factors and ii) about 10% of WT LSU is enough to sequester all available chaperones and iii) LSU probably does not accumulate in the soluble form when unbound by auxiliary proteins.

Interestingly, despite similar rate of *rbcL* translation (Fig. 2 and 6 in previous chapter), levels of CPN60-bound LSU in $\Delta RBCS;5'UTRpsaA:rbcL$ and LSU_8mut are drastically different. It is most probably due to mutations introduced (A143W-R215A-D216A) that change the folding capacity of LSU. It is possible that a quality control step before the chaperonin exist, which would distinguish WT (*5'UTRpsaA:rbcL*) and mutated (LSU_8mut) LSU sequence and direct it to the proteolytic degradation or aggregation. Indeed, the BSD2 orthologue in *Chlamydomonas* has a partial homology to DnaJ chaperone (part of Hsp70 chaperone complex) and was found co-migrating with *rbcL* transcript on polysomes. It might act upstream of CPN60 in LSU control. Altogether we hypothesize that the first bottleneck/quality check to LSU synthesis is upstream of the chaperonin complex.

6.2.3. Mutated LSU is directed to aggregates

We partially tested the fate of LSU in the LSU_8mut background strains (compared to other mutants and WT used previously) by following its repartition between: soluble, insoluble, Triton100-sensitive insoluble and Triton100-insensitive insoluble fractions. We found that in LSU_8mut and $\Delta RBCS; LSU_8mut$ a significant amount of LSU was found in the Triton- insoluble pellet fraction of cell extract. Those aggregates were extremely resistant and less sensitive to solubilization than thylakoid membranes (see cyt. *f* and LHCII partitioning). Together with the observation of the relatively low CPN60-LSU loading, it points to an early aggregation of LSU in LSU_8mut context. Whether it is a result of a specific mechanism of control (see above) or spontaneous aggregation is to

be determined. Interestingly, SSU co-fractionated with LSU in the LSU₈mut. This might suggest that, contrary to what was proposed before, SSU can to some extent interact with LSU at the early stage of assembly or that SSU gets co-precipitated together because of nucleation mechanism. The latter seems in fact to be promoted by SSU presence as we observe less precipitates in ΔRBCS; LSU₈mut. This is in line with the reports that SSU is necessary for Rubisco organization and cell localization²²⁷ but suggests that this process may start very early during the assembly or that *rbcL* translation is localized to specific chloroplast foci where SSU is directed. Higher EPYC1 accumulation in LSU₈mut may further substantiate this mechanism (Fig. 6.6). Similarly, co-localization of some ribosomal S12 with LSU aggregates may also be an effect of early co-translational aggregation. Whether the presence of S12 in these pellet insoluble fractions is really linked or not to LSU aggregation state, or a marker of some other unlinked biogenesis membranes, requires additional experimental support.

We could also observe a slightly higher ratio of pellet-fractionated LSU in ΔRBCS;5'UTRpsaA:*rbcL* compared to WT. But contrary to the LSU₈mut aggregates, it was almost completely solubilized by Triton100 suggesting that the precipitation/aggregation state in this strain is much smaller. This result allows concluding that the excess of unbound native (not mutated) LSU is, similarly to SSU, proteolytically degraded. Finally soluble LSU in LSU₈mut and ΔRBCS; LSU₈mut accumulated to almost the same amount as in ΔRBCS suggesting that the mutations do not prevent LSU folding but rather affect the translation-folding-assembly balance.

6.2.4. EPYC1 accumulates in coordination with Rubisco

While Rubisco accumulation is not perturbed in an *epyc1* mutant, as shown in³²², we demonstrate here that the converse is not true. EPYC1 accumulation of scaffold protein of the pyrenoid is strongly diminished in Rubisco defective mutants. Most probably, EPYC1 alone cannot form the pyrenoid lattice and is directed to proteolysis in absence of Rubisco (similarly to SSU in the absence of LSU). The residual EPYC1 levels vary significantly between the strains. One possible explanation is that because of its Rubisco surface binding properties, EPYC1 could miss-interact with LSU oligomers. Indeed almost no EPYC1 accumulate in LSU₂mut and Δ*rbcL*, correlating with the absence of LSU assembly intermediate accumulation in these strains. The interaction with LSU oligomers would lead to EPYC1 partial stabilization or precipitation. Most likely, this interaction would be reflected by EPYC1 fractionation within the triton-insoluble fraction, a hypothesis that awaits experimental confirmation.

6.2.5. Conclusions from LSU and RAF1 quantifications.

Rubisco is present in around 450 000 copies per cell which would approximately give two enzymes per light reactions chain. Because of this high number, the 1-2% amount of LSU found in SSU-lacking strain still is significant; however it is distributed between three oligomeric forms (Fig. 6.7 C). Assuming that the CES-inhibitor of the *rbcL* translation is the HMWC of LSU (octamer) complexed to RAF1, we can approximate that it would be present in 3,5-7 thousand copies per cell. It is probably close or a little bit more than the total number of *rbcL* transcripts (~4,500). These calculations match surprisingly well with the hypothesis that this HMW-LSU₈ repressor could bind directly (or not) to the mRNA. At the same time, the active RAF1 form is a dimer; from our calculation it is present in ~31,500 copies per cell which is in the same value range than the sum of LSU dimers it is stabilizing (partitioned in both the LSU dimer and octamer: LSU dimer plus 4 dimers per octamer = 18,500 - 35.000). This might suggest that RAF1 is the limiting factor for folded, soluble LSU accumulation. However, the much stronger signal observed for RAF1 in the LMWC suggests that the HMWC comprises less than 4 RAF1 dimers. It also suggests that the dimer would represent the more stable LSU-RAF1 intermediate, and that the HMWC would be stabilized by additional proteins which might be limiting for its abundance. Note also that RAF1 can accumulate without LSU in complexed form in $\Delta rbcL$. Before it was reported to accumulate free as trimers (~150 kDa)²⁶⁹, much less than what we observe. It is not excluded that it could interact transitorily with other chaperones and/or residual SSU as proposed²⁸⁵.

6.3. Additional comments

It has been demonstrated that free LSU (Rubisco-independent) can bind RNA *in vitro* through a RNA binding domain that normally is hidden within the structure of the enzyme²¹⁶. Following this observation different studies have proposed the moonlighting function of LSU in protecting RNA during oxidative stress. Large subunit would bind RNA and participate in the formation of ribonucleo-protein agglomerates called stress granules (cpSGs: for chloroplast stress granules) to prevent RNA oxidation and mistranslation. We have shown that, contrary to previous observations, in the absence of SSU, LSU does not form aggregates, nor does it precipitate. Even in the CES deregulated $\Delta RBCS;5'UTRpsaA:rbcL$ context, where theoretically much more LSU is available for the binding, the pellet-associated LSU is mostly Triton-sensitive, indicating that it originates from vesicles entrapping stroma that are preparation artefacts. The majority of unassembled LSU in the $\Delta RBCS$ strain was localized in the steady state in the soluble protein fraction (Fig. 6.5) and was associated with chaperones (Fig. 6.2). To

account for these different observations, we suggest either that the LSU agglomerates are specific to the oxidative stress (thereby not seen with our light regime) and/or that the hypothesis raised by Zhan and co-authors could result from the observations of LSU bound to *rbcL* transcript – a CES regulation which is visible due to the high number of *rbcL* mRNA present in the cell.

7. General discussion

After 3.5 Gyo of evolution Rubisco has still poor specificity towards carbon dioxide and very low catalytic efficiency. Despite its pivotal role for the cell's fitness, its amelioration throughout the years towards higher carboxylation rates seems to have been surprisingly feeble. One of the reasons for that might be the fact that Rubisco originated in the oxygen-free atmosphere with CO₂ levels at least 200 times higher than today, resulting in low pressure for CO₂ specificity. Changes in the atmosphere towards higher O₂:CO₂ ratios (*nomen omen*, caused by the Rubisco activity) costed Rubisco an oxygenation reaction that leads to unproductive photorespiration. It is unknown if the oxygenation reaction was an inherent characteristic of ancient Rubiscos, but nowadays photorespiratory pathway can diminish net carbon fixation by half. It is discussed whether it has some important role for cell's metabolism but it is clear that it has been evolutionary retained for 2 Gyo now. It might be than that this process is etched in Rubisco function and cannot be lost. Because of its presence, of the high discrepancy in oxygen and carbon dioxide amounts, and of structural similarities between the two gasses, the adaptative flexibility of Rubisco towards higher carboxylation rates is limited. The evolution of Rubisco small subunit could also be linked to changing ratios of atmospheric gases. SSU is present only in Type I Rubiscos of oxygenic phototrophs that nowadays live usually in carbon-limited niches. It has been shown that SSU is required for Rubisco localization to a carbon concentrating structure – the pyrenoid, which improves the local environment of the enzyme. Maybe then, SSU's first function was a priori in the organization of Rubiscos, a first step towards carbon concentrating mechanism of cyanobacteria and algae. The late apparition of SSU could be also a sign that improvements in LSU catalysis have already reached the plateau long ago. Indeed, tradeoffs between specificity (towards carbon dioxide and oxygen) and catalysis have been proposed, despite not being high, to be already almost perfectly tuned to the nowadays atmosphere and cannot progress any further³²³. In this light, improvements in the local CO₂ availability by carbon concentrating mechanisms might be the sole mean to ameliorate carbon fixation. The low sequence divergence between Rubiscos and generally the low number of different forms also suggest that its progress already reached a steady state and its activity can be hardly improved. All single, positive mutations could have been already iterated and did accumulate. Additive effect of multiple mutations could be the only way of improving Rubisco efficiency, yet their exploration might be limited by the assembly chaperones that co-evolved to ensure Rubisco proper formation. GroEL/ES and CPN60/20/10 complexes generally rather promote accumulation of the structural mutations³²⁴, but assembly chaperones have

definitely limited the evolution of LSU. RBCX was shown to decrease the number of permissive mutations of LSU in a directed evolution experiment of the cyanobacterial enzyme³²⁵. Similarly, complementarity between LSU and its chaperones is required for efficient Rubisco assembly. The LSU₂ form of bacteria requires only GroEL/ES for folding and can easily be expressed in chloroplast but foreign Type I Rubisco of closely related species is poorly synthesized in the host²⁷⁴. This limitation can be overcome by co-expression with its cognate chaperones as demonstrated in tobacco-Arabidopsis hybrids²³³ which proves that specificity of the chaperones towards LSU is necessary for the efficient production. Because of the “fixed” catalytic properties of Rubisco, photosynthetic organisms, especially terrestrial ones, produce large quantities of Rubisco to assure their competitiveness. This requires highly controlled expression of their subunits. In bacteria, both Rubisco subunits genes are localized in one operon that ensures their coordinated production. Subsequent assembly, especially of LSUs is made possible by the action of chaperones that stabilize intermediates and prevent disassembly. One of these chaperones, RBCX is even found in the Rubisco operon in some cyanobacteria (e.g. *Synechococcus* sp. PCC7002) where it was reported to be required for Rubisco formation²⁷⁶ (see Introduction: Rubisco folding). Interestingly, its knock-out has no effect on enzymes accumulation in other organisms, even closely related, but the same RBCX can increase cyanobacterial Rubisco overexpression in *E. coli*^{273,277}. Using an artificial micro RNA silencing method³²⁶ we tried to knock-down *RBCX* genes in *Chlamydomonas*, but could not isolate transformants exhibiting a visible defect on Rubisco accumulation (data not shown). Similar efforts have been done by others, and a moderate reduction in the Rubisco accumulation could be observed in the *RBCX*-knocked down lines (Thomas Hauser: thesis). The function of *RBCX* seems to be non-essential, or lost in most phototrophs, at least most probably within eukaryotes. It can be that its presence is redundant with the other described chaperone, *RAF1* (co-existing in the same organisms), that seems to play the same role in stabilization of LSU intermediates. Its effect seems much more pronounced as it is essential for Rubisco synthesis in maize²⁶⁹ and it ameliorates Rubisco solubility and reconstitution *in vitro*²⁸¹. Despite that, it is not essential for *Synechocystis* sp PCC 6803 Rubisco formation²⁸⁴, again suggesting that, at least in cyanobacteria, where translation and assembly of both subunits take place in the same compartment, chaperones could be redundant. On the other hand, in any publication presented so far, native *RBCX* protein could not be detected in cell extracts (we also failed to detect it, even in fractionated extracts (data not shown)). Its low level of accumulation is also observed at transcript level; in *Chlamydomonas* *RBCXs* mRNA are ten times less abundant than that of *RAF1*³²⁷.

Maybe then its function is disappearing in eukaryotes, being replaced by RAF1 but can reemerge when RAF1 is not available or in certain, specific conditions.

The endosymbiosis event led to the creation of dual system of prokaryotic and eukaryotic origins. Genes coding for subunits of the same complex, are now separated and require a more precise control to prevent their unwanted production. Rubisco is a good example of the requirements for nuclear control over chloroplast functions. *RbcL* gene is encoded in multiple copies of the chloroplast genome and is almost constitutively transcribed into thousands (~4,500) copies of mRNA that are stabilized by the MRL1 protein of the OTAF “family”. *rbcL* transcript abundance is linearly correlated with the amount of MRL1 suggesting that the protein is a limiting factor for *rbcL* mRNA accumulation. At the same time as we have demonstrated, MRL1 is not only required for stabilization but also for the translation of the transcript. As no Shine-Dalgarno element is present within the 5'UTR of *rbcL*, MRL1 might be necessary for the re-shaping of the two stem-loops at the very beginning of the 5'UTR where it binds. MRL1 would this way allow translation initiation through sequence modifications and/or even interact with 30S to target the transcript to the ribosome. We failed to find any factors interacting with MRL1 that could complete the M-T factor couple that was observed for many photosynthetic genes (see Introduction: Table 1.2). It does not exclude that such a factor exists, but may support the hypothesis that a single OTAF can perform a dual role in gene expression.

Classic, anterograde control of posttranscriptional gene expression, exerted through trans-acting proteins, many of them (like MRL1) being members of helical-repeat modular protein super family (TPR, OPR, etc.), coexists with assembly-controlled regulation. In *Chlamydomonas*, this process- control by epistasy of synthesis (CES)- is also responsible for Rubisco regulation. We have shown that in the absence of small subunit, LSU autoinhibits its own translation to prevent its wasteful production. This process is assembly dependent as mutations we have introduced to affect the LSU oligomerization prevent it to happen. We have measured LSU translation rates in deregulated context. Our calculations (based on the intensities of radioactive labeling) for now are imprecise, as the different strains we have generated incorporate radioactive acetate at broadly different rates. Nevertheless, we could conclude that when deregulated, *rbcL* translation could reach higher rates than in WT. This suggests that in normal conditions Rubisco formation may be SSU-limited and some amount of free LSU would be constantly available for CES regulation to precisely adjust *rbcL* translation to the availability of its partner. Our results suggest that LSU accumulation is also limited at the level of assembly. Abnormally high translation rates in LSU₈mut and 5'UTR_{psaA}:*rbcL* background result in lower accumulation of LSU, a situation not seen for example for cyt.

f in a mirror situation¹⁶⁴. The reason behind is that unassembled LSU is only stable when in complexed form either with chaperones or as higher order oligomers. Therefore, the observed accumulation levels represent the steady-state balance between assembly-degradation-aggregation. In 5'UTR*psaA:rbcL* (high translation, WT LSU sequence), the total amount of LSU accumulating is limited by the number of chaperones able to stabilize it. The excess is proteolized. In LSU₈mut context (and LSU₂mut: not shown), the high translation of a modified LSU results in a significant amount of LSU aggregating before being able to fold in the chaperonin complex because of the quality control check at the early stage of the assembly or because of the higher propensity for aggregation of mutated LSU. Our observations also suggest that the first stable oligomer of LSU assembly is a RAF1-bound LSU dimer as in LSU₂mut, which is predicted not to be able to dimerize, we could not detect any free LSU. From this result alone one can infer that stabilization of the large subunit is required for any LSU-RNA interaction to occur (for CES). Furthermore, we propose that the last intermediate of the assembly - LSU₈ bound to RAF1- is the effector of the CES-process regulating *rbcL* translation. It is tempting to try testing whether this form is the most stable intermediate of the pathway, this way suggesting its permanent mRNA association. However, ideal immunochase experiment would be tedious to execute because cytoplasmic protein inhibition would also result in the inhibition of chaperones production that are required for the stabilization. Additionally, the stability of LSU-RAF1 intermediates is indicated by the fact that CPN60-independent intermediates accumulate to the same levels in Δ RBCS;5'UTR*psaA:rbcL* and Δ RBCS. While LSU translation rate is at least 20 times lower in Δ RBCS, LSU-RAF1(s) complexes are accumulated at a comparable level, suggesting that in the absence of SSU they are stable. If not, one would expect that in Δ RBCS all dimeric LSU₂ bound to RAF1 would end up degraded or end up in the HMWC. In addition, the similar ratio of low molecular weight LSU-RAF1 intermediate observed in both strains is an indication that RAF1 is limiting for this intermediate's accumulation (higher LSU translation rate does not result in higher accumulation of the LSU₂-RAF1). At the same time, from the intensity of the antibody signal, one could think that HMWC contains less RAF1 than the LMWC. This could be true. The complex can contain other, unidentified proteins. We have designed strains with tagged versions of RAF1 (RAF1-strep) and LSU (LSU-His) to purify the intermediates and analyze their composition by mass spectrometry. Experiments are still ongoing, as once more Rubisco proved not to be easy to collaborate with. Apparently, the histidine-tag has a deleterious effect in LSU oligomerization and therefore as for now, we cannot isolate any of its partners. Interestingly, this situation occurs only in the Δ RBCS mutant background. In WT SSU context, Rubisco-His is accumulating and could be purified (Pierre Crozet, personal communication). One of the possibilities is that the

tag could interfere with the stability of the intermediates which in WT could be compensated by SSU presence that shifts the assembly equilibrium towards the holoenzyme formation. Co-immunoprecipitation coupled to mass spectrometry analysis is required to draw conclusions about the exact composition of the HMWC.

RBCX could be a prime candidate for a possible interactant. As already mentioned, its main role was supposed to aid the LSU oligomerization in a similar way to RAF1³⁰⁸. Notwithstanding, the fact that RBCX knock-downs have no effect on Rubisco accumulation in *Synechocystis* points to the contrary. From the structural data we know that RBCX-LSU interaction sites practically do not overlap with SSU binding positions (Thomas Hauser: thesis and²⁷⁸). One could think that RBCX function would be apparent only in the situation when SSU is less available - it could for example stabilize the octameric structure of the CES-inhibitor.

As for now, the CES process has not been reported in prokaryotic organisms. It could have evolved with the endosymbiotic event as one of the means of the anterograde control over the plast. Convergence with the appearance of OTAFs might have resulted in a dual control system that would assure the tight regulation of multimeric proteins of separate origin. For example MCA1 and TCA1 have already been proposed to act in the CES regulation of the *petA* gene²¹⁰. Previous reports demonstrated that, *in vitro* LSU binds RNA in a sequence-unspecific manner²¹⁶; MRL1 could then participate in the process by providing sequence specificity for *rbcL* transcript. Its amount is linked to the *rbcL* mRNA abundance and from our calculations, the number of LSU₈-RAF1 intermediate that serve as the effector of the CES is in the same range of values (3.5-7 k). We showed that MRL1 is stable for at least 6h after an arrest of cytoplasmic translation. In these conditions, *RBCS* translation is also arrested, therefore, at some point, the translation of *rbcL* should also be arrested in a CES-manner. The fact that MRL1 is stable in the inhibited background does not exclude that it is a part of the inhibitory complex. Altogether, those clues draw a possible scenario for MRL1 participation in the CES process. To tackle this hypothesis, we have generated MRL1-HA strains in the $\Delta rbcL$, $\Delta RBCS$ and $\Delta RBCS;\Delta rbcL$ contexts to follow its behavior related to the CES process. In particular, we want to look at MRL1 possible modifications (e.g. phosphorylation) that could alter its activity and its repartition in the complexes with or without LSU. At the same time, the presence of *rbcL* (or of the reporter gene) transcript in the unassembled LSU complexes will be tested through co-immunoprecipitation.

8. Conclusions

Rubisco biogenesis is tightly controlled by the nucleus through the posttranscriptional regulation exerted via imported proteins. For now, only one factor – MRL1 has been involved in this process. It remains unknown if it is the sole, *rbcL* dedicated OTAF as our efforts to target other putative factors have been fruitless. However, we were able to demonstrate that MRL1 alone is required for both accumulation and translation of *rbcL* transcript suggesting that it might be sufficient for its expression.

At the same time the CES process ensures an additional level of control depending on Rubisco assembly. In this work we have shown that in *Chlamydomonas reinhardtii*, the expression of *rbcL* is controlled by the presence of the SSU. In its absence, large subunit autoinhibits its own transcript's translation, possibly through an interaction with its 5'UTR. This process depends on LSU assembly as mutations aimed to prevent LSU oligomerization allow escaping the regulation. We propose that a high oligomeric state in LSU assembly pathway – an octamer, bound with RAF1 chaperone is the effector of the inhibition. We hypothesize that this oligomer could contain other proteins e.g. directing the RNA recognition. Experiments are ongoing to test the composition of the complex and hopefully identify new Rubisco regulatory factors. One possibility is that the two processes of anterograde signaling (post-transcriptional gene regulation and CES) are interconnected and that MRL1 would be implicated in the CES process as the specificity factor for *rbcL* 5'UTR.

The analysis of the assembly mutants, generated during the time of my thesis, allowed us also to bring some evidence on the fate of Rubisco large subunit. We propose a model where Rubisco synthesis is tuned to the availability of SSU within the stroma. LSU accumulation is limited at the assembly level, by the amounts of its chaperones necessary for its stabilization. Any excess is proteolized, as we could not observe any soluble, un-bound LSU. The first stable form is most probably a LSU dimer stabilized through its interaction with RAF1. When unable to dimerize, monomers are either proteolized or are re-captured by the CPN60 complex. Accumulation of any further oligomers is limited by RAF1 and most probably other factor(s). Further experiments are needed to elucidate the individual function and relationships between factors required for Rubisco formation. Beside RBCX and RAF1 that seem to be *bona fide* LSU assembly chaperones, the RAF2 and BSD2 chaperones are believed to participate in the process at different stages. From our observation of mutated LSU (LSU₈mut) aggregation it is tempting to suggest that LSU needs to interact with chaperones that would guide the nascent polypeptide to the CPN60 complex. On this behalf, unraveling of the BSD2 mechanism of action could shed light on the early stages of LSU synthesis.

9. Materials

Table 9.1: PCR oligonucleotides used; Modified sequences compared to the endogenous sequence are shown in lowercase letters

| Primer name | Sequence |
|------------------|---|
| IP-R15 lin.F | CGTTTCCTTTTCGTTGCTGAAGC |
| IP-R15 lin.R | AGGTGGAATACGAAGGTCTTCAAG |
| IP- LS-A143.F | CTTCGTATTCCACCTtggTACGTATAAACATTCGTA |
| IP-LS-R215D216.R | AACGAAAAGGAAACGtgCAgcCCAACGCATGAA |
| atpB Pst.F | gcgctgcagCTATTAGTAAAGCTGCTTCATT |
| atpB Spe.R | tcgactagtTCACACTCTTATTATTTACTCGCACGT |
| IP-R15 E109.R | GAATAAGTCGATTGGGTAAGCTACG |
| IP-R15R253 Pst.F | gctGTATGTGCTAAAGAATTAGGTG |
| IP-LSE109A Bam.F | CCAATCGACTTATTCGctGAAGGaTCcGTAACATA |
| IP-LS R253A P.R | TTTAGCACATACaGCAGcTTTCATCATTTCTTCACAA |
| PsaAProm.F | cacgtgCTTTTACGAATACACATATGG |
| psaAProm-rbcL.R | AGTTTCTGTTTGTGGAACCATGGATTTCTCCTTATAATAAC |
| psaAPromRbcL.F | GTTATTATAAGGAGAAATCCATGGTTCCACAAACAGAAACT |
| RbcL EcoNI.R | CGACCGTAGTTTTTAGCTGAA |
| IP-PsaAProm.F | GAGAGGAGTGAACAGTCACGTGCTTTTAC |
| IP-RbcL EcoNI.R | CATAAACTGCACGACCGTAGTTTTTAGCTGAAAGAC |
| IP-R15 BseRI.R | ACTGTTCACTCCTCTCCAATATAGTAG |
| IP-R15 EcoNI.F2 | GTCGTGCAGTTTATGAATGTTTAC |
| LSmutA143W.F | TGAAGACCTTCGTATTCCACCTTGG |
| LSA143wt.F | TGAAGACCTTCGTATTCCACCTGCT |
| LSmutD216A.R | GCTTCAGCAACGAAAAGGAAACGTG |
| LSD216wt.R | GCTTCAGCAACGAAAAGGAAACGGTC |
| LSmutE109A.F | AGCTTACCCAATCGACTTATTCGCT |
| LSE109wt.F | CGTAGCTTACCCAATCGACTTATTCGAA |
| LSmutR253A.R | CTAATTCTTTAGCACATACAGCAGC |
| LSR253wt.R | TAATTCTTTAGCACATACTGCACG |

Table 9.2: Antibodies used in this study

| Antibody | Reactivity | Origin |
|---|---|------------------------------|
| α Rubisco | Type 1 LSU, SSU; Type 2 LSU | Gift from Spencer M. Whitney |
| α Cyt <i>f</i> | Chlamydomonas Cytochrome <i>f</i> | Generated in the lab |
| α CPN60 | CPN60 α ; CPN60 β | Gift from Michael Schroda |
| α RAF1 | Chlamydomonas Rubisco Accumulation Factor 1 | Generated by Genescript |
| α Tubulin | Plants, algae α tubulin | Agrisera |
| α RPS12 | pRibosomal subunit 12 | Gift from J.D. Rochaix |
| α EPYC1 | CrEPYC1 | Gift from Martin Jonikas |
| α HA | HA-tag | Covance |
| α PsaD | Photosystem I subunit D | Agrisera |
| α OEE2 | Oxygen evolving complex 2 | Generated in the lab |
| LHCII | PSII Light harvesting antenna | Gift from R. Basi |
| α Rabbit = HRP-conjugated secondary antibodyPolyclonal | Polyclonal, rabbit-raised antibodies | Promega |
| α Mouse = HRP-conjugated secondary antibodyPolyclonal | Monoclonal, mouse-raised antibodies | Promega |

Table 9.3: *E. coli* used in this study

| Strain | Description | Origin |
|-------------|------------------------|---------------------|
| NEB 5-alpha | Thermocompetent strain | New England Biolabs |

Table 9.4: *Chlamydomonas* strains used in this study

| Strain | Mutation | Origin | Reference |
|---|---|---|---------------------------|
| T222+ | Wild type strain used in the laboratory | Derivative of 137c | ChlamyStation Collection* |
| ΔRbcL 1.7.5 | Deletion of the <i>rbcL</i> gene par insertion of the antibiotic resistance cassette | WTN+ wild type | 1 |
| ΔR T1.3+ | Detetion of the <i>rbcL</i> gene par insertion of the antibiotic resistance cassette | ΔRbcL 1.7.5 crossed with WT24- strain | This study |
| ΔSSU | Cal13.1B-; Cal13.5A+; Insertional mutagenesis deletion of a 35kbp region resulting in an absence of <i>RBCS1</i> and <i>RBCS2</i> genes | Mutant from R. Dent photosynthetic mutant collection backcrossed to T222+ | 293 |
| 5'UTR <i>rbcL::petA</i> | 5'UTR <i>rbcL</i> driven <i>petA</i> gene | | 1 |
| 5'UTR <i>rbcL::petA::ΔSSU</i> | 5'UTR <i>rbcL</i> driven <i>petA</i> gene with a deletion of both <i>RBCS</i> genes | 5'UTR <i>rbcL::petA</i> crossed with ΔSSU (Cal13.1B-) | 2 progenies T1.2 and O1.4 |
| 5'UTR <i>psaA::rbcL</i> | 5'UTR <i>psaA</i> driven <i>rbcL</i> gene | | This study |
| 5'UTR <i>psaA::rbcL::ΔSSU</i> | | | This study |
| 5'UTR <i>psaA::rbcL::5'UTRrbcL::petA::ΔSSU</i> | | | This study |
| 5'UTR <i>psaA::rbcL::5'UTRrbcL::petA::ΔSSU</i> WT | Crossed with WTS24- | progenies of cross | This study |
| LSU* | | | This study |
| LSU*:5'UTR <i>rbcL::petA</i> | | | This study |
| LSU*:5'UTR <i>rbcL::petA::ΔSSU</i> | | | This study |
| LSU ₈ mut | <i>rbcL</i> A143W-R215A-D216A | | This study |
| LSU ₈ mut:ΔSSU | <i>rbcL</i> A143W-R215A-D216A in Cal13.5A+ background | | This study |
| LSU ₂ mut | E109A-R253A | | This study |
| LSU ₂ mut:WT | Reference strain | | This study |
| LSU ₂ mut:ΔSSU | <i>rbcL</i> E109-R253 in Cal13.5A+ background | | This study |

* <http://chlamystation.free.fr/>

10. Methods

10.1. Cultures

C. reinhardtii

If not stated otherwise, wild type (T222+ : derivative of 137c strain) and mutant strains of *C. reinhardtii* were grown on solid Tris-acetate-phosphate medium (TAP) supplemented with agar and if necessary selective antibiotics, and in liquid cultures under continuous, dim light ($7 \mu\text{M photons} \times \text{m}^{-2} \times \text{s}^{-1}$) on an orbital shaker (120 rpm) at RT. Cells from exponentially growing cultures ($2 \times 10^6 \text{ cells mL}^{-1}$) were used for all experiments.

E. coli

NEB 5-alpha strain of *E. coli* was used for the plasmids amplification. It was grown on solid and liquid LB medium at 37°C with shaking (liquid cultures) with or without addition of the selective antibiotics.

10.2. Genetics methods

Chlamydomonas mating protocol

Strains of opposing mating types (“+” and “-”) were plated densely on the TAP medium depleted 10x in nitrogen source and left for a 5-days starvation to induce gametogenesis. Subsequently cells were mixed together in a 1:1 ratio in approx. 5 ml of sterile water and let shaking for 30 min. After this, they were placed under $60 \mu\text{mol photons} \times \text{m}^{-2} \times \text{s}^{-1}$ light without shaking. After the first signs of zygote formation (layer of the cells forming on the surface and at the bottom of the flask) drops of the cell mixture were deposited every hour on the TAP plates containing 30 g agar $\times \text{l}^{-1}$. After drying, plates were kept in the darkness for 10-14 days for the zygote maturation. Subsequently, vegetative cells were removed from the plate and zygotes were transferred to a fresh TAP plate and then separated using Singer Micromanipulator. Additionally, each zygotes-containing plate was treated with chloroform foams to kill any resting, vegetative cells. Plates prepared this way were left overnight in dim light for the duration of the meiotic division of the zygotes. Next morning progenies of the division were separated on the plate and left for further growth. For a more descriptive, visual protocol see Jiang and Stern, 2009³²⁸.

Transformation protocols:

Nuclear

Cells are concentrated to a final density of 4×10^8 cells mL⁻¹ in TAP medium supplemented with 40 mM sucrose and incubated at 16°C for 20 min. Subsequently cells are electroporated (using GenePulser XCell™; BioRad) in the presence of up to 2 µg of linearized plasmid and incubated at 16°C for 20 more min. Afterwards cells are plated either on minimal medium (MIN: TAP without acetate) for the autotrophy recovery selection or on TAP plates containing selective antibiotic (in the latter case cells are adapted overnight on a rotary shaker at dim light for the proper expression of the resistance gene).

Chloroplast

Chloroplast transformation was done as described in Boynton *et al.* 1988³²⁹. 1×10^8 cells are plated on a selective medium to create a “cell carpet”. 0.2 µm tungsten beads are used as carriers for the plasmid DNA (up to 4 µg). Transformation was done using build *in house* particle cannon.

UV mutagenesis

15 mL from an exponentially growing culture at 2.10^6 cells/ml were poured in an empty Petri dish together with a sterile paper clip. While stirring, they were exposed to UV treatment in an in-house built UV box for 30s at $3.6 \text{ erg} \cdot \text{mm}^{-2} \cdot \text{s}^{-1}$. Cells were subsequently left to recover for one hour in the dark with stirring, before 250 µl (about 5.10^5 cells) were layered onto a 5FC containing selective TAP plate (2mg/ml). Cell viability after UV treatment was estimated to be around 10%. To recover independent mutants, multiple treatments were made, from which a single aliquot was used and put on a selective plate.

10.3. Biophysics methods

The fluorescence based screen was conducted according to Johnson *et al.* 2009²⁹⁵. TAP-grown cells/transformants were dark adapted for 15 min prior to the measurements. Fluorescence kinetics induction was measured under constant illumination of 120 or 250 µM red light photons $\times \text{m}^{-2} \times \text{s}^{-1}$ for 7 min.

10.4. Molecular biology methods

If not stated otherwise DNA manipulations were done following standard protocols as in Sambrook et al. 1989³¹⁷.

E. coli transformation

Transformation was done through heat shock method following the instructions of the NEB 5-alpha cells provider (New England Biolabs).

Chlamydomonas fast DNA isolation

To perform genotyping, DNA was isolated using a little amount of cells suspended in 10 µl water to which 10 µl of 100% ethanol was added and incubated 5 min at RT. Afterwards 80 µl of 5% Chelex resin was added and the mixture was incubated 8 min at 95°C. The supernatant was separated and served as a template for PCR.

Chlamydomonas RNA isolation, Southern blotting and hybridization

RNA isolation was done as described in Drapier et al. 1998³³⁰. RNA was separated on 1.2% agarose gel with 8% formaldehyde. Subsequently, capillary transfer to a TM Membrane (QBIAGEN) was done and the membranes were hybridized with appropriate probes as described in Drapier et al. 1992³³¹.

Plasmid construction (pLS_{ARD}, aAXdB, pL₂mut, pLStr, paAR)

-pLS_{ARD} plasmid was generated using the In-Fusion PCR Cloning Kit (Clontech, In-Fusion® HD Cloning Plus), following the manufacturer's instructions, from the R15 plasmid backbone (Johnson et al., 2010) amplified with the IP-R15 lin.F and R primers, and a 252 bp amplified region from R15 with primers IP-LS-A143.F and IP-LS-R215D216.R introducing the A143W, and R215A and D216A mutations respectively.

The *aadA* excisable cassette driven by the *psaA* promoter described in²⁰³ was modified to replace *rbcl* 3' regulatory region with the ones of *atpB*. To this end, *atpB* 3'UTR was PCR-amplified and flanked by *Pst*I and *Spe*I restriction sites using the *atpB* Pst.F and *atpB* Spe.R primers, and cloned into *Pst*I-*Spe*I digested paAX plasmid, yielding the paAEXCdB plasmid. It was further digested by *Kpn*I and *Sac*I, blunted by NEB Quickblunting kit and ligated into *Afl*III-digested pLS_{ARD} plasmid, or R15 plasmid. Clones in which the *aadA* cassette inserted in a reverse orientation compared to the *rbcl* genes were selected, yielding pLSARD-X and pR15-X3' respectively.

-pL₂mut plasmid was similarly assembled from an R15 PCR-amplified fragment using the IP-R15 E109.R and IP-R15R253 *Pst*.F primers, and a 473pb amplified fragment containing the mutated *rbcL* region containing the E109A and R253A substitutions introduced with the IP-LS E109A Bam.F and IP-LS R253A P.R primers. The *aadA* marker (*KpnI*-*SacI* fragment of paAXdB, blunted) was thereafter introduced at the *BseRI* site upstream of the *rbcL* promoter in reverse orientation compared to *rbcL*, yielding the pLS2-X plasmid. To check that the cassette insertion is neutral on *rbcL* expression, the aAXdB marker was also introduced in the pR15 plasmid, yielding pR15-X5'

To generate the paAR plasmid, the *psaA* promoter region was first fused to part of *rbcL* CDS sequence by overlapping PCR using the following primers: PsaAProm.F, psaAProm-rbcL.R, psaAPromRbcL.F and RbcL EcoNI.R on the paAXdB and R15 plasmid templates with the Phusion Taq polymerase (NEB). The resulting 814 bp fragment was further amplified using the IP-psaAProm.F and IP-rbcL EcoNI primers, and assembled into the R15 backbone amplified by the IP-R15 *BseRI*.R and IP-R15 EcoNI.F2 primers using the Clontech In-Fusion PCR Cloning kit. Insertion of the aAXdB excisable marker (*KpnI*-*SacI* blunted fragment) was further performed at the *BseRI* restriction site, yielding paAR-X plasmid in which the *aadA* marker is in opposite orientation compared to *rbcL*. All clones were sequenced, and no mutations were found within the chloroplast containing sequences.

10.5. Biochemical analysis:

Chlamydomonas "whole cell" protein preparation

Collected cells (30 × concentrated) were washed in cold 5 mM Hepes, 20 mM EDTA buffer supplemented with EDTA-free protease inhibitors mix (Roche), then resuspended in 0.1M DTT, Na₂CO₃ and frozen in liquid nitrogen.

Soluble fraction isolation

4 × 10⁸ cells were pelleted and resuspended in Native Extraction buffer (20 mM HEPES pH= 7.5, 20 mM KCl, 10% glycerol, 2× protease inhibitor mix). They were broken using a French press cell apparatus (6000 psi) and centrifuged at 267000 rcf at 4°C for 25 min. Soluble fraction was collected and concentrated using Amicon Ultra 30 kDa cutoff centrifugation units (Millipore). Samples were used immediately after preparation.

Fractionation protocol

The experiment was done following a modified protocol described in Zhan et al. 2015³¹³. 4×10^8 cells were resuspended in 5ml of buffer (5 mM Hepes, 20 mM EDTA buffer supplemented with protease inhibitors mix), they were broken using a French press apparatus (6000 psi) and centrifuged at 500 rcf at 4°C for 1 min to remove unbroken cells. An aliquot was collected that would correspond to “whole cell” fraction. The rest of the supernatant was centrifuged at 267000 rcf at 4°C for 25 min. the supernatant was collected and concentrated using Amicon Ultra 10 kDa cutoff centrifugation units (Millipore) to create a “soluble” fraction. The pellet was resuspended in the same amount of buffer supplemented with 2% Triton and incubated at RT with gentle shaking for 15 min. It was finally spun down at 13200 rcf at 4°C for 20 min. The supernatant was designed “Triton solubilized pellet” fraction and the remaining, resuspended pellet made a “Triton insoluble” fraction. All fractions were immediately frozen in liquid nitrogen and were used for SDS-PAGE analysis as described below.

Antibiotics Immunochase

To follow proteins' lifetime, antibiotics arresting either chloroplast or cytoplasmic translation ($200 \mu\text{g} \times \text{ml}^{-1}$ chloramphenicol, $200 \mu\text{g} \times \text{ml}^{-1}$ lincomycin and $10 \mu\text{g} \times \text{ml}^{-1}$ cycloheximide respectively) were added to the cell cultures. At given time points cell aliquots were collected and “whole cell” protein fraction was prepared.

Pulse experiment

Chlamydomonas ^{14}C -acetate pulse experiment was done as described in Kuras and Wollman 1994¹⁹⁷. 5×10^6 cells were washed one time in MIN medium then resuspended in 5 ml of fresh MIN medium and incubated for 1h at RT with vigorous shaking to remove the acetate from the medium. Subsequently, $5 \mu\text{l}$ of $1\text{mg} \times \text{ml}^{-1}$ cycloheximide and 50 μCi of ^{14}C -acetate were added to the cells. After 7 min of vigorous shaking cells were mixed with 35 ml of cold TAP medium supplemented with 40mM acetate and immediately spun down. Afterwards they were washed in 5 mM Hepes, 20 mM EDTA buffer supplemented with EDTA-free protease inhibitors mix (Roche). Then they were resuspended in 0.1M DTT and Na_2CO_3 . Finally samples were suspended 1:1 in 2% SDS, 20% sucrose solution, boiled for 1 minute, then spun down at 12 000g for 15 minutes and subsequently the supernatant was loaded on the gel. The gel protocol and electrophoresis are described in²⁰⁵. Afterwards, gels were dried and exposed to autoradiography screens for at least one month. Phosphorescence signal was measured using Typhoon FLA 9500 Reader (GE Healthcare).

Protein quantification

Protein concentrations were measured calorimetrically by Bradford assay ³²¹ using QuickStart Bradford Dye reagent (Bio-Rad) following manufacturer's protocol.

SDS PAGE:

Standard:

Protein electrophoresis in denaturing conditions was performed according to the modified Laemmli protocol ³¹⁹. For protein loading of "whole cell" samples, chlorophyll fluorescence was used for quantification according to ¹⁹⁷. Samples were suspended 1:1 in 2% SDS, 20% sucrose solution than boiled for 1 minute, then spun down at 12 000g for 15 minutes and subsequently loaded on the gel. Electrophoresis was performed in midi cool slab apparatus (Bio Craft Japan) with a migration buffer (12.14 mM Tris, 134.2 mM glycine, 0.1 % (w/v) SDS, 1 mM EDTA). Table 10.1 shows gel casting scheme.

Table 10.1: Preparation of SDS PAGE gels (amounts for 1 gel)

| Component | Resolving gel | | | Stacking gel |
|-------------------------|---------------|--------|--------|--------------|
| | 8% | 10% | 12% | 4.5% |
| 30% AA/bisAA 37.5/1 | 4 ml | 5 ml | 6 mL | 1.15 ml |
| 3M Tris-HCl pH=8.8 | 1.9 ml | 1.9 ml | 1.9 ml | - |
| 0.5M Tris-HCl pH=6.8 | - | - | - | 1.75 ml |
| H ₂ O MQ | 9.1 ml | 8.1 ml | 7.1 ml | 9.1 ml |
| SDS 20% | 75 µl | 75 µl | 75 µl | 37 µl |
| TEMED | 20 µl | 20 µl | 20 µl | 10 µl |
| 10% APS | 75 µl | 75 µl | 75 µl | 37 µl |

For Pulse experiments

Sample treatment was similar to the one of Standard protocol. Electrophoresis was done following the protocol from Kuras and Wollman 1994 ¹⁹⁷. *In house* gel system was used for the electrophoresis with a migration buffer (12.14 mM Tris, 134.2 mM glycine, 0.1 % (w/v) SDS, 1 mM EDTA). Gels were prepared according to the table 10.2 recipe.

Table 10.2: Preparation of gradient SDS-PAGE gels (amounts for 1 gel)

| Component | Resolving gel - gradient | | Stacking gel |
|-------------------------|--------------------------|--------|--------------|
| | 12% | 18% | 5% |
| 40% AA/bis AA 37.5/1 | 6 ml | 9 ml | - |
| 40% AA | 6 ml | 9 ml | - |
| 30% AA/bis AA 37.5/1 | - | - | 5 ml |
| 3M Tris-HCl pH=8.8 | 10 ml | 10 ml | - |
| 0.5M Tris-HCl pH=6.8 | - | - | 7.5 ml |
| Sucrose | - | 5.3 g | - |
| Urea | 19.4 g | 19.4 g | - |
| H ₂ O MQ | 8.8 ml | 1 ml | 17.3 ml |
| TEMED | 8.4 µl | 9.6 µl | 20 µl |
| 10% APS | 24 µl | 24 µl | 200 µl |

Native PAGE

Native electrophoresis was done according to a modified Schägger protocol ³²⁰. Samples (e.g. Chlamydomonas soluble proteins fractions) were suspended in Native Buffer (20 mM HEPES pH= 7.5, 20 mM KCl, 10% glycerol, 2x protease inhibitor mix, 0.25% bromophenol blue). Electrophoresis was done either in MiniProtean electrophoresis units, using predefined 4-16% gradient gels (Invitrogen) or in Midi cool slab units (Bio Craft Japan) using gels casted according to the Table 10.3. Migration was done at 4°C in a buffer consisting of 50 mM Tricine, 15 mM Bis-Tris pH=7.0 at constant voltage of 60 V for 1h than 120 V.

Table 10.3: Preparation of Native PAGE gels (amounts for 1 gel)

| Component | Resolving gel - gradient | | Stacking gel |
|--|--------------------------|---------------|--------------|
| | 4.5% | 15% | 4% |
| 30% AA/bis AA 37.5/1 | 1.7 ml | 5.2 ml | 1.33 ml |
| 6x Gel buffer (300 mM Bis-Tris pH=7; 3M ACA) | 1.75 ml | 1.75 ml | 1.67 ml |
| Glycerol | - | 2.1 g | - |
| H ₂ O MQ | Up to 10,5 ml | Up to 10,5 ml | Up to 10 ml |
| TEMED | 5 µl | 5 µl | 10 µl |
| 10% APS | 20 µl | 20 µl | 100 µl |

Two-dimensional electrophoresis

Native- to SDS-PAGE electrophoresis was done using strips from 1D-Native gels that were incubated 30 min in equilibration buffer (50 mM Tris-HCl pH 6.8, 30% Glycerol, 2% SDS, 0,67% B-mercaptoethanol) than placed in the Midi cool slab gel unit (Bio Craft Japan) and the second dimension denaturizing electrophoresis was conducted following the protocol described in²⁰⁵.

Gels and membranes staining

For in gel protein detection Coomasie blue staining was performed using Coomasie Staining Solution (0.16% Coomasie brilliant blue R-250, 40% Methanol, and 7% Acetic acid) followed by multiple washes in 10% Methanol, and 7% Acetic acid. For Mass-Spectrometry adapted protein quantification gels were stained using SYPRO Ruby Protein Gel Stain (Thermo Fisher Scientific) according to the manufacturer's protocol. Western blot membranes were stained using Ponceau Red stain (0.2% Ponceau red, 3% tri-chloro acetic acid) followed by multiple washes with MQ water.

Western blotting and immunodetection

Protein transfer into nitrocellulose or PVDF membrane was done following Towbin H. *et al.*³³². Transfers were done in a Semi-Dry apparatus (Bio-Rad) using 3-buffer system developed in the laboratory with a constant current of $1\text{mA} \times (\text{cm}^2)^{-1}$ for 100 min. Membranes were then blocked with 3% dry milk (m/v) in PBS-T (1× PBS, 0.1% TWEEN) buffer for 45 min and incubated in the primary antibodies for 1h, washed 3 times 10 min in PBS-T buffer and finally incubated 1h with the secondary antibody conjugated with horseradish peroxidase. After 3 consecutive 10 min washes in PBS-T membranes were exposed to Clarity™ Western ECL substrate (Bio-Rad) and bands' chemiluminescence was detected using ChemiDoc XR+ (BioRad) apparatus. Quantification of the signal intensities was done using ImageLab software (BioRad).

Bibliography

- 1 Johnson, X. *et al.* MRL1, a conserved Pentatricopeptide repeat protein, is required for stabilization of rbcL mRNA in Chlamydomonas and Arabidopsis. *Plant Cell* **22**, 234-248, (2010).
- 2 Wheatley, N. M., Sundberg, C. D., Gidaniyan, S. D., Cascio, D. & Yeates, T. O. Structure and identification of a pterin dehydratase-like protein as a ribulose-bisphosphate carboxylase/oxygenase (RuBisCO) assembly factor in the alpha-carboxysome. *The Journal of biological chemistry* **289**, 7973-7981, (2014).
- 3 Bracher, A., Whitney, S. M., Hartl, F. U. & Hayer-Hartl, M. Biogenesis and Metabolic Maintenance of Rubisco. *Annual review of plant biology*, (2017).
- 4 Staehelin, L. A. Chloroplast structure: from chlorophyll granules to supra-molecular architecture of thylakoid membranes. (2003).
- 5 Choquet, Y. & FA, W. Chlamydomonas sourcebook - The CES process. *Chlamydomonas sourcebook; Chapter 29; Volume 2*, (2007).
- 6 Cavaiuolo, M., Kuras, R., Wollman, F. A., Choquet, Y. & Vallon, O. Small RNA profiling in Chlamydomonas: insights into chloroplast RNA metabolism. *Nucleic Acids Research*, in press, (2017).
- 7 Engel, B. D. *et al.* Native architecture of the Chlamydomonas chloroplast revealed by in situ cryo-electron tomography. *eLife* **4**, (2015).
- 8 Loizeau, K. *et al.* Small RNAs reveal two target sites of the RNA-maturation factor Mbb1 in the chloroplast of Chlamydomonas. *Nucleic Acids Research* **42**, 3286-3297, (2014).
- 9 Andersson, I. & Backlund, A. Structure and function of Rubisco. *Plant physiology and biochemistry : PPB / Societe francaise de physiologie vegetale* **46**, 275-291, (2008).
- 10 Anthonisen, I. L., Salvador, M. L. & Klein, U. Specific sequence elements in the 5' untranslated regions of rbcL and atpB gene mRNAs stabilize transcripts in the chloroplast of Chlamydomonas reinhardtii. *RNA* **7**, 1024-1033, (2001).
- 11 Whitney, S. M., Houtz, R. L. & Alonso, H. Advancing our understanding and capacity to engineer nature's CO₂-sequestering enzyme, Rubisco. *Plant Physiol* **155**, 27-35, (2011).
- 12 Suay, L., Salvador, M. L., Abesha, E. & Klein, U. Specific roles of 5' RNA secondary structures in stabilizing transcripts in chloroplasts. *Nucleic Acids Research* **33**, 4754-4761, (2005).
- 13 Eberhard, S., Finazzi, G. & Wollman, F. A. The dynamics of photosynthesis. *Annual review of genetics* **42**, 463-515, (2008).
- 14 Ellis, R. J. The most abundant protein in the world. *Trends in Biochemical Sciences* **4**, 241-244, (1979).
- 15 Dyall, S. D., Brown Mt Fau - Johnson, P. J. & Johnson, P. J. Ancient invasions: from endosymbionts to organelles. (2004).
- 16 Ball, S. G. *et al.* Metabolic Effectors Secreted by Bacterial Pathogens: Essential Facilitators of Plastid Endosymbiosis? *The Plant Cell* **25**, 7-21, (2013).
- 17 Marin, B., M. Nowack, E. C. & Melkonian, M. A Plastid in the Making: Evidence for a Second Primary Endosymbiosis. *Protist* **156**, 425-432, (2005).
- 18 Nakayama, T. & Ishida, K.-i. Another acquisition of a primary photosynthetic organelle is underway in Paulinella chromatophora. *Current Biology* **19**, R284-R285, (2009).
- 19 Ochoa de Alda, J. A. G., Esteban, R., Diago, M. L. & Houmard, J. The plastid ancestor originated among one of the major cyanobacterial lineages. **5**, 4937, (2014).
- 20 Ponce-Toledo, R. I. *et al.* An Early-Branching Freshwater Cyanobacterium at the Origin of Plastids. *Current Biology* **27**, 386-391.

- 21 Dagan, T. *et al.* Genomes of Stigonematalean Cyanobacteria (Subsection V) and the Evolution of Oxygenic Photosynthesis from Prokaryotes to Plastids. *Genome Biology and Evolution* **5**, 31-44, (2013).
- 22 Keeling, P. J. The Number, Speed, and Impact of Plastid Endosymbioses in Eukaryotic Evolution. *Annual Review of Plant Biology* **64**, 583-607, (2013).
- 23 Oldenburg, D. J. & Bendich, A. J. The linear plastid chromosomes of maize: terminal sequences, structures, and implications for DNA replication. *Current Genetics* **62**, 431-442, (2016).
- 24 Margulis, L. Symbiosis and evolution. *Scientific American* **Aug;225(2)**, 48-57, (1971).
- 25 Odahara, M., Kobayashi, Y., Shikanai, T. & Nishimura, Y. Dynamic Interplay between Nucleoid Segregation and Genome Integrity in Chlamydomonas Chloroplasts. *Plant Physiology* **172**, 2337-2346, (2016).
- 26 Powikrowska, M., Oetke, S., Jensen, P. E. & Krupinska, K. Dynamic composition, shaping and organization of plastid nucleoids. *Frontiers in Plant Science* **5**, 424, (2014).
- 27 Green, B. R. Chloroplast genomes of photosynthetic eukaryotes. *The Plant Journal* **66**, 34-44, (2011).
- 28 Liere, K. & Link, G. RNA-binding activity of the matK protein encoded by the chloroplast trnK intron from mustard (*Sinapis alba* L.). *Nucleic Acids Research* **23**, 917-921, (1995).
- 29 Boudreau, E. *et al.* A large open reading frame (orf1995) in the chloroplast DNA of Chlamydomonas reinhardtii encodes an essential protein. *Molecular & general genetics : MGG* **253**, 649-653, (1997).
- 30 Drescher, A., Ruf, S., Calsa, T., Jr., Carrer, H. & Bock, R. The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. *The Plant journal : for cell and molecular biology* **22**, 97-104, (2000).
- 31 Kikuchi, S. *et al.* Uncovering the protein translocon at the chloroplast inner envelope membrane. *Science* **339**, 571-574, (2013).
- 32 Maul, J. E. *et al.* The Chlamydomonas reinhardtii Plastid Chromosome: Islands of Genes in a Sea of Repeats. *The Plant Cell* **14**, 2659-2679, (2002).
- 33 Meeks, J. C. *et al.* An overview of the genome of Nostoc punctiforme, a multicellular, symbiotic cyanobacterium. (2001).
- 34 Stegemann, S. & Bock, R. Experimental Reconstruction of Functional Gene Transfer from the Tobacco Plastid Genome to the Nucleus. *The Plant Cell* **18**, 2869-2878, (2006).
- 35 Martin, W. *et al.* Evolutionary analysis of Arabidopsis, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. (2002).
- 36 Cullis, C. A., Vorster, B. J., Van Der Vyver, C. & Kunert, K. J. Transfer of genetic material between the chloroplast and nucleus: how is it related to stress in plants? *Annals of Botany* **103**, 625-633, (2009).
- 37 Wicke, S., Schneeweiss, G. M., dePamphilis, C. W., Müller, K. F. & Quandt, D. The evolution of the plastid chromosome in land plants: gene content, gene order, gene function. *Plant Molecular Biology* **76**, 273-297, (2011).
- 38 Zerges, W., S., W. & Rochaix, J. D. Light activates binding of membrane proteins to chloroplast RNAs in Chlamydomonas reinhardtii. (2002).
- 39 Daley, D. O. & Whelan, J. Why genes persist in organelle genomes. (2005).
- 40 Allen, J. F. Why chloroplasts and mitochondria retain their own genomes and genetic systems: Colocation for redox regulation of gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 10231-10238, (2015).
- 41 Allen, J. F. The function of genomes in bioenergetic organelles. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **358**, 19-37; discussion 37-18, (2003).

- 42 Delannoy, E., Fujii, S., Colas des Francs-Small, C., Brundrett, M. & Small, I. Rampant Gene Loss in the Underground Orchid *Rhizanthella gardneri* Highlights Evolutionary Constraints on Plastid Genomes. *Molecular Biology and Evolution* **28**, 2077-2086, (2011).
- 43 J L Popot, a. & Vitry, C. On the Microassembly of Integral Membrane Proteins. *Annual Review of Biophysics and Biophysical Chemistry* **19**, 369-403, (1990).
- 44 Torabi, S. *et al.* PsbN Is Required for Assembly of the Photosystem II Reaction Center in *Nicotiana tabacum*. *The Plant Cell* **26**, 1183-1199, (2014).
- 45 Levey, T., Westhoff, P. & Meierhoff, K. Expression of a nuclear-encoded psbH gene complements the plastidic RNA processing defect in the PSII mutant hcf107 in *Arabidopsis thaliana*. *The Plant Journal* **80**, 292-304, (2014).
- 46 Smith, D. R. & Lee, R. W. A Plastid without a Genome: Evidence from the Nonphotosynthetic Green Algal Genus *Polytomella*. *Plant Physiology* **164**, 1812-1819, (2014).
- 47 Terashima, M., Specht, M. & Hippler, M. The chloroplast proteome: a survey from the *Chlamydomonas reinhardtii* perspective with a focus on distinctive features. *Current Genetics* **57**, 151-168, (2011).
- 48 Li, H.-m. & Chiu, C.-C. Protein Transport into Chloroplasts. *Annual Review of Plant Biology* **61**, 157-180, (2010).
- 49 Nakai, M. The TIC complex uncovered: The alternative view on the molecular mechanism of protein translocation across the inner envelope membrane of chloroplasts. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 957-967, (2015).
- 50 Sjuts, I., Soll, J. & Bölter, B. Import of Soluble Proteins into Chloroplasts and Potential Regulatory Mechanisms. *Frontiers in Plant Science* **8**, 168, (2017).
- 51 Bruce, B. D. The paradox of plastid transit peptides: conservation of function despite divergence in primary structure. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1541**, 2-21, (2001).
- 52 Keegstra, K. & Cline, K. Protein import and routing systems of chloroplasts. (1999).
- 53 Wollman, F.-A. An antimicrobial origin of transit peptides accounts for early endosymbiotic events. *Traffic* **17**, 1322-1328, (2016).
- 54 López-Juez, E. Plastid biogenesis, between light and shadows. *Journal of Experimental Botany* **58**, 11-26, (2007).
- 55 Liebers, M. *et al.* Regulatory Shifts in Plastid Transcription Play a Key Role in Morphological Conversions of Plastids during Plant Development. *Frontiers in Plant Science* **8**, 23, (2017).
- 56 Eberhard, S., Drapier, D. & Wollman, F. A. Searching limiting steps in the expression of chloroplast-encoded proteins: relations between gene copy number, transcription, transcript abundance and translation rate in the chloroplast of *Chlamydomonas reinhardtii*. (2002).
- 57 Matsuo, M. & Obokata, J. Dual roles of photosynthetic electron transport in photosystem I biogenesis: light induction of mRNAs and chromatic regulation at post-mRNA level. (2002).
- 58 Zoschke, R., Liere, K. & Börner, T. From seedling to mature plant: *Arabidopsis* plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *The Plant Journal* **50**, 710-722, (2007).
- 59 Carter, M. L., Smith, A. C., Kobayashi, H., Purton, S. & Herrin, D. L. Structure, circadian regulation and bioinformatic analysis of the unique sigma factor gene in *Chlamydomonas reinhardtii*. *Photosynthesis research* **82**, 339-349, (2004).
- 60 Irihimovitch, V. & Stern, D. B. The sulfur acclimation SAC3 kinase is required for chloroplast transcriptional repression under sulfur limitation in *Chlamydomonas reinhardtii*. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 7911-7916, (2006).

- 61 Chotewutmontri, P. & Barkan, A. Dynamics of Chloroplast Translation during Chloroplast Differentiation in Maize. *PLoS Genetics* **12**, e1006106, (2016).
- 62 Udy, D. B., Belcher, S., Williams-Carrier, R., Gualberto, J. M. & Barkan, A. Effects of Reduced Chloroplast Gene Copy Number on Chloroplast Gene Expression in Maize. *Plant Physiology* **160**, 1420-1431, (2012).
- 63 Schweer, J., Geimer, S., Meurer, J. & Link, G. Arabidopsis mutants carrying chimeric sigma factor genes reveal regulatory determinants for plastid gene expression. *Plant & cell physiology* **50**, 1382-1386, (2009).
- 64 Drager, R. G., Girard-bascou, J., Choquet, Y., Kindle, K. L. & Stern, D. B. In vivo evidence for 5'→3' exoribonuclease degradation of an unstable chloroplast mRNA. *The Plant Journal* **13**, 85-96, (1998).
- 65 Hicks, A., Drager, R. G., Higgs, D. C. & Stern, D. B. An mRNA 3' Processing Site Targets Downstream Sequences for Rapid Degradation in Chlamydomonas Chloroplasts. *Journal of Biological Chemistry* **277**, 3325-3333, (2002).
- 66 Yehudai-Resheff, S., Hirsh, M. & Schuster, G. Polynucleotide Phosphorylase Functions as Both an Exonuclease and a Poly(A) Polymerase in Spinach Chloroplasts. *Molecular and cellular biology* **21**, 5408-5416, (2001).
- 67 Miranda, R. G., Rojas, M., Montgomery, M. P., Gribbin, K. P. & Barkan, A. RNA-binding specificity landscape of the pentatricopeptide repeat protein PPR10. *RNA* **23**, 586-599, (2017).
- 68 Stern, D. B., Goldschmidt-Clermont, M. & Hanson, M. R. Chloroplast RNA Metabolism. *Annual Review of Plant Biology* **61**, 125-155, (2010).
- 69 Drechsel, O. & Bock, R. Selection of Shine-Dalgarno sequences in plastids. *Nucleic Acids Research* **39**, 1427-1438, (2011).
- 70 Bieri, P., Leibundgut, M., Saurer, M., Boehringer, D. & Ban, N. The complete structure of the chloroplast 70S ribosome in complex with translation factor pY. *The EMBO Journal*, (2016).
- 71 Beligni, M. V., Yamaguchi, K. & Mayfield, S. P. The translational apparatus of Chlamydomonas reinhardtii chloroplast. *Photosynthesis research* **82**, 315-325, (2004).
- 72 Watson, J. C. & Surzycki, S. J. Extensive sequence homology in the DNA coding for elongation factor Tu from Escherichia coli and the Chlamydomonas reinhardtii chloroplast. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 2264-2267, (1982).
- 73 Graf, M. *et al.* Cryo-EM structure of the spinach chloroplast ribosome reveals the location of plastid-specific ribosomal proteins and extensions. *Nucleic Acids Research* **45**, 2887-2896, (2017).
- 74 Manuell, A., Beligni, M. V., Yamaguchi, K. & Mayfield, S. P. Regulation of chloroplast translation: interactions of RNA elements, RNA-binding proteins and the plastid ribosome. *Biochemical Society transactions* **32**, 601-605, (2004).
- 75 Yamaguchi, K. & Subramanian, A. R. The Plastid Ribosomal Proteins: IDENTIFICATION OF ALL THE PROTEINS IN THE 50 S SUBUNIT OF AN ORGANELLE RIBOSOME (CHLOROPLAST). *Journal of Biological Chemistry* **275**, 28466-28482, (2000).
- 76 Yamaguchi, K. & Subramanian, A. R. Proteomic identification of all plastid-specific ribosomal proteins in higher plant chloroplast 30S ribosomal subunit. *European Journal of Biochemistry* **270**, 190-205, (2003).
- 77 Rogalski, M., Karcher, D. & Bock, R. Superwobbling facilitates translation with reduced tRNA sets. *Nat Struct Mol Biol* **15**, 192-198, (2008).
- 78 Alkatib, S., Fleischmann, T. T., Scharff, L. B. & Bock, R. Evolutionary constraints on the plastid tRNA set decoding methionine and isoleucine. *Nucleic Acids Research* **40**, 6713-6724, (2012).

- 79 Chua, N., Blobel, G., Siekievitz, G. & Palade, G. Periodic variations in the ratio of free to thylakoid-bound chloroplast ribosomes during the cell cycle of *Chlamydomonas reinhardtii*. *The Journal of cell biology* **71**, 497-514, (1976).
- 80 Friemann, A. & Hachtel, W. Chloroplast messenger RNAs of free and thylakoid-bound polysomes from *Vicia faba* L. *Planta* **175**, 50-59, (1988).
- 81 Zoschke, R. & Barkan, A. Genome-wide analysis of thylakoid-bound ribosomes in maize reveals principles of cotranslational targeting to the thylakoid membrane. *Proceedings of the National Academy of Sciences of the United States of America* **112**, E1678-E1687, (2015).
- 82 Schottkowski, M. *et al.* Biogenic membranes of the chloroplast in *Chlamydomonas reinhardtii*. (2012).
- 83 Uniacke, J. & Zerges, W. Photosystem II assembly and repair are differentially localized in *Chlamydomonas*. *Plant Cell* **19**, 3640-3654, (2007).
- 84 Uniacke, J. & Zerges, W. Chloroplast protein targeting involves localized translation in *Chlamydomonas*. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 1439-1444, (2009).
- 85 Castandet, B. & Araya, A. RNA editing in plant organelles. Why make it easy? *Biochemistry (Moscow)* **76**, 924, (2011).
- 86 Woodson, J. D. & Chory, J. Coordination of gene expression between organellar and nuclear genomes. *Nature reviews. Genetics* **9**, 383-395, (2008).
- 87 Watkins, K. P. *et al.* APO1 Promotes the Splicing of Chloroplast Group II Introns and Harbors a Plant-Specific Zinc-Dependent RNA Binding Domain. *The Plant Cell* **23**, 1082-1092, (2011).
- 88 Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nature reviews. Molecular cell biology* **8**, 479-490, (2007).
- 89 Ruwe, H., Kupsch, C., Teubner, M. & Schmitz-Linneweber, C. The RNA-recognition motif in chloroplasts. *Journal of Plant Physiology* **168**, 1361-1371, (2011).
- 90 Jacobs, J. & Kück, U. Function of chloroplast RNA-binding proteins. *Cellular and Molecular Life Sciences* **68**, 735-748, (2011).
- 91 Hammani, K. *et al.* Helical repeats modular proteins are major players for organelle gene expression. *Biochimie* **100**, 141-150, (2014).
- 92 Main, E. R., Lowe, A. R., Mochrie, S. G., Jackson, S. E. & Regan, L. A recurring theme in protein engineering: the design, stability and folding of repeat proteins. *Current opinion in structural biology* **15**, 464-471, (2005).
- 93 Bohne, A. V., Schwenkert, S., Grimm, B. & Nickelsen, J. Roles of Tetratricopeptide Repeat Proteins in Biogenesis of the Photosynthetic Apparatus. *International review of cell and molecular biology* **324**, 187-227, (2016).
- 94 Blatch, G. L. & Lassle, M. The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *BioEssays : news and reviews in molecular, cellular and developmental biology* **21**, 932-939, (1999).
- 95 D'Andrea, L. D. & Regan, L. TPR proteins: the versatile helix. *Trends Biochem Sci* **28**, 655-662, (2003).
- 96 Sane, A. P., Stein, B. & Westhoff, P. 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*. *The Plant journal : for cell and molecular biology* **42**, 720-730, (2005).
- 97 Chung, S. *et al.* Crooked neck is a component of the human spliceosome and implicated in the splicing process. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1576**, 287-297, (2002).
- 98 Preker, P. J. & Keller, W. The HAT helix, a repetitive motif implicated in RNA processing. *Trends in Biochemical Sciences* **23**, 15-16, (1998).

- 99 Hammani, K., Cook, W. B. & Barkan, A. RNA binding and RNA remodeling activities of the half-a-tetratricopeptide (HAT) protein HCF107 underlie its effects on gene expression. *Proc Natl Acad Sci U S A* **109**, 5651-5656, (2012).
- 100 Small, I. D. & Peeters, N. The PPR motif - a TPR-related motif prevalent in plant organellar proteins. *Trends Biochem Sci* **25**, 46-47, (2000).
- 101 Lurin, C. *et al.* Genome-Wide Analysis of Arabidopsis Pentatricopeptide Repeat Proteins Reveals Their Essential Role in Organelle Biogenesis. *The Plant Cell* **16**, 2089-2103, (2004).
- 102 Colcombet, J. *et al.* Systematic study of subcellular localization of Arabidopsis PPR proteins confirms a massive targeting to organelles. *RNA Biology* **10**, 1557-1575, (2013).
- 103 Cazalet, C. *et al.* Analysis of the *Legionella longbeachae* genome and transcriptome uncovers unique strategies to cause Legionnaires' disease. *PLoS Genet* **6**, e1000851, (2010).
- 104 Tourasse, N. J., Choquet, Y. & Vallon, O. PPR proteins of green algae. *RNA Biology* **10**, 1526-1542, (2013).
- 105 O'Toole, N. *et al.* On the Expansion of the Pentatricopeptide Repeat Gene Family in Plants. *Molecular Biology and Evolution* **25**, 1120-1128, (2008).
- 106 Shikanai, T. & Fujii, S. Function of PPR proteins in plastid gene expression. *RNA Biol* **10**, 1446-1456, (2013).
- 107 Barkan, A. & Small, I. Pentatricopeptide repeat proteins in plants. *Annual review of plant biology* **65**, 415-442, (2014).
- 108 Beick, S., Schmitz-Linneweber, C., Williams-Carrier, R., Jensen, B. & Barkan, A. The pentatricopeptide repeat protein PPR5 stabilizes a specific tRNA precursor in maize chloroplasts. *Molecular and cellular biology* **28**, 5337-5347, (2008).
- 109 Hattori, M. & Sugita, M. A moss pentatricopeptide repeat protein binds to the 3' end of plastid clpP pre-mRNA and assists with mRNA maturation. *The FEBS journal* **276**, 5860-5869, (2009).
- 110 Schmitz-Linneweber, C., Williams-Carrier, R. & Barkan, A. RNA Immunoprecipitation and Microarray Analysis Show a Chloroplast Pentatricopeptide Repeat Protein to Be Associated with the 5' Region of mRNAs Whose Translation It Activates. *The Plant Cell* **17**, 2791-2804, (2005).
- 111 Zoschke, R., Watkins, K. P. & Barkan, A. A Rapid Ribosome Profiling Method Elucidates Chloroplast Ribosome Behavior in Vivo. *The Plant Cell* **25**, 2265-2275, (2013).
- 112 Barkan, A. *et al.* A combinatorial amino acid code for RNA recognition by pentatricopeptide repeat proteins. *PLoS Genet* **8**, e1002910, (2012).
- 113 Barkan, A. *et al.* A Combinatorial Amino Acid Code for RNA Recognition by Pentatricopeptide Repeat Proteins. *PLoS genetics* **8**, e1002910, (2012).
- 114 Shen, C. *et al.* Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. **7**, 11285, (2016).
- 115 Yin, P. *et al.* Structural basis for the modular recognition of single-stranded RNA by PPR proteins. *Nature* **504**, 168-171, (2013).
- 116 Liu, S., Melonek, J., Boykin, L. M., Small, I. & Howell, K. A. PPR-SMRs: Ancient proteins with enigmatic functions. *RNA Biology* **10**, 1501-1510, (2013).
- 117 Meierhoff, K., Felder, S., Nakamura, T., Bechtold, N. & Schuster, G. HCF152, an Arabidopsis RNA Binding Pentatricopeptide Repeat Protein Involved in the Processing of Chloroplast psbB-psbT-psbH-petB-petD RNAs. *The Plant Cell* **15**, 1480-1495, (2003).
- 118 Schmitz-Linneweber, C. *et al.* A Pentatricopeptide Repeat Protein Facilitates the trans-Splicing of the Maize Chloroplast rps12 Pre-mRNA. *The Plant Cell* **18**, 2650-2663, (2006).
- 119 Aryamanesh, N. *et al.* The Pentatricopeptide Repeat Protein EMB2654 Is Essential for Trans-Splicing of a Chloroplast Small Ribosomal Subunit Transcript. *Plant Physiology* **173**, 1164-1176, (2017).

- 120 Pfalz, J., Bayraktar, O. A., Prikryl, J. & Barkan, A. Site-specific binding of a PPR protein defines and stabilizes 5' and 3' mRNA termini in chloroplasts. *The EMBO journal* **28**, 2042-2052, (2009).
- 121 Prikryl, J., Rojas, M., Schuster, G. & Barkan, A. Mechanism of RNA stabilization and translational activation by a pentatricopeptide repeat protein. *Proc Natl Acad Sci U S A* **108**, 415-420, (2011).
- 122 Zoschke, R. *et al.* The pentatricopeptide repeat-SMR protein ATP4 promotes translation of the chloroplast atpB/E mRNA. *The Plant Journal* **72**, 547-558, (2012).
- 123 Zoschke, R., Qu, Y., Zubo, Y. O., Börner, T. & Schmitz-Linneweber, C. Mutation of the pentatricopeptide repeat-SMR protein SVR7 impairs accumulation and translation of chloroplast ATP synthase subunits in *Arabidopsis thaliana*. *Journal of Plant Research* **126**, 403-414, (2013).
- 124 Rüdinger, M., Fritz-Laylin, L., Polsakiewicz, M. & Knoop, V. Plant-type mitochondrial RNA editing in the protist *Naegleria gruberi*. *RNA* **17**, 2058-2062, (2011).
- 125 Okuda, K. *et al.* Quantitative analysis of motifs contributing to the interaction between PLS-subfamily members and their target RNA sequences in plastid RNA editing. *The Plant Journal* **80**, 870-882, (2014).
- 126 Ichinose, M. & Sugita, M. RNA Editing and Its Molecular Mechanism in Plant Organelles. *Genes* **8**, 5, (2017).
- 127 Rivals, E., Bruyère, C., Toffano-Nioche, C. & Lecharny, A. Formation of the Arabidopsis Pentatricopeptide Repeat Family. *Plant Physiology* **141**, 825-839, (2006).
- 128 Salone, V. *et al.* A hypothesis on the identification of the editing enzyme in plant organelles. *FEBS Letters* **581**, 4132-4138, (2007).
- 129 Okuda, K. *et al.* Pentatricopeptide Repeat Proteins with the DYW Motif Have Distinct Molecular Functions in RNA Editing and RNA Cleavage in Arabidopsis Chloroplasts. *The Plant Cell* **21**, 146-156, (2009).
- 130 Fujii, S. & Small, I. The evolution of RNA editing and pentatricopeptide repeat genes. *New Phytologist* **191**, 37-47, (2011).
- 131 Jalal, A. *et al.* A Small Multifunctional Pentatricopeptide Repeat Protein in the Chloroplast of *Chlamydomonas reinhardtii*. *Molecular plant* **8**, 412-426, (2015).
- 132 Auchincloss, A. H., Zerges, W., Perron, K., Girard-Bascou, J. & Rochaix, J. D. Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in *Chlamydomonas reinhardtii*. *The Journal of cell biology* **157**, 953-962, (2002).
- 133 Boulouis, A. *et al.* Spontaneous dominant mutations in *chlamydomonas* highlight ongoing evolution by gene diversification. *Plant Cell* **27**, 984-1001, (2015).
- 134 Kleinknecht, L. *et al.* RAP, the Sole Octotricopeptide Repeat Protein in Arabidopsis, Is Required for Chloroplast 16S rRNA Maturation. *The Plant Cell* **26**, 777-787, (2014).
- 135 Rahire, M., Laroche, F., Cerutti, L. & Rochaix, J. D. Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I. *The Plant journal : for cell and molecular biology* **72**, 652-661, (2012).
- 136 Eberhard, S. *et al.* Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of atpA transcripts in *Chlamydomonas reinhardtii* chloroplasts. *The Plant journal : for cell and molecular biology* **67**, 1055-1066, (2011).
- 137 Balczun, C. *et al.* Two adjacent nuclear genes are required for functional complementation of a chloroplast trans-splicing mutant from *Chlamydomonas reinhardtii*. *The Plant Journal* **43**, 636-648, (2005).
- 138 Marx, C., Wünsch, C. & Kück, U. The Octatricopeptide Repeat Protein Raa8 Is Required for Chloroplast trans Splicing. *Eukaryotic Cell* **14**, 998-1005, (2015).

- 139 Merendino, L. *et al.* A novel multifunctional factor involved in trans-splicing of chloroplast introns in *Chlamydomonas*. *Nucleic Acids Research* **34**, 262-274, (2006).
- 140 Reifschneider, O. *et al.* A Ribonucleoprotein Supercomplex Involved in trans-Splicing of Organelle Group II Introns. *The Journal of biological chemistry* **291**, 23330-23342, (2016).
- 141 Glanz, S., Jacobs, J., Kock, V., Mishra, A. & Kück, U. Raa4 is a trans-splicing factor that specifically binds chloroplast tscA intron RNA. *The Plant Journal* **69**, 421-431, (2012).
- 142 Wang, F. *et al.* Two *Chlamydomonas* OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b6 f. *The Plant journal : for cell and molecular biology* **82**, 861-873, (2015).
- 143 Simarro, M. *et al.* FAST KINASE DOMAIN-CONTAINING PROTEIN 3 IS A MITOCHONDRIAL PROTEIN ESSENTIAL FOR CELLULAR RESPIRATION. *Biochemical and biophysical research communications* **401**, 440-446, (2010).
- 144 Boehm, E. *et al.* FASTKD1 and FASTKD4 have opposite effects on expression of specific mitochondrial RNAs, depending upon their endonuclease-like RAP domain. *Nucleic Acids Research* **45**, 6135-6146, (2017).
- 145 Kruse, B., Narasimhan, N. & Attardi, G. Termination of transcription in human mitochondria: Identification and purification of a DNA binding protein factor that promotes termination. *Cell* **58**, 391-397, (1989).
- 146 Peralta, S., Wang, X. & Moraes, C. T. Mitochondrial transcription: lessons from mouse models. *Biochimica et biophysica acta* **1819**, 961-969, (2012).
- 147 Camara, Y. *et al.* MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome. *Cell metabolism* **13**, 527-539, (2011).
- 148 Schönfeld, C. *et al.* The Nucleus-encoded Protein MOC1 Is Essential for Mitochondrial Light Acclimation in *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry* **279**, 50366-50374, (2004).
- 149 Linder, T. *et al.* A family of putative transcription termination factors shared amongst metazoans and plants. *Curr Genet* **48**, 265-269, (2005).
- 150 Yakubovskaya, E., Mejia, E., Byrnes, J., Hambardjiev, E. & Garcia-Diaz, M. Helix unwinding and base flipping enable human MTERF1 to terminate mitochondrial transcription. *Cell* **141**, 982-993, (2010).
- 151 Robinson, E. H. & Eichman, B. F. Nucleic acid recognition by tandem helical repeats. *Current opinion in structural biology* **22**, 101-109, (2012).
- 152 Babiychuk, E. *et al.* Plastid gene expression and plant development require a plastidic protein of the mitochondrial transcription termination factor family. *Proc Natl Acad Sci U S A* **108**, 6674-6679, (2011).
- 153 Hammani, K. & Barkan, A. An mTERF domain protein functions in group II intron splicing in maize chloroplasts. *Nucleic Acids Res* **42**, 5033-5042, (2014).
- 154 Meskauskiene, R. *et al.* A mutation in the Arabidopsis mTERF-related plastid protein SOLDAT10 activates retrograde signaling and suppresses (1)O(2)-induced cell death. *The Plant journal : for cell and molecular biology* **60**, 399-410, (2009).
- 155 Romani, I. *et al.* A Member of the Arabidopsis Mitochondrial Transcription Termination Factor Family Is Required for Maturation of Chloroplast Transfer RNA(Ile)(GAU). *Plant Physiology* **169**, 627-646, (2015).
- 156 Majeran, W. *et al.* Nucleoid-Enriched Proteomes in Developing Plastids and Chloroplasts from Maize Leaves: A New Conceptual Framework for Nucleoid Functions. *Plant Physiology* **158**, 156-189, (2012).
- 157 Delannoy, E., Stanley, W. A., Bond, C. S. & Small, I. D. Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochemical Society transactions* **35**, 1643-1647, (2007).
- 158 Schmitz-Linneweber, C. & Small, I. Pentatricopeptide repeat proteins: a socket set for organelle gene expression. *Trends in Plant Science* **13**, 663-670, (2008).

- 159 Chen, Y. & Varani, G. Engineering RNA-binding proteins for biology. *FEBS Journal* **280**, 3734-3754, (2013).
- 160 Tanaka Hall, T. M. De-coding and Re-coding RNA recognition by PUF and PPR repeat proteins. *Current opinion in structural biology* **36**, 116-121, (2016).
- 161 Zhelyazkova, P. *et al.* Protein-mediated protection as the predominant mechanism for defining processed mRNA termini in land plant chloroplasts. *Nucleic Acids Research* **40**, 3092-3105, (2012).
- 162 Zoschke, R., Watkins, K. P., Miranda, R. G. & Barkan, A. The PPR-SMR protein PPR53 enhances the stability and translation of specific chloroplast RNAs in maize. *The Plant journal : for cell and molecular biology* **85**, 594-606, (2016).
- 163 Wostrikoff, K., Choquet, Y., Wollman, F. A. & Girard-Bascou, J. TCA1, a single nuclear-encoded translational activator specific for petA mRNA in *Chlamydomonas reinhardtii* chloroplast. *Genetics* **159**, 119-132, (2001).
- 164 Loiselay, C. *et al.* Molecular identification and function of cis- and trans-acting determinants for petA transcript stability in *Chlamydomonas reinhardtii* chloroplasts. *Molecular and cellular biology* **28**, 5529-5542, (2008).
- 165 Dauvillée, D., Stampacchia, O., Girard-Bascou, J. & Rochaix, J.-D. Tab2 is a novel conserved RNA binding protein required for translation of the chloroplast psaB mRNA. *The EMBO Journal* **22**, 6378-6388, (2003).
- 166 Schwarz, C., Elles, I., Kortmann, J., Piotrowski, M. & Nickelsen, J. Synthesis of the D2 Protein of Photosystem II in *Chlamydomonas* Is Controlled by a High Molecular Mass Complex Containing the RNA Stabilization Factor Nac2 and the Translational Activator RBP40. *The Plant Cell* **19**, 3627-3639, (2007).
- 167 Raynaud, C. *et al.* Evidence for regulatory function of nucleus-encoded factors on mRNA stabilization and translation in the chloroplast. *Proceedings of the National Academy of Sciences* **104**, 9093-9098, (2007).
- 168 Douchi, D. *et al.* A Nucleus-Encoded Chloroplast Phosphoprotein Governs Expression of the Photosystem I Subunit PsaC in *Chlamydomonas reinhardtii*. *Plant Cell* **28**, 1182-1199, (2016).
- 169 Lefebvre-Legendre, L. *et al.* A Nucleus-Encoded Chloroplast Protein Regulated by Iron Availability Governs Expression of the Photosystem I Subunit PsaA in *Chlamydomonas reinhardtii*. *Plant Physiology* **167**, 1527-1540, (2015).
- 170 Feng, P. *et al.* Chloroplast retrograde signal regulates flowering. *Proceedings of the National Academy of Sciences* **113**, 10708-10713, (2016).
- 171 Bradbeer, J. W., Atkinson, Y. E., Borner, T. & Hagemann, R. Cytoplasmic synthesis of plastid polypeptides may be controlled by plastid-synthesised RNA. *Nature* **279**, 816-817, (1979).
- 172 Chan, K. X., Phua, S. Y., Crisp, P., McQuinn, R. & Pogson, B. J. Learning the Languages of the Chloroplast: Retrograde Signaling and Beyond. *Annual Review of Plant Biology* **67**, 25-53, (2016).
- 173 Woodson, J. D. Chloroplast quality control – balancing energy production and stress. *New Phytologist* **212**, 36-41, (2016).
- 174 Susek, R. E., Ausubel, F. M. & Chory, J. Signal transduction mutants of arabidopsis uncouple nuclear CAB and RBCS gene expression from chloroplast development. *Cell* **74**, 787-799, (1993).
- 175 Koussevitzky, S. *et al.* Signals from Chloroplasts Converge to Regulate Nuclear Gene Expression. *Science* **316**, 715-719, (2007).
- 176 Chi, W., Sun, X. & Zhang, L. Intracellular Signaling from Plastid to Nucleus. *Annual Review of Plant Biology* **64**, 559-582, (2013).
- 177 Xiao, Y. *et al.* Retrograde Signaling by the Plastidial Metabolite MEcPP Regulates Expression of Nuclear Stress-Response Genes. *Cell* **149**, 1525-1535, (2012).

- 178 Guo, H. *et al.* Plastid-nucleus communication involves calcium-modulated MAPK signalling. *Nature Communications* **7**, 12173, (2016).
- 179 Perron, K., Goldschmidt-Clermont, M. & Rochaix, J. D. A factor related to pseudouridine synthases is required for chloroplast group II intron trans-splicing in *Chlamydomonas reinhardtii*. *The EMBO journal* **18**, 6481-6490, (1999).
- 180 Rivier, C., Goldschmidt-Clermont, M. & Rochaix, J. D. Identification of an RNA-protein complex involved in chloroplast group II intron trans-splicing in *Chlamydomonas reinhardtii*. *The EMBO journal* **20**, 1765-1773, (2001).
- 181 Lefebvre-Legendre, L. *et al.* A pioneer protein is part of a large complex involved in trans-splicing of a group II intron in the chloroplast of *Chlamydomonas reinhardtii*. *The Plant Journal* **85**, 57-69, (2016).
- 182 Higgs, D. C., Kuras, R., Kindle, K. L., Wollman, F. A. & Stern, D. B. Inversions in the *Chlamydomonas* chloroplast genome suppress a petD 5' untranslated region deletion by creating functional chimeric mRNAs. *The Plant journal : for cell and molecular biology* **14**, 663-671, (1998).
- 183 Murakami, S., Kuehnle, K. & Stern, D. B. 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, (2005).
- 184 Ossenbuhl, F., Hartmann, K. & Nickelsen, J. A chloroplast RNA binding protein from stromal thylakoid membranes specifically binds to the 5' untranslated region of the psbA mRNA. *Eur J Biochem* **269**, 3912-3919, (2002).
- 185 Monod, C., Goldschmidt-Clermont, M. & Rochaix, J. D. Accumulation of chloroplast psbB RNA requires a nuclear factor in *Chlamydomonas reinhardtii*. *Molecular & general genetics : MGG* **231**, 449-459, (1992).
- 186 Nickelsen, J., van Dillewijn, J., Rahire, M. & Rochaix, J. D. Determinants for stability of the chloroplast psbD RNA are located within its short leader region in *Chlamydomonas reinhardtii*. *The EMBO journal* **13**, 3182-3191, (1994).
- 187 Vaistij, F. E., Goldschmidt-Clermont, M., Wostrikoff, K. & Rochaix, J. D. Stability determinants in the chloroplast psbB/T/H mRNAs of *Chlamydomonas reinhardtii*. *The Plant journal : for cell and molecular biology* **21**, 469-482, (2000).
- 188 Komenda, J. *et al.* Accumulation of the D2 Protein Is a Key Regulatory Step for Assembly of the Photosystem II Reaction Center Complex in *Synechocystis* PCC 6803. *Journal of Biological Chemistry* **279**, 48620-48629, (2004).
- 189 Takahashi, Y., Matsumoto H Fau - Goldschmidt-Clermont, M., Goldschmidt-Clermont M Fau - Rochaix, J. D. & Rochaix, J. D. Directed disruption of the *Chlamydomonas* chloroplast psbK gene destabilizes the photosystem II reaction center complex. (1994).
- 190 Ikeuchi, M., Shukla Vk Fau - Pakrasi, H. B., Pakrasi Hb Fau - Inoue, Y. & Inoue, Y. Directed inactivation of the psbI gene does not affect photosystem II in the cyanobacterium *Synechocystis* sp. PCC 6803. (1995).
- 191 Komenda, J., Lupínková, L. & Kopecký, J. Absence of the psbH gene product destabilizes photosystem II complex and bicarbonate binding on its acceptor side in *Synechocystis* PCC 6803. *European Journal of Biochemistry* **269**, 610-619, (2002).
- 192 Mayers, S. R. *et al.* Further characterization of the psbH locus of *Synechocystis* sp. PCC 6803: inactivation of psbH impairs QA to QB electron transport in photosystem 2. (1993).
- 193 van Wijk, K. J. Protein Maturation and Proteolysis in Plant Plastids, Mitochondria, and Peroxisomes. *Annual Review of Plant Biology* **66**, 75-111, (2015).
- 194 Nishimura, K., Kato, Y. & Sakamoto, W. Essentials of Proteolytic Machineries in Chloroplasts. *Molecular plant* **10**, 4-19, (2017).

- 195 Schmidt, G. W. & Mishkind, M. L. Rapid degradation of unassembled ribulose 1,5-
bisphosphate carboxylase small subunits in chloroplasts. *Proceedings of the National
Academy of Sciences* **80**, 2632-2636, (1983).
- 196 de Vitry, C., Olive, J., Drapier, D., Recouvreur, M. & Wollman, F. Posttranslational events
leading to the assembly of photosystem II protein complex: a study using photosynthesis
mutants from *Chlamydomonas reinhardtii*. *The Journal of cell biology* **109**, 991-1006,
(1989).
- 197 Kuras, R. & Wollman, F. A. The assembly of cytochrome b6/f complexes: an approach
using genetic transformation of the green alga *Chlamydomonas reinhardtii*. (1994).
- 198 Girard-Bascou, J., Choquet, Y., Schneider, M., Delosme, M. & Dron, M. Characterization
of a chloroplast mutation in the *psaA2* gene of *Chlamydomonas reinhardtii*. *Current
genetics* **12**, 489-495, (1987).
- 199 Stampacchia, O. *et al.* A nuclear-encoded function essential for translation of the
chloroplast *psaB* mRNA in *chlamydomonas*. *Plant Cell* **9**, 773-782, (1997).
- 200 Takahashi, Y., Goldschmidt-Clermont, M., Soen, S. Y., Franzen, L. G. & Rochaix, J. D.
Directed chloroplast transformation in *Chlamydomonas reinhardtii*: insertional
inactivation of the *psaC* gene encoding the iron sulfur protein destabilizes photosystem
I. *EMBO J* **10**, 2033-2040, (1991).
- 201 Bennoun, P. *et al.* Characterization of photosystem II mutants of *Chlamydomonas
reinhardtii* lacking the *psbA* gene. *Plant molecular biology* **6**, 151-160, (1986).
- 202 Erickson, J. M. *et al.* Lack of the D2 protein in a *Chlamydomonas reinhardtii* *psbD* mutant
affects photosystem II stability and D1 expression. *EMBO J* **5**, 1745-1754, (1986).
- 203 Wostrikoff, K., Girard-Bascou, J., Wollman, F. A. & Choquet, Y. Biogenesis of PSI involves
a cascade of translational autoregulation in the chloroplast of
Chlamydomonas. *The EMBO Journal* **23**, 2696, (2004).
- 204 Minai, L., Wostrikoff, K., Wollman, F. A. & Choquet, Y. Chloroplast biogenesis of
photosystem II cores involves a series of assembly-controlled steps that regulate
translation. *Plant Cell* **18**, 159-175, (2006).
- 205 Drapier, D., Rimbault, B., Vallon, O., Wollman, F.-A. & Choquet, Y. Intertwined
translational regulations set uneven stoichiometry of chloroplast ATP synthase subunits.
The EMBO Journal **26**, 3581-3591, (2007).
- 206 Khrebtukova, I. & Spreitzer, R. J. Elimination of the *Chlamydomonas* gene family that
encodes the small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase.
Proceedings of the National Academy of Sciences of the United States of America **93**,
13689-13693, (1996).
- 207 Choquet, Y. *et al.* Translation of cytochrome f is autoregulated through the 5'
untranslated region of *petA* mRNA in *Chlamydomonas* chloroplasts. *Proceedings of the
National Academy of Sciences of the United States of America* **95**, 4380-4385, (1998).
- 208 Choquet, Y., Zito, F., Wostrikoff, K. & Wollman, F.-A. Cytochrome f Translation in
Chlamydomonas Chloroplast Is Autoregulated by its Carboxyl-Terminal Domain. *The
Plant Cell* **15**, 1443-1454, (2003).
- 209 Kuras, R., Buschlen, S. & Wollman, F. A. Maturation of pre-apocytochrome f in vivo. A
site-directed mutagenesis study in *Chlamydomonas reinhardtii*. *The Journal of biological
chemistry* **270**, 27797-27803, (1995).
- 210 Boulouis, A. *et al.* The Nucleus-Encoded trans-Acting Factor MCA1 Plays a Critical Role in
the Regulation of Cytochrome f Synthesis in *Chlamydomonas* Chloroplasts. *The Plant Cell*
23, 333-349, (2011).
- 211 Ketchner, S. L. *et al.* Chloroplasts can accommodate inclusion bodies. Evidence from a
mutant of *Chlamydomonas reinhardtii* defective in the assembly of the chloroplast ATP
synthase. *The Journal of biological chemistry* **270**, 15299-15306., (1995).

- 212 Rodermel, S., Haley, J., Jiang, C. Z., Tsai, C. H. & Bogorad, L. A mechanism for
intergenomic integration: abundance of ribulose biphosphate carboxylase small-subunit
protein influences the translation of the large-subunit mRNA. *Proceedings of the
National Academy of Sciences of the United States of America* **93**, 3881-3885, (1996).
- 213 Wostrikoff, K. & Stern, D. Rubisco large-subunit translation is autoregulated in response
to its assembly state in tobacco chloroplasts. *Proceedings of the National Academy of
Sciences of the United States of America* **104**, 6466-6471, (2007).
- 214 Wostrikoff, K., Clark, A., Sato, S., Clemente, T. & Stern, D. Ectopic Expression of Rubisco
Subunits in Maize Mesophyll Cells Does Not Overcome Barriers to Cell Type-Specific
Accumulation. *Plant Physiology* **160**, 419-432, (2012).
- 215 Gamble, P. E. & Mullet, J. E. Translation and stability of proteins encoded by the plastid
psbA and psbB genes are regulated by a nuclear gene during light-induced chloroplast
development in barley. *The Journal of biological chemistry* **264**, 7236-7243, (1989).
- 216 Yosef, I. *et al.* RNA binding activity of the ribulose-1,5-bisphosphate
carboxylase/oxygenase large subunit from *Chlamydomonas reinhardtii*. *The Journal of
biological chemistry* **279**, 10148-10156, (2004).
- 217 Drapier, D., Girard-Bascou, J. & Wollman, F. A. Evidence for Nuclear Control of the
Expression of the atpA and atpB Chloroplast Genes in *Chlamydomonas*. *The Plant Cell* **4**,
283-295, (1992).
- 218 Wildman, S. & Bonner, J. The proteins of green leaves; isolation, enzymatic properties
and auxin content of spinach cytoplasmic proteins. *Arch Biochemistry* **3**, 381-413.,
(1947).
- 219 Nisbet, E. G. *et al.* The age of Rubisco: the evolution of oxygenic photosynthesis.
Geobiology **5**, 311-335, (2007).
- 220 Dubbs, J. M. & Robert Tabita, F. Regulators of nonsulfur purple phototrophic bacteria
and the interactive control of CO₂ assimilation, nitrogen fixation, hydrogen metabolism
and energy generation. *FEMS Microbiology Reviews* **28**, 353-376, (2004).
- 221 Sato, T., Atomi, H. & Imanaka, T. Archaeal type III RuBisCOs function in a pathway for
AMP metabolism. *Science* **315**, 1003-1006, (2007).
- 222 Gunn, L. H., Vålegård, K. & Andersson, I. A unique structural domain in
Methanococcoides burtonii ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco)
acts as a small subunit mimic. *The Journal of biological chemistry* **292**, 6838-6850,
(2017).
- 223 Tabita, F. R., Satagopan, S., Hanson, T. E., Kreel, N. E. & Scott, S. S. Distinct form I, II, III,
and IV Rubisco proteins from the three kingdoms of life provide clues about Rubisco
evolution and structure/function relationships. *Journal of experimental botany* **59**, 1515-
1524, (2008).
- 224 Andrews, T. J. Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large
subunits in the complete absence of small subunits. *The Journal of biological chemistry*
263, 12213-12219, (1988).
- 225 Lee, B., Berka, R. M. & Tabita, F. R. Mutations in the small subunit of cyanobacterial
ribulose-bisphosphate carboxylase/oxygenase that modulate interactions with large
subunits. *Journal of Biological Chemistry* **266**, 7417-7422, (1991).
- 226 Ishikawa, C., Hatanaka, T., Misoo, S., Miyake, C. & Fukayama, H. Functional Incorporation
of Sorghum Small Subunit Increases the Catalytic Turnover Rate of Rubisco in Transgenic
Rice. *Plant Physiology* **156**, 1603-1611, (2011).
- 227 Meyer, M. T. *et al.* Rubisco small-subunit alpha-helices control pyrenoid formation in
Chlamydomonas. *Proceedings of the National Academy of Sciences of the United States
of America* **109**, 19474-19479, (2012).

- 228 Taylor, T. C., Backlund, A., Bjorhall, K., Spreitzer, R. J. & Andersson, I. First Crystal Structure of Rubisco from a Green Alga, *Chlamydomonas reinhardtii*. *Journal of Biological Chemistry* **276**, 48159-48164, (2001).
- 229 Chen, Z. X. & Spreitzer, R. J. Chloroplast intragenic suppression enhances the low CO₂/O₂ specificity of mutant ribulose-bisphosphate carboxylase/oxygenase. *The Journal of biological chemistry* **264**, 3051-3053, (1989).
- 230 Laing, W. A. & Christeller, J. T. A model for the kinetics of activation and catalysis of ribulose 1,5-bisphosphate carboxylase. *Biochemical Journal* **159**, 563-570, (1976).
- 231 Taylor, T. C. & Andersson, I. Structure of a product complex of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase. *Biochemistry* **36**, 4041-4046, (1997).
- 232 Sharwood, R. E., Ghannoum, O. & Whitney, S. M. Prospects for improving CO₂ fixation in C₃-crops through understanding C₄-Rubisco biogenesis and catalytic diversity. *Current opinion in plant biology* **31**, 135-142, (2016).
- 233 Whitney, S. M., Birch, R., Kelso, C., Beck, J. L. & Kapralov, M. V. Improving recombinant Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1 chaperone. *Proceedings of the National Academy of Sciences of the United States of America*, (2015).
- 234 Wilson, R. & Whitney, S. Photosynthesis: Getting it together for CO₂ fixation. *Nature Plants* **1**, 15130, (2015).
- 235 Lee, D. W. *et al.* Functional Characterization of Sequence Motifs in the Transit Peptide of Arabidopsis Small Subunit of Rubisco. *Plant Physiology* **140**, 466-483, (2006).
- 236 Lee, J., Wang, F. & Schnell, D. J. Toc Receptor Dimerization Participates in the Initiation of Membrane Translocation during Protein Import into Chloroplasts. *The Journal of biological chemistry* **284**, 31130-31141, (2009).
- 237 Pilon, M. *et al.* Expression in Escherichia coli and purification of a translocation-competent precursor of the chloroplast protein ferredoxin. *Journal of Biological Chemistry* **265**, 3358-3361, (1990).
- 238 Jarvis, P. & Soll, J. Toc, Tic, and chloroplast protein import. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1541**, 64-79, (2001).
- 239 Houtz, R. L., Magnani, R., Nayak, N. R. & Dirk, L. M. Co- and post-translational modifications in Rubisco: unanswered questions. *Journal of experimental botany* **59**, 1635-1645, (2008).
- 240 Rowland, E., Kim, J., Bhuiyan, N. H. & van Wijk, K. J. The Arabidopsis Chloroplast Stromal N-Terminome: Complexities of Amino-Terminal Protein Maturation and Stability. *Plant Physiology* **169**, 1881-1896, (2015).
- 241 Loschelder, H., Schweer, J., Link, B. & Link, G. Dual Temporal Role of Plastid Sigma Factor 6 in Arabidopsis Development. *Plant Physiology* **142**, 642-650, (2006).
- 242 Ishizaki, Y. *et al.* A nuclear-encoded sigma factor, Arabidopsis SIG6, recognizes sigma-70 type chloroplast promoters and regulates early chloroplast development in cotyledons. *The Plant Journal* **42**, 133-144, (2005).
- 243 Shiina, T., Allison, L. & Maliga, P. *rbcl* Transcript levels in tobacco plastids are independent of light: reduced dark transcription rate is compensated by increased mRNA stability. *The Plant Cell* **10**, 1713-1722, (1998).
- 244 Klein, U., Salvador, M. L. & Bogorad, L. Activity of the *Chlamydomonas* chloroplast *rbcl* gene promoter is enhanced by a remote sequence element. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 10819-10823, (1994).
- 245 Kasai, S. *et al.* Effect of coding regions on chloroplast gene expression in *Chlamydomonas reinhardtii*. *Journal of Bioscience and Bioengineering* **95**, 276-282, (2003).

- 246 Salvador, M. L., Suay, L., Anthonisen, I. L. & Klein, U. Changes in the 5'-untranslated
region of the *rbcl* gene accelerate transcript degradation more than 50-fold in the
chloroplast of *Chlamydomonas reinhardtii*. *Current genetics* **45**, 176-182, (2004).
- 247 Johnson, X. Manipulating RuBisCO accumulation in the green alga, *Chlamydomonas*
reinhardtii. *Plant molecular biology* **76**, 397-405, (2011).
- 248 Drapier, D., Girard-Bascou, J., Stern, D. B. & Wollman, F.-A. A dominant nuclear mutation
in *Chlamydomonas* identifies a factor controlling chloroplast mRNA stability by acting on
the coding region of the *atpA* transcript. (2002).
- 249 Barkan, A. Nuclear Mutants of Maize with Defects in Chloroplast Polysome Assembly
Have Altered Chloroplast RNA Metabolism. *The Plant Cell* **5**, 389-402, (1993).
- 250 Mühlbauer, S. K. & Eichacker, L. A. The stromal protein large subunit of ribulose-1,5-
bisphosphate carboxylase is translated by membrane-bound ribosomes. *European*
Journal of Biochemistry **261**, 784-788, (1999).
- 251 Berry, J. O., Carr, J. P. & Klessig, D. F. mRNAs encoding ribulose-1,5-bisphosphate
carboxylase remain bound to polysomes but are not translated in amaranth seedlings
transferred to darkness. *Proceedings of the National Academy of Sciences* **85**, 4190-
4194, (1988).
- 252 Mühlbauer, S. K. & Eichacker, L. A. Light-dependent Formation of the Photosynthetic
Proton Gradient Regulates Translation Elongation in Chloroplasts. *Journal of Biological*
Chemistry **273**, 20935-20940, (1998).
- 253 Singh, B. N. *et al.* A pea chloroplast translation elongation factor that is regulated by
abiotic factors. *Biochemical and Biophysical Research Communications* **320**, 523-530,
(2004).
- 254 Rodermel, S. R., Abbott, M. S. & Bogorad, L. Nuclear-organelle interactions: Nuclear
antisense gene inhibits ribulose bisphosphate carboxylase enzyme levels in transformed
tobacco plants. *Cell* **55**, 673-681, (1988).
- 255 Irihimovitch, V. & Shapira, M. Glutathione Redox Potential Modulated by Reactive
Oxygen Species Regulates Translation of Rubisco Large Subunit in the Chloroplast.
Journal of Biological Chemistry **275**, 16289-16295, (2000).
- 256 Cohen, I., Knopf, J. A., Irihimovitch, V. & Shapira, M. A Proposed Mechanism for the
Inhibitory Effects of Oxidative Stress on Rubisco Assembly and Its Subunit Expression.
Plant Physiology **137**, 738-746, (2005).
- 257 Cohen, I., Sapir, Y. & Shapira, M. A Conserved Mechanism Controls Translation of
Rubisco Large Subunit in Different Photosynthetic Organisms. *Plant Physiology* **141**,
1089-1097, (2006).
- 258 Nordhues, A., Miller, S. M., Mühlhaus, T. & Schroda, M. New Insights into the Roles of
Molecular Chaperones in *Chlamydomonas* and *Volvox*. *International review of cell and*
molecular biology **285**, 75-113, (2010).
- 259 Trösch, R., Mühlhaus, T., Schroda, M. & Willmund, F. ATP-dependent molecular
chaperones in plastids — More complex than expected. *Biochimica et Biophysica Acta*
(*BBA*) - *Bioenergetics* **1847**, 872-888, (2015).
- 260 Goloubinoff, P., Christeller, J. T., Gatenby, A. A. & Lorimer, G. H. Reconstitution of active
dimeric ribulose bisphosphate carboxylase from an unfoiled state depends on two
chaperonin proteins and Mg-ATP. *Nature* **342**, 884-889, (1989).
- 261 Hayer-Hartl, M., Bracher, A. & Hartl, F. U. The GroEL-GroES Chaperonin Machine: A
Nano-Cage for Protein Folding. *Trends in biochemical sciences* **41**, 62-76, (2016).
- 262 Bai, C. *et al.* Protomer Roles in Chloroplast Chaperonin Assembly and Function.
Molecular plant **8**, 1478-1492, (2015).
- 263 Baneyx, F. *et al.* Spinach Chloroplast *cpn21* Co-chaperonin Possesses Two Functional
Domains Fused Together in a Toroidal Structure and Exhibits Nucleotide-dependent

- Binding to Plastid Chaperonin 60. *Journal of Biological Chemistry* **270**, 10695-10702, (1995).
- 264 Tsai, Y.-C. C., Mueller-Cajar, O., Saschenbrecker, S., Hartl, F. U. & Hayer-Hartl, M. Chaperonin Cofactors, Cpn10 and Cpn20, of Green Algae and Plants Function as Hetero-oligomeric Ring Complexes. *The Journal of biological chemistry* **287**, 20471-20481, (2012).
- 265 van der Vies, S. M., Bradley, D. & Gatenby, A. A. Assembly of cyanobacterial and higher plant ribulose biphosphate carboxylase subunits into functional homologous and heterologous enzyme molecules in *Escherichia coli*. *The EMBO Journal* **5**, 2439-2444, (1986).
- 266 Larimer, F. W. & Soper, T. S. Overproduction of *Anabaena* 7120 ribulose-bisphosphate carboxylase/oxygenase in *Escherichia coli*. *Gene* **126**, 85-92, (1993).
- 267 Barraclough, R. & Ellis, R. J. Protein synthesis in chloroplasts. IX. Assembly of newly-synthesized large subunits into ribulose biphosphate carboxylase in isolated intact pea chloroplasts. *Biochimica et biophysica acta* **1**, 13, (1980).
- 268 Kim, S.-R., Yang, J.-I. & An, G. OsCpn60 α 1, Encoding the Plastid Chaperonin 60 α Subunit, Is Essential for Folding of rbcL. *Molecules and Cells* **35**, 402-409, (2013).
- 269 Feiz, L. *et al.* Ribulose-1,5-bis-phosphate carboxylase/oxygenase accumulation factor1 is required for holoenzyme assembly in maize. *Plant Cell* **24**, 3435-3446, (2012).
- 270 Checa, S. K. & Viale, A. M. The 70-kDa Heat-Shock Protein/DnaK Chaperone System is Required for the Productive Folding of Ribulose-Bisphosphate Carboxylase Subunits in *Escherichia Coli*. *European Journal of Biochemistry* **248**, 848-855, (1997).
- 271 Brutnell, T. P., Sawers, R. J., Mant, A. & Langdale, J. A. BUNDLE SHEATH DEFECTIVE2, a novel protein required for post-translational regulation of the rbcL gene of maize. *The Plant Cell* **11**, 849-864, (1999).
- 272 Doron, L., Segal, N., Gibori, H. & Shapira, M. The BSD2 ortholog in *Chlamydomonas reinhardtii* is a polysome-associated chaperone that co-migrates on sucrose gradients with the rbcL transcript encoding the Rubisco large subunit. *The Plant journal : for cell and molecular biology*, (2014).
- 273 Saschenbrecker, S. *et al.* Structure and function of RbcX, an assembly chaperone for hexadecameric Rubisco. *Cell* **129**, 1189-1200, (2007).
- 274 Zhang, X.-H. *et al.* Hybrid Rubisco of tomato large subunits and tobacco small subunits is functional in tobacco plants. *Plant Science* **180**, 480-488, (2011).
- 275 Li, L. A. & Tabita, F. R. Maximum activity of recombinant ribulose 1,5-bisphosphate carboxylase/oxygenase of *Anabaena* sp. strain CA requires the product of the rbcX gene. *Journal of bacteriology* **179**, 3793-3796, (1997).
- 276 Onizuka, T. *et al.* The rbcX gene product promotes the production and assembly of ribulose-1,5-bisphosphate carboxylase/oxygenase of *Synechococcus* sp. PCC7002 in *Escherichia coli*. *Plant & cell physiology* **45**, 1390-1395, (2004).
- 277 Emlyn-Jones, D., Woodger, F. J., Price, G. D. & Whitney, S. M. RbcX can function as a rubisco chaperonin, but is non-essential in *Synechococcus* PCC7942. *Plant & cell physiology* **47**, 1630-1640, (2006).
- 278 Bracher, A., Hauser, T., Liu, C., Hartl, F. U. & Hayer-Hartl, M. Structural Analysis of the Rubisco-Assembly Chaperone RbcX-II from *Chlamydomonas reinhardtii*. *PloS one* **10**, e0135448, (2015).
- 279 Kolesinski, P., Piechota, J. & Szczepaniak, A. Initial characteristics of RbcX proteins from *Arabidopsis thaliana*. *Plant molecular biology* **77**, 447-459, (2011).
- 280 Bracher, A., Starling-Windhof, A., Hartl, F. U. & Hayer-Hartl, M. Crystal structure of a chaperone-bound assembly intermediate of form I Rubisco. *Nature structural & molecular biology* **18**, 875-880, (2011).

- 281 Kolesinski, P., Belusiak, I., Czarnocki-Cieciura, M. & Szczepaniak, A. Rubisco Accumulation
Factor 1 from *Thermosynechococcus elongatus* participates in the final stages of
ribulose-1,5-bisphosphate carboxylase/oxygenase assembly in *Escherichia coli* cells and
in vitro. *The FEBS journal*, (2014).
- 282 Whitney, S. M., Birch, R., Kelso, C., Beck, J. L. & Kapralov, M. V. Improving recombinant
Rubisco biogenesis, plant photosynthesis and growth by coexpressing its ancillary RAF1
chaperone. *Proceedings of the National Academy of Sciences of the United States of
America* **112**, 3564-3569, (2015).
- 283 Kanevski, I., Maliga, P., Rhoades, D. F. & Gutteridge, S. Plastome Engineering of Ribulose-
1,5-Bisphosphate Carboxylase/Oxygenase in Tobacco to Form a Sunflower Large Subunit
and Tobacco Small Subunit Hybrid. *Plant Physiology* **119**, 133-142, (1999).
- 284 Kolesinski, P., Rydzy, M. & Szczepaniak, A. Is RAF1 protein from *Synechocystis* sp. PCC
6803 really needed in the cyanobacterial Rubisco assembly process? *Photosynth Res*,
(2017).
- 285 Feiz, L. *et al.* A protein with an inactive pterin-4a-carbinolamine dehydratase domain is
required for Rubisco biogenesis in plants. *The Plant journal : for cell and molecular
biology*, (2014).
- 286 Wheatley, N. M., Sundberg, C. D., Gidaniyan, S. D., Cascio, D. & Yeates, T. O. Structure
and Identification of a Pterin Dehydratase-like Protein as a Ribulose-bisphosphate
Carboxylase/Oxygenase (RuBisCO) Assembly Factor in the α -Carboxysome. *The Journal
of biological chemistry* **289**, 7973-7981, (2014).
- 287 Smith, S. M. & Ellis, R. J. Processing of small subunit precursor of ribulose bisphosphate
carboxylase and its assembly into whole enzyme are stromal events. *Nature* **278**, 662-
664, (1979).
- 288 Grimm, R. *et al.* Postimport methylation of the small subunit of ribulose-1,5-
bisphosphate carboxylase in chloroplasts. *FEBS Letters* **408**, 350-354, (1997).
- 289 Ellis, R. J. & Van Der Vies, S. M. The Rubisco subunit binding protein. *Photosynth Res* **16**,
101-115, (1988).
- 290 Roy, H. & Andrews, T. J. in *Photosynthesis: Physiology and Metabolism* (eds Richard C.
Leegood, Thomas D. Sharkey, & Susanne von Caemmerer) 53-83 (Springer Netherlands,
2000).
- 291 Joshi, J., Mueller-Cajar, O., Tsai, Y.-C. C., Hartl, F. U. & Hayer-Hartl, M. Role of Small
Subunit in Mediating Assembly of Red-type Form I Rubisco. *Journal of Biological
Chemistry* **290**, 1066-1074, (2015).
- 292 Barkan, A. Expression of Plastid Genes: Organelle-Specific Elaborations on a Prokaryotic
Scaffold. *Plant Physiology* **155**, 1520-1532, (2011).
- 293 Dent, R. M., Haglund, C. M., Chin, B. L., Kobayashi, M. C. & Niyogi, K. K. Functional
genomics of eukaryotic photosynthesis using insertional mutagenesis of
Chlamydomonas reinhardtii. *Plant Physiol* **137**, 545-556, (2005).
- 294 Li, X. *et al.* An indexed, mapped mutant library enables reverse genetics studies of
biological processes in *Chlamydomonas reinhardtii*. *The Plant Cell*, (2016).
- 295 Johnson, X. *et al.* A new setup for in vivo fluorescence imaging of photosynthetic activity.
Photosynth Res **102**, 85-93, (2009).
- 296 Hong, S. & Spreitzer, R. J. Nuclear Mutation Inhibits Expression of the Chloroplast Gene
That Encodes the Large Subunit of Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase.
Plant Physiology **106**, 673-678, (1994).
- 297 Young, R. E. & Purton, S. Cytosine deaminase as a negative selectable marker for the
microalgal chloroplast: a strategy for the isolation of nuclear mutations that affect
chloroplast gene expression. *The Plant journal : for cell and molecular biology* **80**, 915-
925, (2014).

- 298 Merchant, S. S. *et al.* The Chlamydomonas Genome Reveals the Evolution of Key Animal
and Plant Functions. *Science (New York, N.Y.)* **318**, 245-250, (2007).
- 299 Goodstein, D. M. *et al.* Phytozome: a comparative platform for green plant genomics.
Nucleic Acids Research **40**, D1178-D1186, (2012).
- 300 Wang, H. *et al.* The Global Phosphoproteome of Chlamydomonas reinhardtii Reveals
Complex Organellar Phosphorylation in the Flagella and Thylakoid Membrane. *Molecular
& cellular proteomics : MCP* **13**, 2337-2353, (2014).
- 301 Majeran, W., Wollman, F.-A. & Vallon, O. Evidence for a Role of ClpP in the Degradation
of the Chloroplast Cytochrome b(6)f Complex. *The Plant Cell* **12**, 137-150, (2000).
- 302 Morita, K., Hatanaka, T., Misoo, S. & Fukayama, H. Unusual Small Subunit That Is Not
Expressed in Photosynthetic Cells Alters the Catalytic Properties of Rubisco in Rice. *Plant
Physiology* **164**, 69-79, (2014).
- 303 Börner, T. The discovery of plastid-to-nucleus retrograde signaling—a personal
perspective. *Protoplasma*, (2017).
- 304 Jarvis, P. Targeting of nucleus-encoded proteins to chloroplasts in plants. *New
Phytologist* **179**, 257-285, (2008).
- 305 Brutnell, T. P., Sawers Rj Fau - Mant, A., Mant A Fau - Langdale, J. A. & Langdale, J. A.
BUNDLE SHEATH DEFECTIVE2, a novel protein required for post-translational regulation
of the rbcL gene of maize. (1999).
- 306 Tarnawski, M., Gubernator, B., Kolesinski, P. & Szczepaniak, A. Heterologous expression
and initial characterization of recombinant RbcX protein from Thermosynechococcus
elongatus BP-1 and the role of RbcX in RuBisCO assembly. *Acta biochimica Polonica* **55**,
777-785, (2008).
- 307 Kolesinski, P. *et al.* Insights into eukaryotic Rubisco assembly - crystal structures of RbcX
chaperones from Arabidopsis thaliana. *Biochimica et biophysica acta* **1830**, 2899-2906,
(2013).
- 308 Liu, C. *et al.* Coupled chaperone action in folding and assembly of hexadecameric
Rubisco. *Nature* **463**, 197-202, (2010).
- 309 Hauser, T. *et al.* Structure and mechanism of the Rubisco-assembly chaperone Raf1.
Nature structural & molecular biology, (2015).
- 310 Roy, H., Bloom, M., Milos, P. & Monroe, M. Studies on the assembly of large subunits of
ribulose biphosphate carboxylase in isolated pea chloroplasts. *The Journal of cell
biology* **94**, 20-27, (1982).
- 311 Shapira, M. *et al.* Differential regulation of chloroplast gene expression in
Chlamydomonas reinhardtii during photoacclimation: light stress transiently suppresses
synthesis of the Rubisco LSU protein while enhancing synthesis of the PS II D1 protein.
Plant molecular biology **33**, 1001-1001, (1997).
- 312 Knopf, J. A. & Shapira, M. Degradation of Rubisco SSU during oxidative stress triggers
aggregation of Rubisco particles in Chlamydomonas reinhardtii. (2005).
- 313 Zhan, Y. *et al.* Localized control of oxidized RNA. *Journal of cell science* **128**, 4210-4219,
(2015).
- 314 Hubbs, A. & Roy, H. Synthesis and Assembly of Large Subunits into Ribulose
Biphosphate Carboxylase/Oxygenase in Chloroplast Extracts. *Plant Physiology* **100**, 272-
281, (1992).
- 315 Hubbs, A. E. & Roy, H. Assembly of in vitro synthesized large subunits into ribulose-
biphosphate carboxylase/oxygenase. Formation and discharge of an L8-like species.
Journal of Biological Chemistry **18**, 6, (1993).
- 316 Harris, E. H. 2000 (Academic Press, 2009).
- 317 Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular cloning: a laboratory manual*. (Cold
spring harbor laboratory press, 1989).

- 318 Fischer, N., Stampacchia, O., Redding, K. & Rochaix, J.-D. Selectable marker recycling in
the chloroplast. *Molecular and General Genetics MGG* **251**, 373-380, (1996).
- 319 Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of
bacteriophage T4. (1970).
- 320 Wittig, I. & Schagger, H. Advantages and limitations of clear-native PAGE. *Proteomics* **5**,
4338-4346, (2005).
- 321 Bradford, M. M. A rapid and sensitive method for the quantitation of microgram
quantities of protein utilizing the principle of protein-dye binding. (1976).
- 322 Mackinder, L. C. *et al.* A repeat protein links Rubisco to form the eukaryotic carbon-
concentrating organelle. *Proceedings of the National Academy of Sciences of the United
States of America*, (2016).
- 323 Tcherkez, G. G. B., Farquhar, G. D. & Andrews, T. J. Despite slow catalysis and confused
substrate specificity, all ribulose biphosphate carboxylases may be nearly perfectly
optimized. *Proceedings of the National Academy of Sciences* **103**, 7246-7251, (2006).
- 324 Williams, T. A. & Fares, M. A. The Effect of Chaperonin Buffering on Protein Evolution.
Genome biology and evolution **2**, 609-619, (2010).
- 325 Durão, P. *et al.* Opposing effects of folding and assembly chaperones on evolvability of
Rubisco. *Nat Chem Biol* **11**, 148-155, (2015).
- 326 Molnar, A. *et al.* Highly specific gene silencing by artificial microRNAs in the unicellular
alga *Chlamydomonas reinhardtii*. *The Plant journal : for cell and molecular biology* **58**,
165-174, (2009).
- 327 Blaby, I. K. *et al.* Systems-Level Analysis of Nitrogen Starvation–Induced Modifications of
Carbon Metabolism in a *Chlamydomonas reinhardtii* Starchless Mutant. *The Plant Cell*
25, 4305-4323, (2013).
- 328 Jiang, X. & Stern, D. Mating and tetrad separation of *Chlamydomonas reinhardtii* for
genetic analysis. LID - 10.3791/1274 [doi] LID - 1274 [pii]. (2009).
- 329 Boynton, J. E. *et al.* Chloroplast transformation in *Chlamydomonas* with high velocity
microprojectiles. (1988).
- 330 Drapier, D. *et al.* The chloroplast *atpA* gene cluster in *Chlamydomonas reinhardtii*.
Functional analysis of a polycistronic transcription unit. (1998).
- 331 Drapier, D., Girard-Bascou J Fau - Wollman, F. A. & Wollman, F. A. Evidence for Nuclear
Control of the Expression of the *atpA* and *atpB* Chloroplast Genes in *Chlamydomonas*.
(1992).
- 332 Towbin H Fau - Staehelin, T., Staehelin T Fau - Gordon, J. & Gordon, J. Electrophoretic
transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and
some applications. (1979).

Summary

The necessity to coordinate the expression of genes originating from different genomes within the plant cell resulted in the appearance of mechanisms imposing nuclear control over organelle gene expression. Anterograde signaling through sequence-specific trans-acting proteins (OTAFs) coexists in the chloroplast with an assembly dependent control of chloroplast synthesis (CES process) that coordinates the stoichiometric formation of photosynthetic complexes.

Ribulose biphosphate carboxylase/oxygenase (Rubisco) is a chloroplast-located carbon fixing enzyme constituted of two subunits. Large subunit (LSU) and small subunit (SSU) are encoded in the chloroplast and nuclear genomes respectively. In the stroma they assemble to form a hexadecameric holoenzyme (LSU₈SSU₈). In this study I tried to highlight major regulatory points of its synthesis in *Chlamydomonas reinhardtii* focusing on the posttranscriptional regulation of LSU.

I showed that the MRL1 PPR protein is a limiting factor for *rbcL* mRNA accumulation. Whereas it has been previously designated as a stabilization factor for the abovementioned transcript, MRL1 appeared also to have a function in *rbcL* translation.

Most notably, I have demonstrated that in *Chlamydomonas reinhardtii* Rubisco expression is controlled by the small subunit (SSU) presence. In its absence *rbcL* undergoes an inhibition of translation through its own product – the unassembled Rubisco large subunit. This process depends on LSU-oligomerization state as I was able to show that the presence of a high order LSU assembly intermediate bound to the RAF1 assembly chaperone is essential for the regulation to occur. In parallel I shed light on the fate of unassembled LSU in a deregulated CES context, thereby improving our understanding of the process of its folding and assembly.

Résumé

La nécessité de coordonner l'expression des gènes provenant de génomes différents chez les plantes a conduit à l'émergence de mécanismes imposant un contrôle nucléaire sur l'expression génétique de l'organelle. Des signaux antérogrades, exercés par des protéines reconnaissant des séquences spécifiques, existent en parallèle avec un contrôle des synthèses chloroplastiques dépendant de l'assemblage (CES). Ensemble, ils coordonnent la formation stoichiométrique des complexes photosynthétiques.

La Ribulose biphosphate carboxylase/oxygénase (Rubisco) est une enzyme localisée dans le chloroplaste qui contient deux sous-unités. La grande sous-unité (LSU) et la petite sous-unité (SSU) sont codées par les génomes chloroplastique et nucléaire respectivement. Elles s'assemblent dans le stroma du chloroplaste pour former une holoenzyme hexadécamérique (LSU₈SSU₈). Pendant mon travail au laboratoire, j'ai tenté de décrire les étapes régulatrices majeures de la synthèse de la Rubisco chez *Chlamydomonas reinhardtii* en me focalisant sur la régulation post-transcriptionnelle de la LSU.

J'ai montré que la protéine PPR – MRL1 est un facteur limitant pour l'accumulation de l'ARN messager de *rbcL*. Bien qu'il ait été décrit précédemment comme un facteur stabilisateur du transcrit susnommé, MRL1 s'est révélé avoir un rôle dans la traduction.

J'ai par ailleurs démontré que chez *Chlamydomonas*, l'expression de la Rubisco est contrôlée par la présence de la SSU. En son absence, la traduction de *rbcL* est inhibée par son propre produit – la grande sous-unité non assemblée. J'ai pu montrer qu'un intermédiaire d'assemblage, constitué de LSU en complexe avec sa chaperonne RAF1, est nécessaire pour cette régulation, ce qui prouve que ce processus dépend de l'état d'oligomérisation de la LSU. Parallèlement, j'ai caractérisé le devenir de la LSU non assemblée quand la régulation CES est perturbée, et grâce à cela ait contribué à améliorer la connaissance de son processus de repliement et d'assemblage.