UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE

Laboratoire de Glycobiologie Structurale et Fonctionnelle UMR CNRS / USTL n°8576

DOCTORATE THESIS

Microbial Molecular Genetics

Jenifer NIRMAL RAJ

For obtaining the doctoral degree of the Ecole doctorale: Biologie Santé
Université de Lille I - Sciences et Technologies

Selection of Mutants Defective for Starch Biosynthesis in Marine Single-cell Cyanobacteria

16th April 2010

Jury:

President	:	Dr. Jean-Claude Michalski	Université de Lille 1, France
Examiners	:	Pr. Alan Myers	Iowa State University, USA
		Dr. Jean-Luc Putaux	Cermav-CNRS, Grenoble, France
Director	:	Pr. Steven Ball	Université de Lille 1, France

Abstract

Starch defines a semi-crystalline insoluble polysaccharide that aggregates into granules either in the cytosol or the chloroplast of respectively the glaucophytes, red algae and the true plants. Because of the selective distribution of starch within lineages that descend from primary endosymbiosis of the plastid, we investigated the origin of this particular organization. We thus contributed to the finding of comparable structures in a particular subgroup of cyanobacteria, and made a structural characterization of this cyanophycean starch from marine cyanobacteria related to *Crocosphaera watsonii* and *Cyanobacterium*.

We were able to show that, in these cyanobacteria, the solid starch granules could be degraded *in vitro* by a combination of starch bound glucan phosphorylase and endo-amylase; thereby defining a pathway entirely different from that characterizing the green plants and algae. In addition, we demonstrated that starch metabolism, in our particular strain, also responded to circadian clock control. Finally, we were able to generate and select mutants defective for starch biosynthesis. Among these we report the first mutants defective for cyanobacterial soluble starch synthase activity.

Resumé

L'amidon s'accumule en grains insolubles et semi-cristallins dans d'une part le cytoplasme des algues rouges et des glaucophytes et dans d'autre part les chloroplastes des algues et plantes vertes. Vu la distribution séléctive de l'amidon dans ces trois lignées issues de l'endosymbiose primaire du plaste, nous avons recherché lequel des deux partenaires de cette dernière est responsable de l'origine de cette structure. Nous avons, dans notre travail, contribué à l'identification de polysaccharides de ce type chez un sous groupe particulier de cyanobactéries et avons réalisé une caractérisation détaillée de l'amidon d'une souche de cyanobactérie marine apparentée à *Crocosphaera watsonii* et *Cyanobacterium*. Nous avons aussi pu montrer que chez ce sous groupe les grains pouvaient être dégradés par une combinaison d'acitivités phosphorolytiques et hydrolytiques qui leur est associée; définissant de cette manière une voie entièrement différent de celle qui caractérise les plantes te algues vertes. De plus nous avons démontré que le métabolisme de l'amidon répond dans notre souche à un contrôle par l'horloge circadienne. Enfin nous avons été capables d'engendrer et sélectionner des mutants défectueux pour la biosynthèse de l'amidon. Parmi ceux-ci nous rapportons les premiers mutants défectueux pour l'expression d'une amidon synthase chez les cyanobactéries.

Acknowledgments

I owe myself to the Almighty who is guiding Light, strength and wisdom in me.

This thesis arose by the direction of Prof. Steven Ball, years of research that has been done. By that, I have worked with a great number of people whose contribution in assorted ways to the research and the making of the thesis deserved special mention. It is pleasure to convey my gratitude to them all in my humble acknowledgment.

In first place, my grateful thanks address to Prof. Steven Ball who welcomed and accepted me in their research Group as well as giving me a supervision and guidance throughout the work from the very early stage of this research. I would like to record my gratitude to Dr. Christophe Colleoni for his supervision, advice and guidance. Who encourage me all the time and crucial contribution in this project. I sincere for his tiring and tedious help during my whole thesis.

I express my deep sense of gratitude to Dr. Jean-claude Michalski, Director of UGSF, who accepted to be a member of president jury of my thesis.

I want to thank prof. Alan Myers and Dr. Jean-Luc Putaux for having accepted to revise and judge this PhD work and accepted to be members of the jury.

A grateful acknowledgement goes to my colleagues David, Catherine, Charlotte, Phillipe, Nico, Aline, Ugo, Hande and Special thanks to Dr.Christophe D'hulst and members from his lab, who significantly contributed to the experimental realization of part of this work and also I want to thank all the people I worked with UGSF.

I convey special acknowledgement to remember and thank all of my friends particularly Suresh, Sara, Oliver and their family whose they really gave me a lot of sweet memories and support extended at the right moment. My Humble thanks to Archana chaudhary, Reni and Deepak.

I am grateful to my parents Nirmal raj, Premkumari and sister Joshi clement and my nephew Nithin for their moral support without which I could not have completed this project. Especially for Sr. Herbert Mary who believe in me all the time.

I would like to express my appreciation to my sweet heart and fiancée Stella, who supported me in all Occasions, whose dedication, love and persistent confidence in me, has taken the load off my Shoulders.

Finally, I would like to thank everybody who was important to the successful realization of thesis, as well as expressing my apology that I could not mention personally one by one.

Table of Contents

Figures

Introduction

Materials and Methods

Results

Tables

Introduction

Materials and Methods

Results

Abbreviations

3PGA : 3 – Phosphoglycerate

ADP : Adenosine diphosphate

ADP-Glc or ADPG : Adenosine diphosphate Glucose

ATP : Adenosine triphosphate

BE : Branching Enzyme

cDNA : Complementary De-oxy Ribonucleic acid

DNA : De-oxy Ribonucleic acid

DBE : Debranching Enzyme

DMSO : Dimethyl sulfoxide

DP : Degree of Polymerization

Dpe : Disproportionating enzyme

DTT : Dithiothreitol

EDTA : Ethylenediaminetetraacetic acid

EGT : Endosymbiotic gene transfer

GBSS : Granule bound Starch Synthase

GDP : Guanosine diphosphate

GMP : Guanosine monophosphate

Glc-1-P or G-1-P : Glucose-1-Phosphate

Glc-6-P or G-6-P : Glucose-6-Phosphate

GT : Glycosyl transferase

GWD : Glucan Water Dikinase

HEPES : (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

ISo : Isoamylase

Kb : Kilo basepair

kDa : KiloDalton

Km : Michaelis Menton Constant

LGT : Lateral gene transfer
mM : Millimolar
MEX : Maltose export
MOS : Malto-oligosaccharides
MS : Mass spectrometry
NDP : Nucleotide diphosphate
NST : Nucleotide sugar translocator
OD : Optical Density
PEP : Phosphoenolpyruvate
PHO : Phosphorylase
Pi : Orthophosphate, Inorganic Phosphate
PPi : Inorganic pyrophosphate
pPTs : Plastidial phosphate translocators
PUL : Pullulanase
PWD : Phosphoglucan water dikinase
RNA : Ribonucleic acid
mRNA : Messenger Ribonucleic acid
Rubisco : Ribulose 1,5-biphosphate carboxylase oxydase
SDS : Sodium dodecyl sulfate or Sodium lauryl sulfate
SS : Starch Synthase
SSS : Soluble Starch Synthase
TPT : Triose phosphate translocator
TRIS : Tris-(hrdroxymethyl)-aminomethane
UDP-Glu or UDPG: Uridine diphosphate glucose
WSP : Water Soluble Polysaccharide
WT : Wild Type
λmax : Maximum Absorbance
µg : Microgram

Introduction

Introduction

Sometime between 0.7 to 1.5 billion years ago (Cavalier-Smith, 2006; Yoon *et al.*, 2004) an ancestor of present-day cyanobacteria was internalized probably through phagocytosis by a heterotrophic eukaryotic cell. That this was a unique event is suggested by the fact that both protein sequences derived from the cyanobiont (the cyanobacterial endosymbiont) and those from the eukaryotic host are monophyletic and thus can be traced back to a pair of unique ancestors (McFadden *et al*., 2004; Rodriguez-Ezpeleta *et al.*, 2005). Although nothing is known about the nature of the ancient endosymbiotic link, it is reasonable to assume that the latter was based on the export of photosynthate from the cyanobiont to the host cytosol. Endosymbiosis of the plastid thus brought the ability to perform oxygenic photosynthesis into the eukaryotic world. As the cyanobiont slowly evolved to become a true organelle, the majority of cyanobacterial genes were lost as they were neither involved in neither oxygenic photosynthesis nor essential for maintenance and division of the symbiont. During this process a complex machinery of protein targeting from the cytosol to the evolving plastid appeared, thereby facilitating a process by which the remaining genes were transferred to the nucleus and their protein products synthesized on cytosolic ribosomes to be retargeted to the organelle. In addition, a number of other protein products and pathways were rewired to the evolving organelle which were not all necessarily present in the ancestral cyanobiont. Three eukaryotic lineages emerged after or during this metabolic integration of the plastid (Figure 1): the Chloroplastida (green algae and land-plants), the Rhodophyceae (red algae) and the Glaucophyta (glaucophytes). These three lineages generated through primary endosymbiosis contain the original "old" plastids with two membranes and were therefore recently named "Archaeplastida" (Adl et al., 2005).

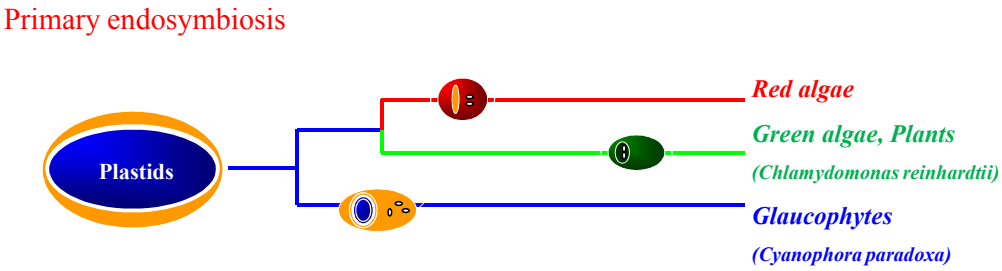

Figure 1: Emergence of Photosynthetic lineages. Archaeplastida are composed of 3 eukaryotic lines: The glaucophytes, the rhodophyceae and the chloroplastida. All 3 lines originate from a single endosymbiotic event involving a cyanobiont related to present-day cyanobacteria and a unicellular heterotrophic eukaryotic host. Simultaneously, all 3 lines gained the ability to synthesize starch, a novel form of storage polysaccharide related to glycogen, the otherwise most widespread form of storage polysaccharide found in living cells. Rhodophyceae and glaucophytes accumulate starch in their cytoplasm, whereas the latter is exclusively found in the plastids of chloroplastida (green algae and land plants).

Some single cell members or ancestors from these lineages were internalized probably also through phagocytosis by other heterotrophic eukaryotes thereby generating a variety of secondary endosymbiosis lines with derived plastids (Keeling, 2009). These secondary plastids are always surrounded by more than 2 and most of the time by 4 membranes. This generated a number of other important photosynthetic eukaryotes such as the brown algae, diatoms, dinoflagellates, cryptophytes and haptophytes. In addition to photosynthesis, eukaryotes have gained a number of other important biochemical features not found in heterotrophic eukaryotes unrelated to Archaeplastida. Among these, is the ability to store starch an insoluble and semi-crystalline form of storage polysaccharide which until quite recently was only reported in Archaeplastida and some but not all of their secondary endosymbiosis derivatives. Plant biologists are familiar with a form of starch found in the chloroplast or amyloplast of land plants and green algae. However

this polysaccharide is only found in the cytosol of red algae, glaucophytes, dinoflagellates and the non photosynthetic sister lineages of the latter: the apicomplexa parasites. In the cryptophytes, starch is found in the periplastidial space a compartment corresponding to the cytosol of the archaeplastidal alga that was internalized through secondary endosymbiosis to generate among others the cryptophyte lineage. Cytosolic starch was historically first studied in Florideophycidae, a complex group of multicellular red algae (for review see Viola et al., 2001). The term floridean starch was thus coined to describe this form of storage material (cytosolic or periplastidial starch will thus be referred to as "floridean" starch in this review). Therefore plastidial starch remains the exception rather than the rule among the diversity of starch storing lineages. This review is centered on the evolution of the starch pathway. Developments and refinements in the evolution of starch metabolism in grasses have been recently reviewed (Comparat-Moss and Denyer, 2009). In this Darwin review, we will focus on the means by which storage polysaccharide metabolism from the cyanobiont and its eukaryotic host merged to generate the starch pathway. We will propose that this merging of pathways was central to the success of primary endosymbiosis as it established the first biochemical link between the two unrelated partners.

1. Starch and glycogen define two different physical states of α-glucan storage polysaccharide metabolism

Living cells store carbohydrates in the form of a variety of polymers and oligomers. Among these glycogen defines by far the most widespread form of storage as it is found in Archaea, Bacteria and Eukaryotes. Glycogen is made of α-1,4 linked chains of glucose (α-1,4 glucans) that are branched together through α-1,6 linkages. The α-1,6 branches amount from 7 to 10% of the linkages and are symetrically distributed within the glycogen particle (for a review of glycogen structure see Shearer and Graham, 2002). Each chain with the exception of the outer unbranched

chains supports 2 branches. This branching pattern allows for spherical growth of the particle corresponding to tiers (a tier corresponds to a spherical sheath of chains generated between two consecutive branches located at similar distance from the center of the particle). This type of growth leads to an increase in the density of chains in each tier leading to a progressively more crowded structure towards the periphery (Figure 2B). Mathematical modelling predicts a maximal value for the particle size above which further growth is impossible as there would not be sufficient space for interaction of the chains with the catalytic sites of glycogen metabolism enzymes. This generates a particle consisting of 12 tiers corresponding to a 42 nm maximal diameter including 55000 glucose residues. 36% of this total number rests in the outer (unbranched) shell and is thus readily accessible to glycogen catabolism without debranching (Shearer and Graham 2002). In vivo, glycogen particles are thus present in the form of these limit size granules (macroglycogen) and also smaller granules representing intermediate states of glycogen biosynthesis and degradation (proglycogen) (Shearer and Graham 2002). Glycogen particles are entirely hydrosoluble and define thus a state where the glucose is rendered less active osmotically yet readily accesible to rapid mobilization through the enzymes of glycogen catabolism as if it were in the soluble phase.

Starch defines a solid semi-cristalline state composed of a mixture of two different polysaccharides with the same basic chemical linkages as glycogen (for a review of starch structure see Buléon *et al.*, 1998). Amylopectin, the major polysaccharide fraction is indispensable for starch granule formation and contains 4-6% branches while the minor fraction amylose contains less than 1% α-1,6 linkages. Amylose requires a preexisting amylopectin containing granule for its formation (Dauvillée *et al.*, 1999). Mutants deprived from this fraction can be readily isolated in green plants and algae (for review see Ball *et al.,* 1998). These mutants build wild-type amounts of normally organized granules. On the other hand some floridean starch accumulating lineages such as florideophycideae red algae (Viola *et al*., 2001) or apicomplexan

parasites (Coppin *et al*., 2005) lack amylose while sister lineages of the latter (such as the Porphyridales red algae (Shimogana et al., 2007; 2008) or the dinoflagellates (Deschamps et al., 2008d) typically include this polysaccharide fraction. Amylose however is always found in the granules synthesized within plastids by wild-type green algae and land plants (Ball *et al*., 1998). Amylopectin defines one of if not the largest biological polymer known and contains from 10^5 to 10^6 glucose residues (Buléon *et al*., 1998). There is no theoretical upper limit to the size reached by individual amylopectin molecules. This is not due to the slightly lesser degree of overall branching of the molecule when compared to glycogen. Rather it is due to the way the branches distribute within the structure. As displayed in Figure 2A, the branches are concentrated in sections of the amylopectin molecule leading to clusters of chains that allow for indifinite growth of the polysaccharide. Another major feature of the amylopectin cluster structure consists in the dense packing of chains generated at the root of the clusters where the density of branches locally reaches or exceeds that of glycogen. This dense packing of branches generates tightly packed glucan chains that are close enough to align and form parallel double helical structures. The helices within a single and neighbouring clusters align and form sections of crystalline structures separated by sections of amorphous material (containing the branches) thereby generating the semi-cristalline nature of amylopectin and of the ensuing starch granule (Buléon *et al*., 1998). Indeed the crystallized chains become insoluble and typically collapse into a macrogranular solid. This osmotically inert starch granule affords for the storage of unlimited amounts of glucose that become metabolically unavailable. Indeed the enzymes of starch synthesis and mobilization are unable to interact directly with the solid structure with the noticeable exception of granule-bound starch synthase the sole enzyme required for amylose synthesis. This enzyme is indeed able to extend amylose chains by synthesizing α-1,4 glucosyl linkages processively within the polysaccharide matrix (reviewed in Ball *et al.*, 1998). Because no other enzyme is significantely active within granules, this will lead to the formation of long unbranched polysaccharides. On the

other hand in Archaeplastida, glucan-water dikinase initiates amylopectin degradation by phosphorylating selective glucose residues within the clusters thereby disrupting the crystal and facilitating access and attack by hydrosoluble enzymes of starch catabolism (reviewed in Fettke *et al*., 2009). The solid state of starch thereby generates glucose stores which are not as readily accessible as those of glycogen.

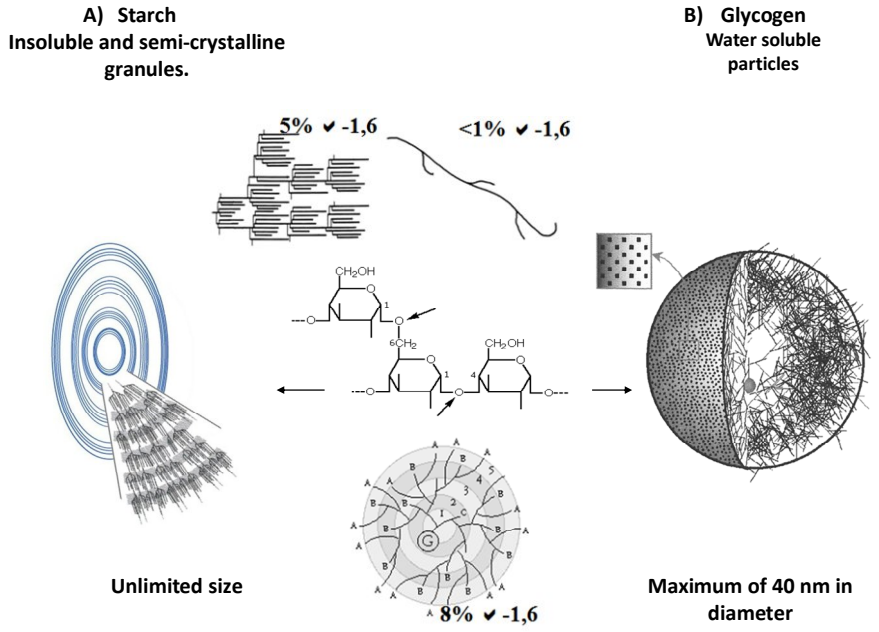

Figure 2: Structure of Starch and Glycogen. Both glycogen and starch are made of units of glucose linked in alpha-1, 4 and branched in alpha1-6 but they differ in their structures and ensuing physicochemical properties. Glycogen consists of tiny homogeneous hydrosoluble particles whose diameter is limited to a maximum of 40 nm. Starch can be defined as a heterogeneous mixture of 2 distinct fractions: the high mass moderately branched amylopectin and the lower mass and less branched amylose. This fraction is produced by the granule-bound starch synthase (GBSS), the sole enzyme required for amylose synthesis. Amylpectin and amylose aggregate in semicrystalline granules of unlimited size. Unlike glycogen, where the alpha-1,6 linkages are uniformly distributed, the branches of amylopectin are concentrated in certain regions of the molecule. The chains resulting from these branches are organized in clusters.

On the other hand starch can be seen as a very efficient intracellular sink immobilising vast amounts of carbon out of cellular metabolism. Mobilizing starch is thus anything but trivial. Indeed because starch defines the most important source of calories in the human diet, human

populations have duplicated genes encoding salivary α-amylase as a function of their local diet (Shadan 2007). Only a small fraction of damaged uncooked starch granules are mobilized during digestion. Because starch granules swell and melt at high temperatures thereby loosening the crystal structure, cooking meals has vastly improved the amount of calories that human can extract from such polysaccharides in their diet.

As previously mentioned the distribution of starch polysaccharides seemed until recently to be limited to Archaeplastida and some of their secondary endosymbiosis derivatives. Therefore the large amounts of carbohydrates and energy available through photosynthesis do not, per se, explain the appearance of this form of storage material. Indeed most photosynthetic bacteria including cyanobacteria were reported to accumulate glycogen and not semi-crystalline starch.

2. Comparative biochemistry of glycogen metabolism in bacteria and opistokonts

As we will see, the enzymes of glycogen and starch metabolism are clearly related. In addition, in Archaeplastida, the pathways of starch biosynthesis and degradation define a mosaic of enzymes phylogenetically related either to bacterial (mostly cyanobacterial) or eukaryotic glycogen metabolism (Coppin et al., 2005; Patron and Keeling, 2006; Deschamps et al., 2008a). The obvious explanation for this observation would be that both partners of plastid endosymbiosis had the ability to synthesize related storage polysaccharides before endosymbiosis. These certainly consisted of α-1,4-linked glucans branched through α-1,6 linkages. Glycogen metabolism defines well studied and conserved pathways within gram-negative bacteria and opistokonts (fungi and animals) who define those eukaryotes that have by far been the most intensively studied.

To understand the merging of these pathways that occurred after endosymbiosis we will briefly outline their common and distinctive features. Figure. 3 summarizes the basic common pathway of storage polysaccharide synthesis in gram negative bacteria (and cyanobacteria) (for review see Preiss 1984) and opistokonts (for review see Roach, 2002).

Briefly, glucose is polymerized within these polysaccharides thanks to its activation in the form of a nucleotide-sugar through the action of NDP-glucose pyrophosphorylase. All eukaryotes known (with the exception of Archaeplastida synthesize glycogen from UDP-glucose while all gram-negative glycogen accumulating bacteria use ADP-glucose. ADP-glucose is a bacterial-specific metabolite not found in heterotrophic eukaryotes. Unlike UDP-glucose which is used by all living cells to synthesize a large number of different molecules, ADP-glucose is devoted to the synthesis of glycogen in bacteria (and also to the osmoprotectant glucosylglycerol in cyanobacteria) (Preiss 1984; Miao et al. 2003; 2006).

Thus, the synthesis of ADP-glucose defines the first committed step of glycogen synthesis in bacteria while glucan elongation defines the first committed step of eukaryotic glycogen synthesis. The glucose from the glycosyl-nucleotide is then transferred to the non-reducing end of a growing α-1,4 linked chain through an elongation reaction catalyzed by glycogen synthase. Branching proceeds differently through an hydrolytic cleavage of a pre-existing α-1,4-linked glucan synthesized through glycogen synthase and an intra or intermolecular transfer of a segment of chain in α-1,6 position.

The branched polymers are sujected to degradation through a combination of glycogen phosphorylase and debranching enzyme. Glycogen phosphorylase defines an enzyme which releases glucose-1-P from the non reducing-end of glycogen in the presence of orthophosphate. This enzyme is unable to cleave the α-1,6 branch and is known to stop 4 glucose residues away from the branch (Dauvillée et al., 2005; Alonso-Casajús et al., 2006). Therefore the short 4 glucose residues long external chains need to be further digested through the action of

debranching enzymes. Debranching enzymes in eukaryotes and bacteria operate differently. In eukaryotes indirect debranching enzyme defines a bifunctional enzyme containing both an α-1,4 glucanotransferase and an α-1,6 glucosidase catalytic site. The transferase will first hydrolyse the last α-1,4 linkage before the branch and thus transfer three glucose residues (maltotriose) to an outer neighbouring chain within the glycogen particle. Glycogen phosphorylase will further recess this 7 glucose residue long chain back to 4 while the second catalytic site will hydrolyze the α-1,6 linkage from the residual unmasked glucose at the branch (for review see Roach, 2002). The net result will consist of complete degradation of glycogen to glucose-1-P and glucose. Bacteria operate through a simpler debranching enzyme that directly cleaves the α-1,6 branch thereby producing a 4 glucose residue long malto-oligosaccharide (maltotetraose) (Dauvillée et al., 2005).

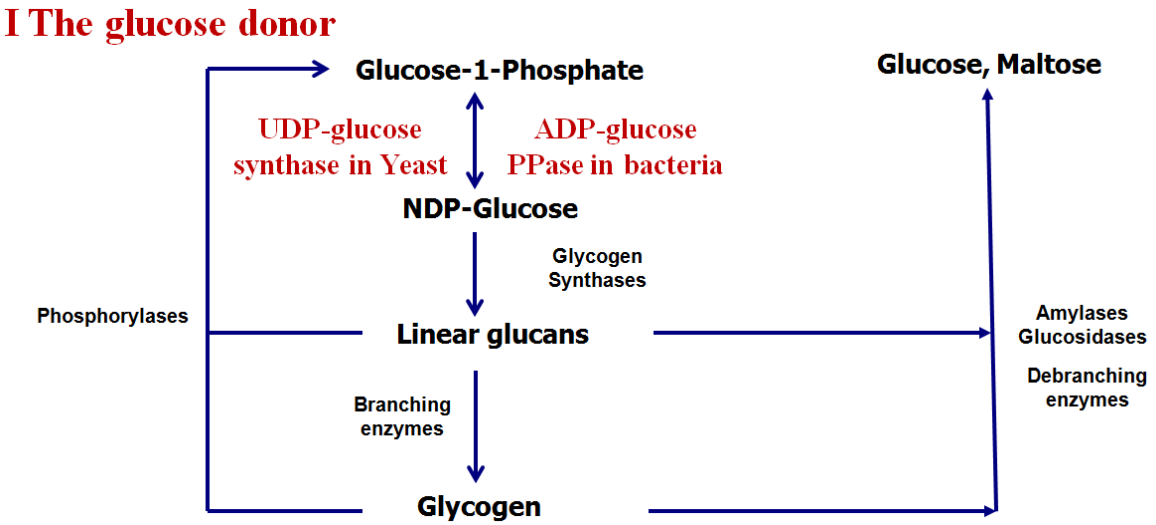

Figure 3: A common basic starch metabolism pathway comparision between Eukaryotes and Bacteria.

These malto-oligosaccharides are then degraded through a combination of α-1,4 glucanotransferase and a maltodextrin phosphorylase distinct from the glycogen phosphorylase (reviewed in Boos and Schuman, 1998). Here again the transferase elongates an acceptor

maltotetraose with a donor malrotriose enabling maltodextrin phosphorylase to further recess the chains. Thus direct debranching in bacteria implies the coupling of glycogen and malto-oligosaccharide (MOS) metabolism (Boos and Schuman, 1998) while MOS metabolism is not needed and indeed not found in opistokont genomes.

In addition to this phosphorolytic pathway of glycogen mobilization there is good evidence for the presence of hydrolytic pathway in opistokonts and circumstantial evidence for the the presence of such a pathway in gram-negative bacteria. Fungi and animals indeed contain an enzyme able to hydrolyze both the α-1,4 and α-1,6 linkages responsible for the degradation of a significant pool of cellular glycogen (reviewed by Roach, 2002; François and Parrou, 2001). However this enzyme is contained in the lysosome (or yeast vacuole) leading to a clear partitioning between the locations of both glycogen synthesis or phosphorolysis which occurs in the cytosol and glycogen hydrolysis which is confined to the lysosome (or yeast vacuole).

In yeast autophagy clearly further impacts regulation of glycogen metabolism (Wang et al., 2001). Undisputable functional proof of the importance of glycogen hydrolysis pathway has been obtained both in yeasts where it is triggered both during sporulation or late stationary phase and also in humans where its absence is known to lead to Pompe's disease (glycogen storage disease type II) (reviewed by Roach, 2002; François and Parrou, 2001). In bacteria and cyanobacteria α-amylase-like sequences are often found in the genomes suggestive of the presence of such a pathway but mutant evidence is lacking (Wing-Ming *et al*., 1994; Reyes-Sosa *et al*., 2010).

It is striking to note that mutations abolishing analogous enzyme activities in model organismes such as *E. coli* and yeast lead to similar or identical phenotypes establishing that all enzymes play analogous functions in the storage polysaccharide metabolism network. This even remains true for very different enzymes such as indirect or direct DBEs from bacteria and opistokonts (Teste *et al*., 2000; Dauvillée et al., 2005). Nevertheless, the use of distinct nucleotide sugars for glycogen polymerization will impact very differently the regulation of the prokaryotic and

eukaryotic pathways. The synthesis of ADP-glucose by ADP-glucose pyrophosphorylase being the first committed step of bacterial glycogen synthesis, this enzyme will be subjected to tight allosteric regulations with effectors that vary according to metabolic specilaization of the bacterial species. Cyanobacterial ADP-glucose pyrophosphorylase in particular is known to be activated by 3-PGA and inhibited by orthophosphate. This regulation in addition to the presence of ATP and glucose-1-P as substrates further couples ADP-glucose synthesis to carbon fixation through the Calvin cycle and thus to photosynthesis, a regulation which was conserved in the case of plastidial starch synthesis in Chloroplastida (for review see .Ballicora *et al.* 2003) In opistokonts protein phosphorylation and dephosphorylation through protein kinases and phosphatases has been known for years to activate or inhibit glycogen synthase and glycogen phosphorylase by modifying their sensitivity to allosteric effectors.

Historically, protein kinases and phosphatases were discovered by studying the physiology of glycogen metabolism in animals (Krebs, 1983). The glycogen synthase of opistokonts is a complex enzyme belonging to a distinct class of glycosyltransferase (GT3 according to the Cazy classification) than that of the bacterial enzyme (GT5). The GT3 opistokont wild-type enzyme is unable to prime the reaction and requires a separate malto-oligosaccharide primer. The "natural" primer for the fungal or animal enzyme is a small protein capable of autoglucosylation: glycogenin. Functional evidence for the importance of glycogenin in glycogen metabolism has been produced in yeast and animals (Roach, 2002).

However in bacteria biochemical evidence suggests that the GT5 ADP-glucose requiring starch synthase is capable of autoglucosylation and therefore does not need the presence of another protein to prime glycogen synthesis (Ugalde et al., 2003). In total, the glycogen pathways of bacteria and opistokonts consists of a network of 6 to 12 enzymes of related function.

3. Comparative biochemistry of glycogen metabolism in opistokonts, amoebozoa and other heterotrophic eukaryotes.

Over 99% of the studies performed on glycogen metabolism in eukaryotes concerns fungi and animals (for reviews see François and Parrou 2001; Shearer and Graham, 2002; Roach 2002). Fungi, animals and lesser known related lineages such as the choanoflagellates define a monophyletic lineage named the opistokonts. This, while of great importance to humans, defines only a small subset of the diversity that typifies eukaryotes (see Figure 4). Because the eukaryotic ancestor that hosted the cyanobiont is not presently thought to define an opistokont or an opistokont ancestor, it becomes important to investigate the nature of storage metabolism to ascertain that the model generated by available studies also applies to other lineages. Among the non-opistokont glycogen accumulating lineages a number of genomes have recently appeared that are relevant to this question. Amoebozoa define an important and diverse group of organisms thought to be located closer to the proposed root of the eukaryotic tree of life (Richards and Cavalier-Smith, 2005). *Dictyostellium discoideum* defines an interesting model familiar to cell biologists and geneticists. The genome of this organism has been sequenced (Eichinger et al. 2005). Among the surprising features displayed by this genome is the presence of a greater number of distinct protein domains than that found either in fungi and animals (Eichinger et al. 2005). A logical explanation for this increase would be the conservation of the initially greater diversity of genes that typified the ancestors of eukaryotes. Glycogen metabolism also displays this increase in complexity. Indeed not only does *Dictyostellium* harbor the full suite of genes found in fungi and animals for glycogen metabolism but in addition it includes a second type of glycogen synthase belonging to the GT5 CAZy family (Deschamps et al., 2008a; Cantarel *et al*., 2009).

Interestingly, it also contains an α-1,4 glucan transferase named dpe2 and related amoeba such as the pathogen *Entamoeba histolytica* contain both dpe2 and β-amylase (Loftus *et al.*, 2005; Deschamps et al., 2008a) as do all Archaeplastida where these enzymes were first described (see the next section for a description of the function of these enzymes). Dpe2 is found together with β-amylase in other eukaryotic lineages unrelated to amoebozoa including the parabasalid *Trichomonas vaginalis* (Carlton *et al.* 2007; Deschamps et al., 2008a). As to the GT5 glycogen synthase, this enzyme is also found in place of the GT3 enzyme in ciliates, other amoebas, parabasalids and also in Glaucophyta, Rhodophyceae (red algae) and lineages thought to derive from them by secondary endosymbiosis such as the apicomplexa parasites (Aury *et al*., 2006; Eisen *et al*., 2006; Carlton *et al.* 2007; Coppin *et al*., 2005; Deschamps et al., 2008a). Because of the very wide distribution of these additional enzymes of glycogen metabolism among eukaryotic lineages separated by over a billion years of evolution (Song et al., 2005), we strongly argue that their presence cannot be explained by lateral gene transfer from Archaeplastida lineages where these activities were first reported and studied.

The most logical explanation would consist of the existence of richer suite of genes of glycogen metabolism in the eukaryotic ancestor that was followed by different histories of selective gene losses in distinct eukaryotic lineages. For instance opistokonts would have lost β-amylase, dpe2 and the GT5 glycogen synthase while parabasalids and archamoebas would have lost only the GT3 enzyme. Ciliates would have lost β-amylase, dpe2 and the GT3 enzyme. Amoebas in general and mycetozoa in particular such as *Dictyostellium discoideum* would have experienced less gene losses than other eukaryotes. The detailed function of the GT5 UDP-glucose utilizing enzyme (although the suspected substrate specificity remains to be formally proven) in the glycogen metabolism network as well as the cytosolic or lysosomal location of the putative β-amylase-dependent pathway of glycogen hydrolysis remains to be ascertained.

An interesting question concerning the sole presence of a GT5 UDP-glucose utilizing glycogen synthase in many eukaryotic lineages consists of its dependence on glycogenin for priming and its possible regulation through the well known set of protein kinases and phosphatases that normally control the GT3 enzyme. Because of the maintenance of a richer suite of enzymes involved in glycogen metabolism in Amoebozoa we have chosen *Entamoeba histolytica* as our reference genome to exemplify the status of glycogen metabolism as it possibly existed in the eukaryotic partner of endosymbiosis before the latter engulfed the cyanobiont.

4. A brief overview of starch metabolism in Chloroplastida

Decades of research and a wealth of studies concerning starch metabolism in Chloroplastida have led to the identification of a very well conserved pathway from the earliest diverging prasinophyte single cell alga such as Ostreococcus to the most sophisticated multicellular terrestrial plants such as maize or rice (Ral *et al*. 2004; Derelle *et al.*, 2006; Deschamps *et al.*, 2008b) (Figure 4). Many reviews are accessible for the interested reader that concern our present detailed understanding of starch biosynthesis and degradation (Myers *et al.*, 2000; Ball and Morell, 2003; Tomlinson and Denyer**,** 2003; Tetlow *et al*., 2004; Morell and Myers, 2005; Zeeman *et al.,* 2007; Fettke *et al*., 2009). A general feature of the plastidial pathways of starch metabolism is defined by its astonishing apparent complexity.

Over 40 genes (not including regulatory genes) seem involved in the building and mobilization of starch in plastids while less than 12 genes are comparatively directly involved in glycogen metabolism both in eukaryotes and bacteria. This apparent increase in complexity is largely due to the high number of isoforms that catalyze each of the steps that we have outlined in the preceding sections. For instance a minimum of 5 starch synthases participate in polymer

elongation, three branching enzymes are reported to introduce the α-1, 6 linkage, 4 direct debranching enzymes are involved in different facets of starch metabolism etc. These enzymes play only partly redundant functions with one another and are often responsible for distinctive roles in the building or degradation of different substructures of starch. Because the starch granule defines a highly organized structure, it was believed by many that this was required to explain the underlying complexity of the granule architecture. The chloroplastidal pathway relies on the sole use of ADP-glucose (Zabawinski *et al*., 2001).

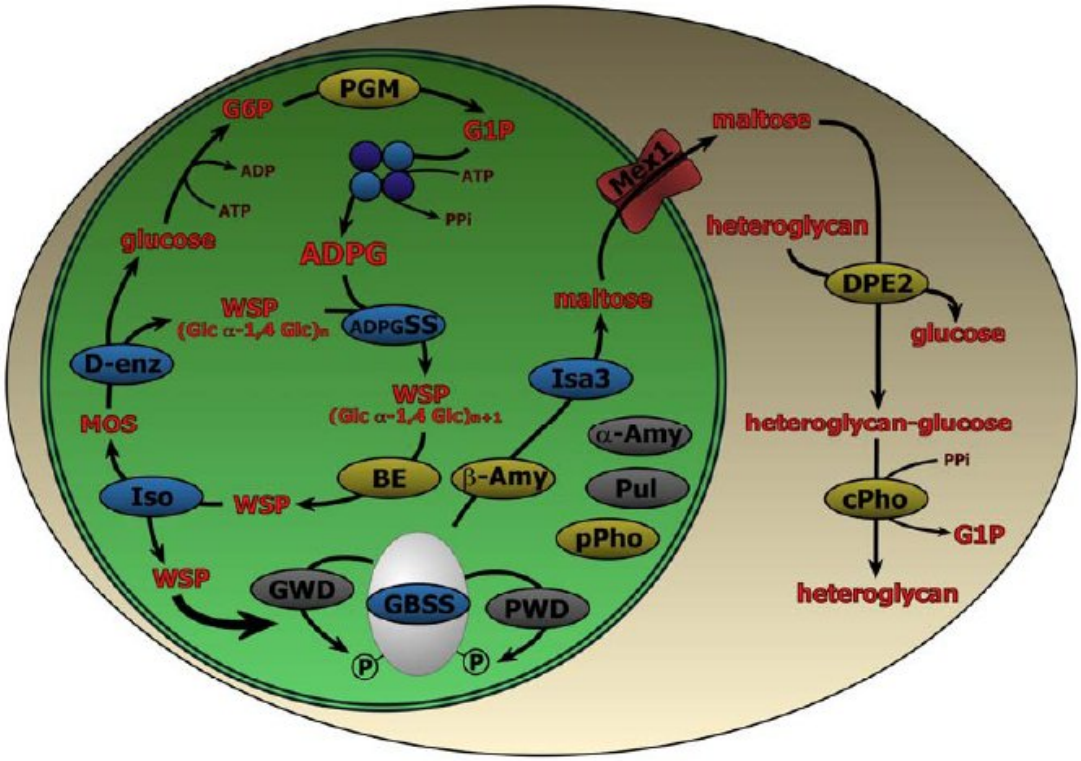

Figure 4: The pathway of starch synthesis and degradation in Chloroplastida. Glucose-6-P (G6P) is concerted into glucose-1-P (G1P) through the action of plastidial phosphoglucomutase (PGM). G1P reacts with ATP to generate ADP-glucose (ADPG) through the major rate-controlling heterotetrameric ADPG pyrophosphorylase (symbolized by four spheres). The glucose from ADPG is then transferred to the non-reducing ends of growing α-1,4 linked chains through the action of soluble starch synthases ($_{ADPG}$SS) to generate a water-soluble oligosaccharide (WSP(Glc α-1,4 Glc) $_{n+1}$). α-1,6 branches are introduced through the cleavage by branching enzymes (BEs) of an α-1,4 linkage from a growing chain and transfer of a segment to a neighboring chain through an α-1,6 glycosidic bond. The randomly branched polymer (WSP) is incapable of aggregating into semi-crystalline insoluble granules if those branches that prevent crystallization of amylopectin are not spliced out through the action of a debranching enzyme (Iso). The latter displays an isoamylase type of specificity. The spliced polysaccharide aggregates into the starch granule (in white) while the oligosaccharides are recycled through the combined action of disproportionating enzyme (D-enz)

and various other hydrolases and phosphorylases. D-enz transfers maltose or longer glucans to accept or oligo or polysaccharide. Some of these oligosaccharides might be transferred directly back to the pool of WSP by D-enz. Amylose (the second unbranched fraction of starch) is exclusively synthesized through the only enzyme working within the starch granule insoluble polysaccharide matrix: granule bound starch synthase (GBSS). Starch degradation is initiated through the phosphorylation of starch at the C6 and C3 positions of glucose by, respectively, glucan-water dikinase and phosphoglucan-water dikinase (GWD and PWD). The loosened structure is then attacked by β-amylases (β-Amy), which generates β-maltose, and a special type of debranching enzyme (Isa3). Maltose is exported to the cytosol through the maltose export transporter (MEX1). In the cytosol, glucose is both released and transferred from maltose by transglucosidase (DPE2). It is thought that a heteroglycan is used as acceptor of this transfer reaction and that the outer α-1, 4 linked chains of this heteroglycan are further recessed by cystosolic phosphorylase (cPho). Phosphorylases define enzymes that recess oligo- or polysaccharides, generating G1P from the non-reducing end in the presence of orthophosphate. Three types of plastidial enzymes are thought to have a role in other intraplastidial pathways of starch mobilization. However, functional evidence for the role of plastidial phosphorylase (pPho), α-amylase (α-Amy) and pullulanase (Pul) is presently lacking. Enzymes are colored with respect to their phylogenetic origin: blue for cyanobacterial, yellow for eukaryotic host and purple for unknown. For reviews of starch synthesis, degradation and structure in Chloroplastida, (see Ball, S.G and Morell, M.K., 2003 and Lu, Y. and Sharkey, T.D., 2006).

The enzymes of ADP-glucose production and those that elongate glucans with this substrate display a distinctive cyanobacterial phylogeny which apparently correlates with the plastidial location of starch in the green lineage (Coppin et al., 2005; Patron and Keeling, 2006, Deschamps et al., 2008a). Most importantly ADP-glucose pyrophosphorylase has conserved the major regulatory properties of the cyanobacterial enzyme and has thus remained throughout its history an enzyme which is tightly coupled to the Calvin cycle and photosynthesis (Ballicora et al., 2003). The starch pathway resembles that of cyanobacterial glycogen metabolism with two major differences. The first difference pertains to the means by which plants achieve the asymmetric distribution of branches within the amylopectin clusters that explains the solid semi-crystalline state of starch and most of its physical properties. Mutant work in Chlamydomonas, cereals and later in Arabidopsis (James *et al.* 1995; Mouille *et al*., 1996; Nakamura *et al.*, 1997; Zeeman *et al*., 1998; Wattebled *et al.*, 2005) have strongly suggested that a form of direct debranching enzyme named isoamylase debranches the loosely spaced α-1,6 linkages only within the hydrosoluble precursor of amylopectin thereby generating the tight spacing of branches required at the root of clusters for polymer crystallization (Ball *et al*., 1996). In the absence of this activity, mutants of Chlamydomonas revert entirely to the synthesis of glycogen (Mouille *et al*., 1996). A second major difference consists in the way the starch granule is degraded.

Chloroplastida enzymes of starch catabolism are unable to directly attack the solid granule. In order to mobilize starch the latter must first be phosphorylated through an enzyme named glucan water dikinase (GWD) that carries both a starch binding and a dikinase domain (Ritte *et al*., 2002, reviewed in Blennow et al., 2002; Fettke et al., 2009). The β-phosphate from ATP is thus bound to the C6 of a few glucose residues within the crystalline lamellae (Ritte *et al*., 2002). The pre-phosphorylated lamellae are then further phosphorylated through PWD (phosphoglucan water dikinase) which cannot initiate starch phosphorylation but requires the prior action of GWD. PWD introduces phosphates at the C3 position. This phosphorylation is sufficient to locally loosen the tight crystal packing of glucans within the granule and allow for degradation of amylopectin through the concerted action of β-amylases (Scheidig *et al*., 2002) and a specialized form of direct debranching enzyme named isa3 (Edner *et al*., 2007; reviewed in Fettke *et al.,* 2009). It is suspected but not proven that other enzymes of starch catabolism may be active at this stage (α-amylase and phosphorylases). The phosphate is then released through the action of sex4 a phosphatase which is functionally equivalent to laforin (see below for definition) but with a different organization of starch binding and phosphatase domains (Gentry et al., 2007; Kotting *et al.*, 2009). β-amylase is an exo-hydrolase producing maltose processively from the non-reducing end of an amylopectin cluster. The maltose cannot be degraded by plastidial enzymes but will be exported by a specialized transporter named mex (maltose export) to the cytosol (Niittylä *et al*., 2004). In the cytosol the maltose will be metabolized thanks to the action of dpe2 (Fettke *et al*., 2009). Dpe2 is an α-1,4 glucanotransferase that will cleave the α-1,4 linkage of maltose with concomitant transfer of one glucose residue to a required acceptor glucan. The acceptor glucan is believed to be a cytosolic heteroglycan whose outer chains consist of α-1,4 linked glucose residues. A cytosolic phosphorylase is thought to recess these outer chains and thereby release glucose-1-P in the presence of orthophosphate (Fettke *et al*., 2009). All components of the starch degradation machinery with the noticeable exception of isa3 are either of eukaryotic phylogeny

or of unknown phylogeny (such as α-amylase and pullulanase) (Deschamps et al. 2008a). Isa3 itself is not of clear cyanobacterial phylogeny and may be more related to Chlamydiae than cyanobacteria (see below). On the whole the degradation pathway is completely unrelated to polysaccharide degradation in bacteria. In addition we have no indication that storage polysaccharides are phosphorylated in cyanobacteria and no equivalent to GWD, PWD, β-amylase and dpe2 can be evidenced in extant bacteria and cyanobacteria. In short Chloroplastida display exceedingly complex pathways of starch synthesis and degradation that only very superficially resembles cyanobacterial storage polysaccharide metabolism; Phylogenetically the pathways define a mosaic of enzymes of distinctive host and cyanobacterial origin (Coppin *et al.*, 2005; Patron and Keeling, 2006; Deschamps *et al*., 2008a) (see table 1).

5. The very simple pathways of floridean starch synthesis and degradation

Much less is known about the pathway of starch synthesis and degradation in the two other Archaeplastida lineages: the Rhodophyceae and Glaucophyta. In both instances starch accumulates in the cytosol of these organisms. A growing body of biochemical and molecular evidence point to the existence of an UDP-glucose based pathway both in Rhodophyceae (for review see Viola *et al*., 2001) and Glaucophyta (Plancke *et al*., 2008). A UDP-glucose pathway is also suspected to be at work in those lineages that are thought to be derived from red algae through secondary endosymbiosis such as the dinoflagellates, apicomplexa parasites and cryptophytes (Coppin et *al.*, 2005; Deschamps *et al*., 2006; Deschamps *et al.*, 2008d). Rhodophyceae are very poor biochemical and genetic models and no starch accumulating red alga can fulfill the prerequisite to become an efficient system allowing for the functional dissection of starch metabolism. The only exception to this pessimistic view comes from the study of the secondary endosymbiont *Crypthecodinium cohnii* (Deschamps *et al*., 2008d;

Dauvillée *et al*., 2009). This homothallic heterotrophic dinoflagellate species does allow for the selection of mutants and crossing. Mutants of *Crypthecodinium* have very recently been reported that have decreased starch amounts and (or) a modified polysaccharide structure (Deschamps *et al*., 2008d; Dauvillée *et al*., 2009). Severely impaired mutants of *C cohnii* were demonstrated to have a decreased and strongly modified UDP-glucose requiring starch synthase (Dauvillée *et al*., 2009).

The defect in starch amount and alteration in amylopectin structure cosegregated in crosses with the modification in enzyme activity. Because no other assayable enzyme of starch metabolism was affected in these mutants, we believe this brings functional proof that floridean starch in this case is indeed synthesized through the UDP-glucose substrate. 4 Rhodophyceae genomes have been recently sequenced including two unicellular cyanidiales and two complex multicellular species (Matsuzaki *et al*., 2004). As with the Chloroplastida, the pathways are very well conserved throughout the lineage. The gene content is displayed in Table 1. The most striking feature of the red lineage pathway is the paucity of enzymes required to synthesize and mobilize starch. Less than 12 genes seem required to operate starch metabolism making it no more complex than glycogen metabolism. Yet Rhodophyceae do accumulate complex starch granules with all major features found in Chloroplastida starch. Some red alga lineages such as the Porphyridiales also accumulate amylose at variance with the initial report that floridean starch lacked this fraction (Nakamura *et al.*, 2005; Shimogana *et al*., 2007).

This very important result proves that a complex pathway is not required to explain the biogenesis of the starch granule architecture. Another striking feature of the pathway is that with the noticeable exception of the enzymes producing or using the nucleotide sugar substrate all other steps of starch synthesis and degradation are analogous in Rhodophyceae and Chloroplastida. Indeed phylogenetic trees show a common origin for all enzymes of starch metabolism in complete agreement with the monophyletic nature of Archaeplastida (Coppin *et*

al., 2005; Patron and Keeling, 2006; Deschamps *et al*., 2008a; Plancke et al., 2008). The only major difference is defined by the absence of ADP-glucose pyrophosphorylase and of the cyanobacterial type of GT5 ADP-glucose requiring starch synthase in Rhodophyceae and Glaucophyta. However GBSS the enzyme of cyanobacterial phylogeny responsible for amylose synthesis within granules is present in Glaucophyta, Porphyridiales red algae, in cryptophytes and in dinoflagellates (Plancke *et al*., 2008; Deschamps *et al*., 2006; 2008a; 2008d; Shimogana *et al*., 2007). In addition the floridean starch GBSS shows a marked preference for UDP-glucose while remaining capable to use ADP-glucose in Glaucophyta, cryptophytes and Porphyridiales.

The soluble starch synthase used by the Rhodophyceae for amylopectin synthesis seems to be unique (no other candidate genes are found in these genomes) and to consist of the GT5 type of glycogen synthase found in many eukaryotic lineages distinct from the opistokonts (Deschamps *et al*., 2008a). The same enzyme was found and its sequence cloned during the preliminary characterization of starch metabolism in the the glaucophyte *Cyanophora paradoxa* (Plancke *et al*., 2008). Remarkably this enzyme is thus able to fulfill all functions which in Chloroplastida seem to require 4 different soluble starch synthases. This enzyme was initially thought by Patron and Keeling (2006) to descend from the cyanobacterial GT5 enzymes. However at the time of their study these authors were unable to realize that this in fact it represented one of the two major forms of glycogen synthase found in heterotrophic eukaryotes. All major steps of starch synthesis and degradation are represented by a single enzyme in the rhodophycean pathway (Table 1).

The only interesting exception to this is defined by starch debranching enzyme (isoamylase) which is represented by two isoforms of bacterial phylogeny. Interestingly in the glaucophyte *Cyanophora paradoxa* isoamylase is known to be synthesized as a large size multimeric complex as in green plants and algae suggesting that this enzyme may have a similar function as that proposed for Chloroplastida. The absence in the rhodophycean genome of dpe1 (D-enzyme), an

enzyme required for assimilitation of maltooligosaccharides longer than maltose, may suggest that in this lineage the other α-1,4 glucanotransferase (dpe2) possibly supplies an equivalent function in addition to its function in maltose assimilation.

6. The evolutionary origin of starch-like structures

The appearance of starch in Archaeplastida begs the question of the origin of this structure. Was a starch-like polymer synthesized before endosymbiosis by either the host or the cyanobacterium or did starch result accidently from the merging of related yet dissimilar pathways? The existence of such polymers in the eukaryotic ancestors seems highly unlikely. Indeed this would suggest that among the diversity of extant heterotrophic eukaryotes one would expect several lineages unrelated to primary endosymbiosis to contain such polymers. However each time an heterotrophic eukaryote was reported to contain starch-like polymers it turned out to define lineages which have lost photosynthesis either among Archaeplastida (the white algae such Polytomella, Polytoma, Prototheca and Helicosoporidium (Hamana et al. 2004; De Koning and Keeling, 2006; Pombert and Keeling, 2010) or among secondary endosymbiosis lines. The most striking case is defined by several apicomplexa parasite species such as *Toxoplasma gondii* which had been known for years to accumulate amylopectin granules (Coppin et al., 2005). It was indeed subsequently found that apicomplexa harbored a cryptic plastid that resulted from secondary endosymbiosis of an Archaeplastida ancestor. As to cyanobacteria, all species examined were reported to contain glycogen and no convincing report or claim of the presence of starch had appeared until very recently. Because the enzyme responsible for generating the crystalline structure of starch displays a bacterial phylogeny and because GBSSI the only enzyme able to elongate glucans within the starch granule itself also displays a cyanobacterial origin, it

remained possible that the cyanobiont's ancestor synthesized such polymers. Nakamura *et al.* (2005) were the first to report the existence of starch like polymers organized into insoluble granules within one group of cyanobacteria which was named subgroup V according to the classification by Honda *et al.* (1999). Because in their survey they had not found *bona fide* large size granules containing amylose they used the term semi-amylopectin to name this type of polymer. Prior to this survey the studies by Schneegurt *et al.* (1994), (1997) established that Cyanothece sp strain ATCC 51142 another subgroup V cyanobacterium synthesized a branched glucan which they thought to represent a novel sort of glycogen molecule based on a measured branching ratio of 9%. They had nevertheless noted that the granule size exceeded the theoretical limits imposed on individual particles of glycogen and concluded that the granules contained several distinct glycogen molecules. Looking back on the data supporting this conclusion, we believe it is possible that chemical methylation would have yielded a slight overestimate of the branching ratio. Indeed a mere 20% overestimate would have been sufficient to turn an amylopectin-like candidate into a putative glycogen structure. In fact Cyanothece sp strain ATCC 51142 contains granules with a semi-amylopectin virtually identical to those reported by Nakamura *et al.* (2005) (Deschamps et al., 2008a). In their studies of nitrogen fixation in unicellular cyanobacteria, Schneegurt *et al.* (1994) noted that the carbodydrate granules were synthesized during the day and were being mobilized during the night. They also showed that nitrogen fixation occurred exclusively in darkness and was under circadian clock regulation. Nitrogenase, the enzyme of Nitrogen fixation, is known to define an enzyme exquisitevely sensitive to the presence of O_2 which inactivates it. Because cyanobacteria produce energy through oxygenic photosynthesis there is a conflict between energy production and its utilization for nitrogen fixation. Many cyanobacteria have resolved this conflict through separating in space diazotrophy from oxygenic photosynthesis in distinct specialized cells within a multicellular filament. However unicellular diazotrophic cyanobacteria of subgroup V are unable to do so and

therefore have resorted to separate these processes in time through circadian clock regulation. Schneegurt et al. (1994) proposed that the energy stored in the carbohydrate granules is used both to supply the energy and reducing power required for nitrogenase and to further lower the O_2 level through respiration. Because diazotrophic unicellular cyanobacteria of subgroup V need to store significantly larger amounts of carbohydrates to feed cellular growth, division and diazotrophy, Deschamps *et al*., (2008a) proposed that this yielded a selection pressure for the change of glycogen metabolism into the synthesis of semi-cristalline polymers. Indeed this would enable the storage of larger amounts of osmotically inert carbon with lesser turnover during the light phase. This could indeed explain the appearance of a starch like structures in this particular taxonomic group which contains many important unicellular marine diazotrophic species. Interestingly Wing-Ming et al. (1994) also noted the presence of "irregular polyglucan granules" in another subgroup V isolate: Synechococcus RF-1. However they only noted the unusual large size of the granules without any detailed structural analysis.

Deschamps *et al.* (2008a) made a detailed structural characterization of the carbohydrate granules contained by a marine unicellular cyanobacterium Clg1 isolated by Falcon *et al.* (2004) related to both the genus *Cyanobacterium* and *Crocosphaera* both of subgroup V. Their attention was drawn by the presence of significantly larger granules than those present in Cyanothece sp strain ATCC 51142. A very detailed characterization of the granules was made. Two polysaccharide fractions resembling amylopectin and amylose were purified with chain-length and mass distributions undistinguishable from the plant starch fractions. In addition the granules displayed wide-angle powder X-ray diffraction patterns reminiscent of cereal starches (the so called A-type diffraction pattern (see Buléon *et al.*, (1998)) demonstrating the presence of the same 3-D spatial organization of the amylopectin crystals. Moreover small-angle X-ray scattering demonstrated the presence of the same 9 nm value that typifies the unit amylopectin cluster size (Deschamps *et al*., 2008a). The carbohydrate granules of Cyanothece sp strain ATCC 51142 also displayed an

A-type diffraction pattern further proving that this storage polysaccharide had properties much closer to amylopectin than glycogen (Deschamps *et al*., 2008a). The presence of amylose in the Clg1 starch prompted Deschamps *et al*., (2008a) to look for the enzyme of amylose biosynthesis. GBSSI, an enzyme of cyanobacterial phylogeny never previously reported within cyanobacteria, was thus found bound to the starch granules and was demonstrated to synthesize amylose *in vitro* (Deschamps *et al*., 2008a). Interestingly GBSSI was more highly selective for ADP-glucose than the Archaeplastidal enzymes which in most instances proved to prefer either ADP-glucose or UDP-glucose but nevertheless were able to polymerize amylose from both. We believe this to reflect a distinct history of the Archaeplastidal enzymes. There is thus now enough evidence to support a cyanobacterial origin to starch. In addition, Deschamps *et al*., (2008a) clearly proposed that the plastid ancestor was indeed a cyanobacterial ancestor of subgroup V.

	Cyanobacteria (*C.watsonii*)	Eukaryotes (*E.histolytica*)	Minimal set for the common ancestor.	Green lineage (*O. tauri*)	Red lineage Compiled minimum set
ADP-glucose pyrophosphorylase	1	-	1	2	-
Soluble starch synthase (ADPG)	2	-	2	5	-
Soluble starch synthase (UDPG)	-	1	1	-	1
GBSS I	1	-	1	1	1
Branching enzyme	3	1	1	2	1
isoamylase	1	-	1	3	1
Indirect debranching enzyme	-	1	1	-	1
Phosphorylase	2	2	1	2	1
Glucanotransferase	1	-	1	1	1
Trans-glucosidase	-	2	1	1	1
beta-amylase	-	4	1	2	1
Glucan water dikinase	-	-	(1)	(4)	(1)

Table 1: The number of isoforms found for each class of glycogen/starch metabolism enzymes was listed. Using phylogenetics, we could determine the origin of each isoform in the red and green lineages except for GWDs. Enzymes of cyanobacterial phylogeny are highlighted in blue. Enzymes of eukaryotic origin are highlighted in beige. Enzymes of uncertain origin are listed between brackets. The cyanobacterial eukaryotes and green plants display highly conserved sets of enzymes. We chose *Crocosphaera watsonii*, *Entamoeba histolytica*, and Ostreococcus as paradigm genomes for, respectively, cyanobacteria, heterotrophic eukaryotes, and green plants. The information concerning rhodophytes was compiled from several genomes as explained in the text.

7. Reconstructing starch metabolism in the common ancestor of Archaeplastida

If one accepts a simple vertical inheritance model for the genes of starch metabolism, the monophyly of Archaeplastida allows for the reconstruction of a minimal gene set that must have been present in the ancestor of Archaeplastida to explain the present distribution of genes involved in storage polysaccharide metabolism in the three Archaeplastida lineages. This minimal gene set is displayed in Table 1. In reconstructing this set, Deschamps *et al.*, (2008a) have minimized the number of genes originating from the green lineage to those that clearly displayed a common unique origin in phylogenetic trees, as we believe that most isoforms were generated by gene duplication when the Chloroplastidae and Rhodophyceae diverged.

Table I also displays the phylogenetic (host, cyanobiont or unknown) origin of the pathway enzymes. We chose as a paradigm of the status of storage polysaccharide metabolism of the eukaryote host the enzyme network evidenced in *Entamoeba histolytica* (Loftus *et al.* 2005). We have previously discussed the relevance of this choice by the finding of a richer set and diversity of important enzymes in amoebas. The starch metabolism network of *Crocosphaera watsonii* was chosen as a model subgroup V starch accumulating diazotrophic cyanobacterium. Table I clearly shows that the starch metabolism network of Rhodophyceae and Chloroplastida define a very similar mosaic of enzymes of host and cyanobiont origin (Deschamps *et al*., 2008a). In addition in phylogenetic trees, the common chloroplastidal and rhodophycean enzyme sequences display a common origin (Coppin *et al*., 2005; Patron and Keeling, 2006; Deschamps *et al.,* 2008a). These observations are in complete agreement with Archaeplastida monophyly. The only difference consists of the presence of enzymes of ADP-glucose synthesis and utilization in Chloroplastida and the sole presence of the glycogen synthase from heterotrophic eukaryotes in Rhodophyceae (Table 1). However the common ancestor must have contained all of these distinctive enzymes.

Some enzymes such as GWD, PWD –α-amylase and pullulanase have unknown phylogenetic origins. Despite all efforts we have been unable to locate GWD or PWD-like sequences in lineages independent from the Archaeplastida. It is quite possible that this function evolved shortly after endosymbiosis in the host cytosol. GWD is responsible for amylopectin phosphorylation and thereby initiates starch degradation. There is good evidence for the presence of a glycogen phosphorylation pathway in heterotrophic eukaryotes and no evidence for such a pathway in cyanobacteria. The precise function of this pathway is unknown. Nevertheless, a dysfunctional phosphoglucan phosphatase activity carried by a mutant laforin protein may help explain why highly phosphorylated anomalous glycogen (called lafora bodies) accumulates during lafora's disease in human tissues including the brain yielding a fatal progressive myoclonic epilepsy (Tagliabracci *et al*., 2008). However the nature of the enzyme responsible for glycogen phosphorylation in humans is unknown. Human laforin is known on the other hand to complement the defect in starch mobilization in Arabidopsis due to a mutation in the related sex4 protein (Gentry *et al.* 2007; Kotting *et al.,* 2009). The sex4-laforin function is known to be required during starch mobilization to dephosphorylate the phosphoglucans generated by GWD (Kotting et al., 2009). It is possible that GWD evolved since this unknown kinase was unable to phosphorylate the hydrophobic crystalline amylopectin clusters. The lafora protein was then immediately recruited by GWD to establish a novel pathway of starch mobilization in the host cytosol. This invention can be seen as host mediated since it appeared in the cytosol and entirely relies on other components of host phylogeny such as laforin and β-amylase. Clearly a better knowledge of the function of laforin in glycogen metabolism is required before one can suggest useful scenarios for the appearance of this very important and intriguing pathway.

8. Subcellular localization of storage polysaccharides in the common ancestor of Archaeplastida

The minimal ancestral enzyme set does not tell us where the enzymes were located shortly after endosymbiosis. Three mutually exclusive scenarios can be considered. In a first scenario both the cyanobiont and the host cytosol synthesized storage polysaccharides, in a second scenario only the cyanobiont synthesized such polysaccharides while in a third scenario only the cytosol contained this material.

We very strongly argue that the third scenario defines the only plausible situation. There are several complex reasons for this that is outlined below:

First Henrissat *et al.* (2002) after making a gene content survey of the genomes of several pathogenic bacteria noted that a strong correlation existed between glycogen metabolism loss and a highly dependent parasitic relationship with the infected host. Becoming an endosymbiont would, according to this view automatically lead to the loss of storage polysaccharide metabolism by the symbiont. This seems largely have confirmed by surveying most endosymbiont genomes (Gil et al., 2004). There is only one known example in the literature which is independent of primary endosymbiosis of the plastid that is based on photosynthate export from a cyanobacterium engulfed by a protist. Paulinella indeed carries two cyanobionts (called chromatophores) which are replicated with the host and cannot live as independent organisms (Bodyl *et al*., 2007). The chromatophore genome has recently been sequenced (Nowack *et al*., 2008). It apparently still contains over a thousand genes and still displays a typical cyanobacterial cell morphology and organization. Interestingly another heterotrophic species related to Paulinella that does not shelter chromatophores seems to prey on cyanobacteria that ressemble the latter. This observation pleads for a phagotrophic origin of these symbionts.

The Paulinella genome was proven to lack enzymes of glycogen metabolism while these are universally present in the genomes of free-living cyanobacteria (Nowack *et al*., 2008). Although one might argue that these genes could have been at least in theory transferred to the host nucleus, the present gene content of the chromatophore does not support transfers as extensive as those evidenced during evolution of plastids.

A second argument comes from a close examination of the phylogenetic origin of the minimal ancestral enzyme set displayed in Table 1. This ancestral metabolism includes a nearly complete set of host glycogen metabolism (with the noticeable exception of indirect debranching enzymes which were substituted by direct DBEs). However important enzymes of starch biosynthesis and degradation in cyanobacteria are missing. This involves 3 branching enzymes and 2 phosphorylases. If the gene losses occurred before routine targeting of cytosolic proteins evolved then the cyanobiont would not have been able to sustain storage polysaccharide metabolism. On the other hand if the losses occurred after such a system became routine one can argue that a duplicated branching enzyme gene of host origin may have had its product targeted to the cyanobiont and substituted for the 3 cyanobacterial enzymes which were subsequently lost. In addition the starch phosphorylase would have required a similar replacement by an enzyme of host origin. In this case we would argue that these enzymes are certainly not functionally equivalent to the cyanobacterial phosphorylase. Indeed the latter is able to attack directly solid cyanobacterial starch and release glucose-1-P, a feat which cannot be achieved by the enzymes of eukaryotic origin (Dauvillée *et al*., 2006).

A third argument comes from the unexplainable complexity of the Chloroplastida starch metabolism network (Table 1). Why many enzymes, but not all, have experienced one to two rounds of gene duplications followed by enzyme subfunctionalizations remains a complete mystery. Rhodophyceae which contain equally complex starch granules have not undergone such gene duplications and subfunctionalizations (Table 1). If on the one hand, one assumes that the

cyanobiont has lost the ability to store starch, then the plastidial localization of starch in green algae will have to be generated through a rewiring of the starch metabolism network from the cytosol to the plastid. We will detail in following sections the means by which this could be achieved. Suffice it to say now that this process requires the amount of gene duplications and enzyme subfunctionalizations that is seen in the Chloroplastida network.

If on the other hand storage polysaccharide metabolism had been maintained in the cyanobiont then the transfer to the nucleus of the genes required for their synthesis and degradation would have occurred one gene at the time. If the transferred gene by chance acquired a transit peptide then the corresponding gene on the cyanobacterial genome would have been lost. This process which is similar to that experienced by Calvin cycle genes or other photosynthesis genes does not require any gene duplication and enzyme subfunctionalization and these are indeed not observed in such networks.

Therefore the complexity of the starch metabolism network pleads for an ancient cysolic pathway. Less convincing arguments than the three outlined above equally support the cytosolic localization of the ancient network. Among these is the fact that two out the three Archaeplastidal lineages still synthesize, today their storage polysaccharides exclusively in the cytosol and that the lineage thought to more closely resemble the ancestral Archaeplastida (the Glaucophyta) is among them.

9. Compartmentalization of the ancient pathway of starch metabolism in the common ancestor of Archaeplastida

If the storage polysaccharides were located in the cytosol then most of the enzymes of starch synthesis and degradation must have been expressed in this compartment. This includes the ADP-glucose requiring starch synthases and all other enzymes of cyanobacterial phylogeny. How

could such enzymes initially encoded by the cyanobiont genome be expressed in the cytosol at such an early stage of endosymbiosis? The process by which genes are transferred from the cyanobiont genome to the host nucleus is called endosymbiotic gene transfer (EGT) and is one of the major causes of lateral gene transfers in eukaryotic genomes. To the naive reader this would define an unusual and mysterious phenomenon that is expected to occur at exceptionally low frequency. In fact EGTs are likely to occur at quite high frequencies.

In yeast marker genes that allow growth only when expressed in the cytosol were introduced in the mitochondrial genome. Mutations at the corresponding nuclear locus were introduced that would lead to the absence of growth on selective media. Restoration of growth due to the transfer of the mitochondrial copy in the yeast nuclear genome was observed at the frequency of 10^6, a frequency comparable to that of spontaneous mutations in a given gene (Thorsness and Fox 1990). Similar experimental results with analogous frequencies were obtained with plastidial markers in tobacco (Stegemann *et al*., 2003). Of course EGTs followed by expression of a protein in the cytosol will probably be observed at lower frequencies since the organelle DNA inserted in the nucleus must by chance be located downstream of active promoter sequences. Nevertheless such results leave little doubt that these events were indeed sufficiently frequent. Of course to explain that the nuclear DNA is virtually not filled with organelle sequences, one has to imagine that most of these events will be counter-selected.

Nevertheless, EGTs followed by cytosolic expression of proteins from the cyanobiont will define the first kind of EGT recorded and can take place immediately at the time of endosymbiosis (or even before if the host preys through phagocytosis on the future endosymbionts). The more classical type of EGT that requires the expression of the protein product in the cyanobiont (such as enzymes of photosynthesis) will have to await the later development of a complex protein targeting machinery able to routinely readdress the corresponding proteins to the evolving organelle. Thus if the cytosolic expression of the cyanobacterial genes of starch metabolism gives

some advantage the corresponding EGT will be selected and maintained. The problem therefore is to understand what would have been the advantage for the host cytosol to harbor a dual substrate biochemical pathway based on both ADP-glucose and UDP-glucose. Indeed if one assumes that ADP-glucose pyrophosphorylase, the enzyme of ADP-glucose synthesis was also subjected to EGT and the enzyme transferred to the cytosol, then it becomes very hard to understand what benefit the cell would get from having cytosolic glucose-1-P funneled into both UDP-glucose and ADP-glucose.

However we argue that ADP-glucose pyrophosphorylase is a highly unlikely target for cytosolic expression since its substrate and allosteric effectors couples the enzyme tightly to photosynthesis. Indeed the regulation of the enzyme through the 3-PGA/Pi ratio has been maintained throughout evolution (Ballicora *et al.*, 2003). If the coupling of the enzyme activity to the Calvin cycle and to ATP production is maintained then it is more reasonable to consider that it remained expressed in the cyanobiont from the cyanobacterial genome and that it did not define a likely candidate for an early EGT leading to its expression in the cytosol.

Therefore, for the enzymes of cyanobacterial origin to be immediately functional for starch synthesis in the cytosol, one has to assume that the ADP-glucose produced within the cyanobiont was exported from the cyanobiont to the cytosol. For such a transport to be effective Deschamps *et al*. (2008a) proposed that a nucleotide sugar translocator (NST) of host origin able to exchange ADP-glucose for AMP or ADP was targeted to the inner membrane. The flux created by this proposal is displayed in Figure 5. It is the only hypothesis that yields a selective advantage for the presence of a cytosolic dual substrate pathway for starch biosynthesis. Indeed the flux depicted in Figure 5 is no more no less than the biochemical flux through which photosynthate was exported from the cyanobiont to its host. This would thus be the carbon flux on which the success of plastid endosymbiosis would rest.

10. The proposed flux of carbon through starch metabolism explains the establishment of plastid endosymbiosis

At the core of plastid endosymbiosis lies a biochemical link whereby the carbon produced through photosynthesis by the cyanobiont is exported to the eukaryotic cytosol to feed the host metabolism. This link was essential for establishment of the plastid and introduction of oxygenic photosynthesis in eukaryotes. The problem with the establishment of such a link relies on the fact that both partners of endosymbiosis are both completely unrelated and independent implying that there is no connection cross talk or possibility of regulation between the two unrelated biochemical networks. Yet the carbon must be exported only when the cyanobiont allows for it and used only when the host needs it. If carbon flows out the cyanobiont in the form of just any metabolite translocated at any time because of the targeting of some host transporter on the cyanobiont inner's membrane then the net result for the homeostasis of the cyanobiont's metabolism is likely to be at best toxic at worst lethal. A sudden burst of metabolites originating from the cyanobiont's metabolism into the host cytosol can, depending on the nature of the latter also affect negatively the host physiology.

Clearly some buffer is needed between the two unrelated biochemical networks because of their asynchronous supply and demand for carbon. It is very hard to imagine a flux better adapted to such constraints than that displayed in Figure 5. ADP-glucose is one of the very few metabolites that can freely flow out of the cyanobiont without affecting it. Indeed the carbon flowing through ADP-glucose pyrophosphorylase is anyhow that part of bacterial metabolism which is committed to storage that is to escape cyanobacterial metabolism and become at least temporarily unavailable into solid starch. Translocating ADP-glucose out of the cyanobiont and having the carbon stored in the cytosol will make very little difference.

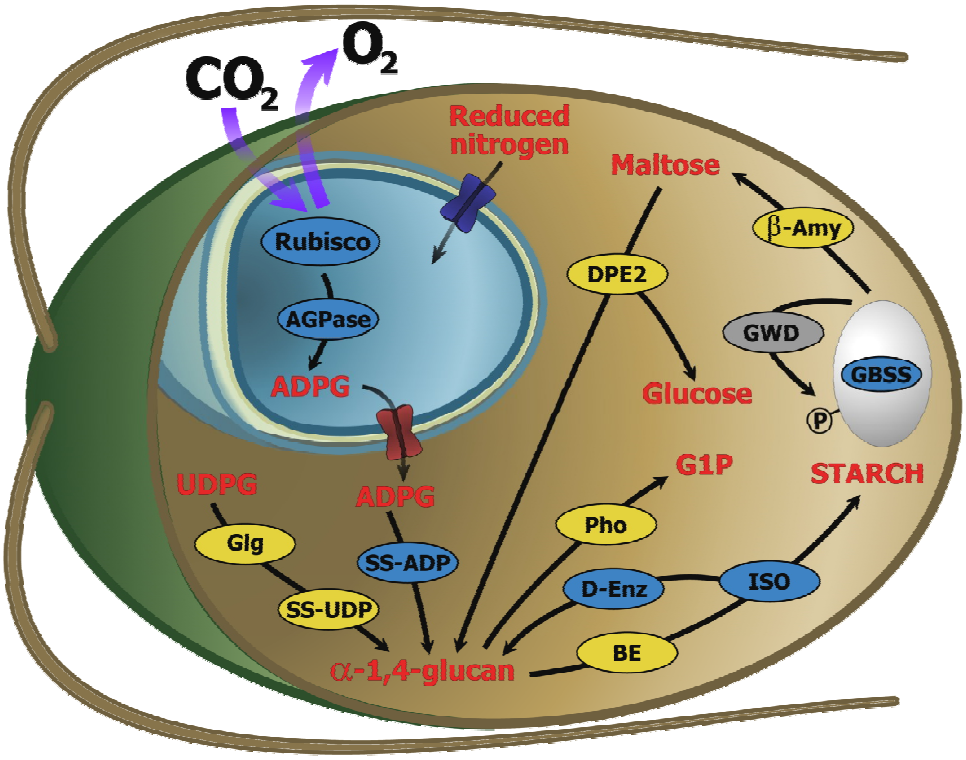

Figure 5: Ancient symbiotic fluxes in the ancestor of all plants. Within the cyanobiont, ADP-glucose pyrophosphorylase (AGPase) responds to photosynthate availability by synthesizing ADP-glucose (ADPG) that is normally committed to storage. The glycosyl-nucleotide is transported through a nucleotide–sugar/triose phosphate translocator that originated from the host endomembrane system as was recently proposed (Weber, AP. et al., 2006). ADP-glucose is polymerized into starch with no interference with the host pathways. Synthesis involves an ADP-glucose requiring soluble starch synthase (SSADPG) that is branched by branching enzyme (BE) and further matured into packaged starch through the action of isoamylase (iso) and D-enzyme (D-enz). Independently from photosynthesis and the cyanobiont, the host is still able to feed glucose into storage through the use of a glg-primed (glycogenin) UDP-glucose requiring soluble starch synthase. Glucose mobilization from starch will depend entirely from the host needs through host enzymes that include phosphorylases (Pho), α -amylases (α -Amy), and a maltose-specific a-1,4 glucanotransferase (DPE2). The efflux of ADPG from the cyanobiont will render the latter unable to fix nitrogen during the night. This in turn will require the host to feed the cyanobiont with reduced nitrogen. The phylogenetic origin of each enzyme is displayed either by a blue (cyanobacterial origin) or by a red (host origin) color. The cyanobiont is represented in blue and is thought to display morphology and pigment composition similar to the cyanelles (plastids) from glaucophytes. The red- labeled α-1, 4 glucan represents a small size pool of maltooligosaccharides generated during starch biosynthesis and degradation. The circled P represents starch phosphorylated by GWD, an enzyme of unknown origin required for degradation. The starch granule is represented in white with GBSS. GBSS defines the only enzyme active within the polysaccharide semicrystalline matrix and is responsible for amylose synthesis.

In addition as ADP-glucose enters the cytosol it has virtually no impact on host metabolism since this glycosyl-nucleotide is not neither produced nor recognized by eukaryotes. Because of this absence of recognition ADP-glucose's only fate will be to be used through a cyanobacterial enzyme that has been transferred by EGT to the host cytosol. The GT5 ADP-glucose requiring starch synthase will transfer glucose to the non-reducing end of outer chains of storage

polysaccharides and release ADP. ADP will reenter the cyanobiont through the exchange reaction mediated through the ADP-glucose NST working on the cyanobiont inner's membrane. This will be either in the form of ADP directly or more likely AMP since most known NSTs are documented to exchange the glycosylnucleotide with the corresponding nucleotide monophosphate. The remaining phosphate will in the latter case re-enter the cyanobiont by other cyanobacterial transporters. The cycle is thus balanced with respect to purine nucleotides and orthophosphate and results in the net incorporation of one glucose residue to a cytosolic storage polysaccharide. This glucose comes from carbon fixation from the cyanobiont while the energy required to polymerize it into storage comes from cyanobacterial photophosphorylation responsible for producing ADP-glucose, the activated form of the carbohydrate. Having the carbon stored into glycogen or starch will efficiently buffer demand and supply for carbon. Indeed the pathway of storage polysaccharide breakdown is entirely under control of host enzymes and will therefore respond specifically to host demand. We believe that having both the first ADP-glucose requiring starch synthase working in the cytosol together with the ADP-glucose NST targeted to the inner cyanobiont membrane would have been sufficient to successfully establish the endosymbiotic link. This would have initially leaded to the synthesis of additional amounts of cytosolic glycogen. The switch to starch would have been brought about later thanks to EGTs of other cyanobacterial or chlamydial genes (see below) such as isoamylase responsible for the synthesis of crystalline amylopectin and dpe1 the α-1,4 glucanotransferase needed to process the chains released by isoamylase. This was followed by EGT of GBSSI a cyanobacterial gene responsible for amylose synthesis (an enzyme initially specific for ADP-glucose) within the solid starch granule (GBSSI is unable to use glycogen (Dauvillée et al., 1999)). These EGTs were facilitated by the fact that the endosymbiont was related to extant subgroup V cyanobacteria and thus contained the required genes. However during this whole process the cyanobiont itself remained unable to synthesize storage polysaccharides possibly

because it had already lost critical genes such as the starch branching enzymes and starch phosphorylases. The host however remained able to feed carbon into glycogen or starch from UDP-glucose. Indeed cytosolic glycogen was initially present in the host to enable it to siphon out of cellular metabolism the excess carbon produced through host metabolism for later use. This useful physiological function was maintained throughout and explains the presence of a dual substrate pathway. The change from cytosolic glycogen to starch was brought about because semi-crystalline polysaccharides were better adapted to trap the vast amounts of glucose that became available through endosymbiosis. This also ensured that a permanent strong carbon sink existed in the cytosol even if vast amounts of glucose had already been stored. The switch to starch required the only innovation in the storage polysaccharide pathways that came about during plastid endosymbiosis: the evolution of GWD. This innovation would have been required if one considers that cyanobacterial starch phosphorylase genes which unlike the corresponding eukaryotic enzymes display the ability to digest solid starch had already been lost. As mentioned earlier this innovation was built on a pre-existing eukaryotic pathway of glycogen phosphorylation whose precise function in glycogen metabolism requires further research (Tagliabracci et al., 2008).

11. Discovering the missing link of eukaryotic photosynthesis

When the flux displayed in Figure 5 was first proposed we had no evidence for the existence of an ancient ADP-glucose translocator. Nevertheless this ancient NST would have been central to the success of endosymbiosis as it established the first link between the endosymbiont and its host. Weber *et al.* (2006) published phylogenies of the major carbon transporters of the inner plastid membrane of Rhodophyceae and Chloroplastida. These plastidial phosphate translocators (pPTs) exchange a number of distinct triose-phosphates, hexose-phosphates, xylulose-phosphate,

PEP for orthophosphate. All these diverse pPTs from both Rhodophyceae and Chloroplastida can be traced back to a unique ancestor that was proposed to define the translocator that was responsible for supplying the host cytosol with photosynthate from the cyanobiont during endosymbiosis (Weber *et al.*, 2006). The common ancestor of the pPTs defined a sister lineage to a group of transporters from the eukaryotic host endomembrane system. Weber et al. (2006) thus suggested that these host transporters were recruited and targeted to the cyanobiont inner membrane to establish the endosymbiotic link. However, the eukaryotic endomembrane transporters all defined nucleotide sugar translocators (NSTs) that exchanged glycosyl-nucleotides for the corresponding nucleotide monophophate and not sugar-phosphate-orthophosphate translocators (Rollwitz et al., 2006). In addition the particular NST3 group concerned (family NST3 according to Martinez Duncker *et al.*, 2003) defines the only NST family reported to contain, among others, purine sugar nucleotide transporters such as GDP-mannose, GDP-arabinose or GDP-fucose. It thus occurred to us that if the carbon translocator responsible for establishing the endosymbiotic link was so close phylogenetically to purine sugar nucleotide transporters then it might very well have defined the ADP-glucose translocator required by the flux proposed in Figure 5. NSTs in general display high selectivity with respect to the nature of the base and modest selectivity with respect to the sugar (Handford *et al.*, 2006). ADP-glucose is not reported in eukaryotes but other purine-nucleotide sugars such as GDP-mannose, GDP-arabinose or GDP-fucose are common. If such a GDP-sugar translocator displayed a fortuitous ability to translocate ADP-glucose it could have been recruited on the cyanobiont's inner membrane during endosymbiosis. This would imply that if ancient NST3-like translocators displayed such properties then maybe extant NST3 transporters should still display them. Colleoni *et al.* (2010) thus very recently investigated the kinetic properties of two GDP-mannose transporters from yeast and plants in yeast membrane derived liposomes. They demonstrated that AMP acted at physiological concentrations as a very potent competitive

inhibitor of both the yeast and the plant GDP-mannose translocator. In addition they demonstrated that the plant enzyme was able to exchange ADP-glucose for AMP at rates similar to those of the GDP-mannose GMP exchange Colleoni *et al.* (2010). Nevertheless the plant enzyme displayed a relatively high Km for ADP-glucose (7 mM). Would such a low affinity for ADP-glucose have allowed the export of photosynthate from the cyanobiont? In the discussion of their findings Colleoni *et al.* (2010) argue that on the periplasmic side of the cyanobiont's inner membrane the high sensitivity of the translocators to AMP would have been sufficient to out compete GMP and GDP-mannose. On the cyanobiont's side, the authors argue that ADP-glucose would have defined the only purine nucleotide sugar available. They also emphasize that the concentrations of the latter would have been likely to exceed 1 mM because of a block in ADP-glucose utilization. Indeed having lost the ability to synthesize storage polysaccharides, the cyanobiont can be compared to a mutant blocked in the utilization of ADP-glucose. Therefore the present biochemical properties of Golgi GDP-mannose translocators are sufficient to generate an efflux of ADP-glucose provided of course the protein is correctly targeted to the symbiont's inner membrane. Several options are discussed by Colleoni et al. (2010) concerning this targeting: these include the existence of either a primitive vesicular transport system or ancestral simplified TOC machinery. It must be recalled that the plant TPT (triose phosphate translocator) localizes to mitochondrial membranes when expressed in yeast without its transit peptide (Loddenkötter et al., 1993).

The results reported by Colleoni et al. (2009) of course do not imply that the ancestors of the corresponding GDP-mannose translocators defined the hyptothetical missing link for the establishment of eukaryotic photosynthesis. In fact, most probably any NST3 purine nucleotide sugar transporter would have probably been suited for this purpose provided it displayed a sufficiently high affinity for ADP-glucose. How then, can the disappearance of this translocator and its replacement by the pPTs be explained? Colleoni et al. (2010) proposed that as the

metabolic integration of the evolving plastid proceeded, an opportunity arose to duplicate the genes corresponding to existing plastid inner membrane transporters and to evolve more integrated routes of carbon exchange through the evolution of novel activities. The ADP-glucose translocator and its derived pPT family of transporters co-existed as long as starch was being polymerized in the cytosol from ADP-glucose. However when the three Archaeaplastida lineages diverged this ability was lost by the Rhodophyceae and Glaucophyta when both lineages lost the ability to elongate starch from ADP-glucose. As to the Chloroplastida the ADP-glucose transporter was lost at the final stage of starch metabolism rewiring to the chloroplast.

12. The rewiring of starch metabolism to chloroplasts involves extensive gene duplications and enzyme subfunctionalizations

The simple cytosolic pathway depicted in Figure 5 generated the corresponding simple floridean starch synthesis and degradation pathways through the selective losses of ADP-glucose pyrophosphorylase, of the ADP-glucose translocator and of the ADP-glucose specific starch synthases. Indeed the rhodophycean pathway is characterized by one isoform for each step of starch synthesis and degradation (Table 1). A simple pathway was also evidenced in secondary endosymbiosis lineages such as the apicomplexan parasites (Coppin *et al*., 2005) and is suspected to at work in the Glaucophyta (Plancke *et al*., 2008). The Chloroplastida on the other hand experienced a rewiring of starch metabolism back to the compartment where it originated: the chloroplast (Deschamps *et al*., 2008b). The means by which a gene will have its product rewired to plastids are quite clear: it must by chance acquire a transit peptide or another adequate targeting sequence and it will then be expressed in the organelle.

However an isolated enzyme from a biochemical pathway rewired by chance to a novel location has little chance to yield a benefit to the cell. Most likely its isolated presence will yield a penalty

and will be counter-selected. The means by which an entire pathway becomes targeted to plastids is thus obscure. This problem was first raised by Michels and Opperdoes (1991) when they were trying to understand how sections of the glycolytic pathway could have been rewired to the peroxisome thereby generating the glycosome in trypanosomes. Nevertheless biochemical pathways do get redirected to novel locations. As Martin (2010) recently pointed out it is reasonable to assume that dual targeting and mistargeting of small amounts of enzyme activities could explain the means by which pathways can be reconstructed since an entirely isolated enzyme would thus not exist. It was argued that the redirected enzyme will find minute yet sufficient amounts of mistargeted partners from its original pathway to ensure maintenance of the redirected enzyme until the next wave of transit peptide acquisitions (Martin, 2010).

We agree that very simple pathways consisting of few distinct steps where a defined balance of enzyme activities is in addition not critical may be rewired in this fashion. We however disagree that such a process would apply to starch metabolism as a whole although it might apply to sections of it. First of all starch metabolism defines a non linear biochemical pathway where for instance the product of elongation becomes a substrate for branching which regenerates a different type of substrate for the elongation enzyme in a more cyclic than linear fashion (Ball and Morell, 2003). Polysaccharide debranching also produces substrates for further elongation and branching and debranching (Ball and Morell, 2003). In addition balance of enzyme activities are certainly critical and the presence of multienzyme complexes may be required to ensure that the right stochiometry and balance of enzyme activities is indeed achieved (Tetlow et al., 2004b; Hennen-Bierwagen et al., 2008). Everything we know concerning starch metabolism tells us that the simultaneous acquisition of transit peptide by all the required enzymes and the achievement of the proper activity balance in one step is absolutely impossible. Yet unlike most biochemical pathways the same set or subset of enzymes can yield several very different end-products that are indeed found in different lineages. In fact storage polysaccharides can be found either in the form

of unbranched malto-oligosaccharides, in the form of glycogen or in the form of starch. Deschamps et al. (2008c) argued that this flexibility of the pathway enabled the evolving green algae to redirect storage polysaccharide metabolism in three steps of increasing biochemical complexity corresponding to the three different types of end-products. Interestingly these three forms of storage glucans correspond to three pools of carbohydrates of increasing size. Linear MOS consist of α-1,4 glucan chains up to a degree of polymerization of 20. Chains of longer size will spontaneously align and crystallize making them inaccessible and potentially cytotoxic. Nevertheless these carbohydrate stores will increase the amount of glucose readily accessible to cellular metabolism. Small MOS on the other hand cannot be accumulated to very high levels as they are quite active osmotically. Glycogen is comparatively much less active osmotically and readily hydrosoluble. Glycogen will thus define a stage where larger pools of carbohydrates can be stored in an osmotically tolerable form. Finally starch being semi-cristalline and solid will define a stage where huge amounts of carbohydrates can be stored in an osmotically inert form although mobilization of these stores will be more complex than those of glycogen. Whatever reason(s) prompted the return of storage glucan(s) to plastids, this (these) reason(s) seem to have acted selectively in the emerging Chloroplastida lineage (Deschamps *et al*., 2008b). Indeed three events seemed to correlate with the emergence of the Chloroplastida: the evolution of novel chlorophyll b-containing light harvesting complexes, the return of starch metabolism in plastids and an increased complexity of starch metabolism. Indeed a bioinformatic analysis carried out in 6 green alga genomes demonstrated that in the earliest diverging prasinophyte algae (the mamelliales), and at variance with other pathways, the full complexity of starch metabolism witnessed in higher plants is already present (Deschamps *et al*., 2008b).

Deschamps *et al.* (2008c) proposed that the evolution of novel LHCs in green algae might have propelled the return of storage glucans to plastids. Indeed higher plants defective for the import of ATP at night were demonstrated to experience photooxidative stresses that depended on the

size of the plastidial starch pools (Reinhold et al., 2007). This was due to the accumulation of protoporphyrin IX an intermediate of chlorophyll biosynthesis due to the high requirement of magnesium chelatase for ATP (Reinhold *et al*., 2007). Because the flux to chlorophyll is suspected to have increased in Chloroplastida so would the need for a plastidial source of ATP at night (Deschamps *et al*., 2008c). Whether or not resistance to photooxidative stresses prompted the return of storage glucans to plastids, this return seems in any case to have required an increase in the pathway's complexity.

Deschamps et al. (2008b); (2008c) argues that the return to the chloroplast and starch metabolism complexity are mechanistically linked. The problems inherent to a rewiring of whole suite of enzymes have been outlined above. However such a return can be imagined if the whole process happened in sequential steps with 2 fully functional intermediate stages involving first MOS production then glycogen synthesis (Deschamps *et al*., 2008c). Until starch biosynthesis was finally achieved the plastidial carbohydrate pools remained significantly smaller than those present in the form of cytosolic starch. However the appearance of starch in the plastid correlated with the loss of this polymer in the cytosol as two major pools of storage carbohydrates in distinct cellular compartments have indeed never been reported.

The key to understand the extent of subfunctionalizations witnessed in Chloroplastida is to imagine that the two intermediate stages, the accumulation of malto-oligosaccharides or glycogen lasted long enough for those enzymes that have been redirected to the plastid to accumulate mutations that would have optimized their activities with respect to the products they synthesized. For instance if a starch synthase gene from the cytosolic starch metabolism network was duplicated and acquired a transit peptide for its localization within plastids at the time of MOS synthesis, then this gene would accumulate mutations optimizing its function for that purpose. This would entail in this case the ability to prime the reaction at increased rates and to elongate short oligosaccharides that remain hydrosoluble.

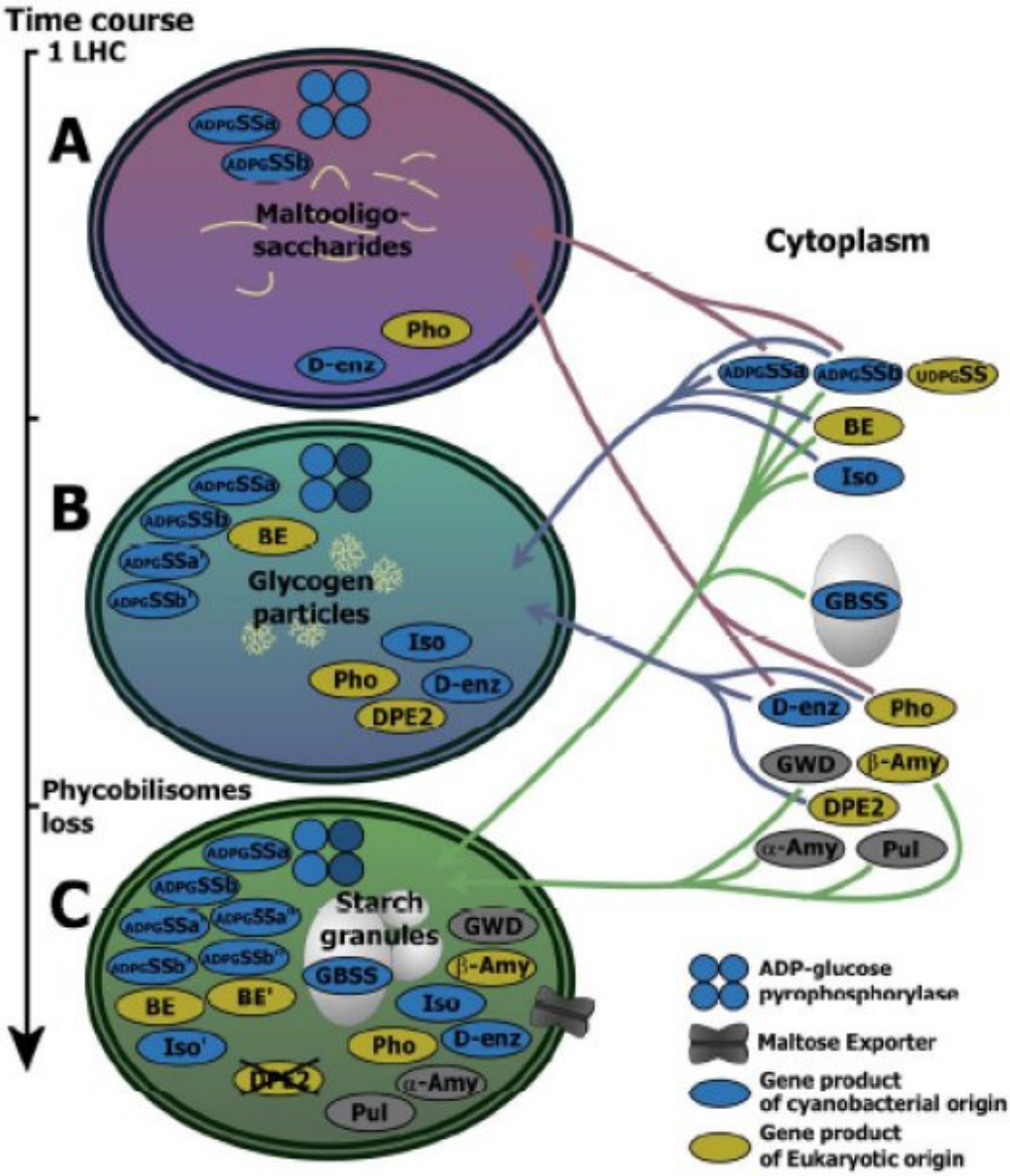

Figure 6: Starch metabolism rewiring during evolution of Chloroplastida. The three panels represent plastids corresponding to three distinct stages of Chloroplastidae evolution. (a) The common ancestor of Rhodophyceae and Chloroplastida, containing phycobilisomes chlorophyll a and cytosolic starch. (b) An intermediate stage of LHC diversification of an ancestor that still contains phycobilisomes and cytosolic starch in addition to more diverse chlorophyll-*a*- and chlorophyll-*b*-containing LHCs. (c) The 'final' stage of chloroplastidae evolution, where both with phycobilisomes and cytosolic starch have been lost. In (a) and (b), cytosolic starch is synthesized from both UDP- glucose (UDPG) and ADP-glucose (ADPG), the latter being exported from plastids, as detailed in Figure 5. The color of each enzyme represents its phylogenetic origin. Each arrow corresponds to stage-specific duplications of genes encoding the cytosolic paralogs concerned. The purple arrows represent those gene duplications encoding cytosolic paralogs that were required for the synthesis and degradation of a small pool of malto-oligosaccharide (MOS) (a). The blue arrows display those gene duplications encoding cytosolic paralogs that were required for the synthesis and degradation of an average-sized pool of glycogen (b). Superscripts 'and "are used to symbolize the successive rounds of duplications and subfunctionalizations. The green arrows represent those gene duplications

encoding cytosolic paralogs that were required for the synthesis and degradation of a large pool of plastidial starch (c). The appearance of the maltose export protein was accompanied by the disappearance of cytosolic starch and of plastidial transglucosidase (symbolized by a cross on the corresponding enzyme). The arrow on the left represents time, which is accompanied by a progressive increase in chlorophyll biosynthesis and light-harvesting capacity. The ensuing progressive increase in plastidial polysaccharide synthesis modified the requirements for the size of the plastidial ADPG pool and was accompanied by duplication and subfunctionalization of ADPG pyrophosphorylase into large (dark blue) and small (blue) subunits.

When a selection pressure reappeared to further increase the plastidial storage carbohydrate pools by evolving to the synthesis of glycogen then it is easier to imagine that an enzyme from the cytosolic starch metabolism network was duplicated and its product redirected to plastids. Indeed it would probably have been longer to accumulate mutations in the aforementioned MOS synthase to turn the enzyme into a polysaccharide synthase rather than just duplicate a cytosolic enzyme gene whose product was already able to cope with branched substrates. When the novel polysaccharide synthase landed within the plastids it will not have lead to the loss of the MOS synthase. Indeed the latter would have been more efficient in glucan priming and the synthesis of short oligosaccharides leading to subfunctionalization of the two enzymes into their respective specializations. Such reasoning applies to all enzymes of starch metabolism and the reader is referred to a recent review for further details (Deschamps *et al*., 2008c). The general consequence of this kind of networking is that the amount of gene duplications and subfunctionalizations experienced will depend on the one hand on the initial complexity of the cytosolic starch metabolism pathway and on the other hand on the number of times a certain type of activity is required to move from one stage to the next. For instance elongation by glucan synthases is required at all three stages while branching is only required for glycogen and starch synthesis. The expected numbers of enzymes corresponding to starch synthase or branching enzyme isoforms will be respectively of 6 and 2 since the initial cytosolic network contained respectively 2 and 1 isoform of each. Similarly debranching will only be required for two stages while disproportionating enzyme or GBSSI will only be required respectively at the MOS or starch stage yielding two or one isoforms. As mentioned previously the appearance of starch in plastids

correlated with its loss in the cytosol. Most of the genes of the cytosolic pathway were lost except those that were recruited to establish the cytosolic pathway of maltose degradation (cytosolic phosphorylase and dpe2) (reviewed by Fettke *et al.*, 2009).

Activity	Theoretical isoform numbers	Observed number of isoforms (Ostreococcus)
ADP-glucose pyrophosphorylase	3	2
Soluble starch synthase (ADPG)	6	5
GBSS I	1	1
Branching enzyme	2	2
isoamylase	4 or 2	3
Phosphorylase	2	2
Glucanotransferase	1	1
Transglucosidase	1	1
☙-amylase	2	2
Glucan water dikinase	2	2

Table 2: A) Theoretical isoform numbers observed from Chloroplastida starch metabolism network. **B)** Ostreococcus contain a comparable number of genes involved in starch metabolism. C) Phosphoglucan-water dikinases were counted with the glucan-water dikinases.

The establishment of starch in the plastid thus required the evolution of a novel transporter exporting maltose the major product of starch catabolism to the cytosol. This mex (maltose export) protein is apparently the only innovation of the Chloroplastidae with respect to starch metabolism as it is not found elsewhere (Niittylä *et al*., 2004).

Figure 6 summarizes the rewiring process while Table 2 displays the predictions of isoform numbers with this evolution scenario. There is a general good agreement between the theoretical isoform numbers and those observed in the Chloroplastida starch metabolism network. Such an unexpected agreement strongly supports the proposal that the ancient starch metabolism pathway was indeed cytosolic.

13. Chlamydial genes in the starch pathway: evidence for an ancient "ménage à trois"

One of the most surprising recent findings of phylogenetic inference consists in the unexpected relationship existing between the genes of Archaeplastida and Chlamydiae (reviewed in Horn, 2008). Chlamydiae define obligate intracellular gram-negative bacterial pathogens intitially documented in animals and more recently found as "symbionts" of protists notably amoebas (Horn, 2008). Some of these protists are also thought to be able to act as reservoirs of pathogenic Chlamydiae. Chlamydiae enter cells at the "elementary body" stage through phagocytosis and multiply only within host vacuoles in the form of reticulate bodies. Chlamydiae typically import their ATP from the host cytosol through an ATP/ADP translocase. An increasing diversity of Chlamydia like microorganisms have been recently identified often showing greater metabolic capabilities (and genomes) than the human pathogens but with similar obligate intracellular life cycles (Horn, 2008). Bioinformatic analysis have yielded 150 out of 2031 protein coding sequences which in the genome of *Pseudochlamydia amoebophila* display phylogenies grouping the latter with the *Archaeplastida* (Horn et al., 2004; Horn 2008).

The more recent acquisition of the red alga *Cyanidioschizon merolae* and of diatom genomes sequence reveal between 39 to 55 genes in these genomes of probable Chlamydial ancestry (Huang and Gogarten, 2007; Becker *et al*., 2008; Moustafa *et al*., 2008). The data are suggestive of a very ancient origin of this particular relationship probably at the time of endosymbiosis. However it is very hard to decide if the Chlamydia donated genes to the Archaeplastida or if they took up such genes during this ancient relationship. One obvious case where the Chlamydia donated a function which can be considered as important for the establishment of endosymbiosis has been documented when analyzing the phylogeny of the NTT transporters (Linka *et al*., 2003). These transporters have been recently proved to import ATP at night within plastids (Reinhold et

al., 2007). Such a function derived from the ATP/ADP translocase at the heart of the parasitic way of life of the Chlamydiae (Trentmann et al., 2007) would have been indeed of importance for a cyanobiont that we argue was deprived of storage polysaccharides. The starch pathway contains two candidates for horizontal transfer of Chlamydial genes: the isoamylases (Huang and Gogarten, 2007; Becker et al., 2008; Moustafa et al., 2008) and one particular soluble starch synthase (the ancestor of SSIII and SSIV) (Moustafa et al., 2008). Isoamylase might very well define a case of LGT (lateral gene transfer) similar to the ATP/ADP translocase. The transfer of the isoamylase gene which switched glycogen to starch synthesis was of importance since it increased the pool of polysaccharides resulting from the export of photosynthate to the host cytosol. As to the soluble starch synthase, phylogenies grouping only the ancestral Chlamydiae with the cyanobacteria within the soluble starch synthase trees argue either for an LGT from cyanobacteria or Archaeplastida to the Protochlamydia genome after separation of the Protochlamydia lineage from the pathogenic Chlamydia or less likely for an LGT from Protochlamydia to plants (Moustafa *et al.*, 2008).

One of the most parcimonious explanations for all these surprising observations would be that the protist lineage, which would enter into a symbiosis with a particular cyanobacterium, was routinely infected by an ancestor of extant Chlamydia (Huang and Gogarten, 2007). It is reasonable to assume that the presence of a cyanobiont exporting photosynthate would have benefited both the host and its energy-parasite Chlamydia. In this “ménage à trois” the cyanobacterium that had lost its storage polysaccharides donated a number of critical genes for expression in the host cytosol enabling it to export photosynthate (possibly ADP-glucose-requiring starch synthases) while the host recruited one of its NSTs and targeted it to the cyanobiont inner membrane. The chlamydial parasite donated its ATP/ADP translocase to the cyanobiont enabling it to withstand in darkness the negative consequences of the loss of its storage polysaccharides. Finally the Chlamydia parasite may have been the source of the

Archaeplastida isoamylases turning the pool of cytosolic glycogen into starch thereby further increasing photosynthate export. In addition if the relationship of the SSIII-SSIV resulted from an LGT from Chlamydia to plants (and not the reverse) then one could even consider that the Chlamydiaea not only facilitated but was essential to plastid endosymbiosis. Indeed as we mentioned previously the cyanobiont was devoid of storage polysaccharides and could already have lost a number of important genes of starch metabolism when it was internalized. In such a context the supply of critical related functions by the energy parasite could have proven essential to the success of endosymbiosis. Clearly, in many instances, further detailed phylogenetic analysis is required to reveal the LGT's direction (from Chlamydia to Archaeplastida or the reverse). Present-day plants and algae are not known to be subjected to infection by Chlamydia. Therefore the ménage à trois was terminated at a very early stage of Archaeplastida evolution possibly before divergence of Chloroplastida, Rhodophyceae and Glaucophyta. According to Huang and Gogarten (2007), this happened when the Archaeplastida's ancestor switched from a mixotrophic way of life involving both phagocytosis and taping carbon and energy from its cyanobiont to an autotrophic way of life. This could have resulted very simply by the loss of phagocytosis. This in turn would have prevented penetration of the energy parasite which is known to use phagocytosis to enter their hosts. The building of rigid and complex cell walls would further strengthen this immunity to infection by Chlamydia. It is tempting to speculate that recurrent infection of a protist by Chlamydia may have played an important role in endosymbiosis not only by supplying critical functions in the establishment of the endosymbiotic link but also by selecting out a lineage that would better tolerate the energy parasitism strain exerted by the Chlamydiae. Those protists that successfully achieved endosymbiosis would indeed be at an advantage over those that did not achieve it and relied entirely on phagocytosis to withstand the strain imposed by the parasites.

Objective of the Thesis

As outlined above, an ancestor of unicellular diazotrophic cyanobacteria of subgroup V may have defined the cyanobiont during endosymbiosis. Starch-like polymers are indeed found within this subgroup. We believe that functions of enzymes in a biochemical network depend more on their evolutionary history rather than any necessity pertaining to the function. The starch metabolism network in plants has a complex historical record. In order to fully understand it, we need to know how the storage polysaccharide metabolism networks functioned in the cyanobiont on the one hand and the host on the other hand before endosymbiosis. In order to get this understanding we have chosen to dissect genetically starch metabolism in subgroup V cyanobacteria. The goal of my thesis was thus to characterize the structure of starch in these cyanobacteria and to master their culture to the point where starch metabolism defective mutants could be selected. Because of the extremely long generation time and ensuing slow growth, we have succeeded in selecting the mutants but will have to leave the characterization of these useful tools to others.

Introduction Bibliography

Introduction bibliography

Adl SM et al. 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *Journal of Eukaryotic Microbiology* **52**, 399-451.

Alonso-Casajús N, Dauvillée D, Viale AM, Muñoz FJ, Baroja-Fernández E, Morán-Zorzano MT, Eydallin G, Ball S, Pozueta-Romero J. 2006. The role of glycogen phosphorylase in glycogen breakdown and structure in *Escherichia coli. The Journal of Bacteriology* **188**, 5266-5272.

Aury JM et al. 2006. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature* **444**, 171-178.

Ball SG, Guan HP, James M, Myers A, Keeling P, Mouille G, Buléon A, Colonna P, Preiss J.1996. From glycogen to amylopectin: a model explaining the biogenesis of the plant starch granule. *Cell* **86**, 349-352.

Ball S, van de Wal M, and Visser R. 1998. Progress in understanding the biosynthesis of amylose. *Trends in Plant Sciences* **3**, 462-467.

Ball SG, Morell MK. 2003. Starch biosynthesis. *Annual Review of Plant Biology* **54**, 207-233.

Ballicora MA, Iglesias AA and Preiss J. 2003. ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogensynthesis. *Microbiology and Molecular Biology Reviews* **67**, 213–225.

Becker B, Hoef-Emden K, Melkonian M. 2008. Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. *BMC Evolutionary Biology* **8**, 203.

Blennow A, Nielsen TH, Baunsgaard L, Mikkelsen R, Engelsen SB. 2002. Starch phosphorylation: A new front line in starch research. *Trends in Plant Science* **7**, 445–450.

Bodyl A, Makiewicz P, Stiller JW. 2007. The intracellular cyanobacterial of *Paulinella chromatophora*: endosymbionts or organelles? *Trends in Microbiology* **15**, 295-296.

Boos W, Shuman H. 1998. Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiology and Molecular Biology Reviews* **62**, 204-229.

Buléon A, Colonna P, Planchot V, Ball S. 1998. Starch granules: structure and biosynthesis. *International Journal of Biological Macromolecules* **23**, 85-112.

Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. 2009. The Carbohydrate-Active Enzymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research* **37**, D233-238

Carlton JM, Hirt RP, Silva JC et al. 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* **315**, 207-212.

Cavalier-Smith T. 2006. Cell evolution and Earth history: stasis and revolution. *Philosophical Transactions of the Royal Society* **B 361**, 969-1006.

Colleoni C, Dauvillée D, Mouille G, Morell M, Samuel M Slomiany MC, Liénard L, Wattebled F, D'Hulst C, Ball S. 1999. Biochemical characterization of the *Chlamydomonas reinhardtii* alpha-1,4 glucanotransferase supports a direct function in amylopectin biosynthesis. *Plant Physiology* **120**, 1005-1014.

Colleoni C, Linka M, Deschamps P, Handford MG, Dupree P, Weber APM, Ball SG. 2010. Identification of a missing link required for the endosymbiotic origin of eukaryotic photosynthesis. *Molecular Biology and Evolution*. Submitted.

Comparat-Moss S, Denyer K. 2009. The evolution of the starch biosynthetic pathway in cereals and other grasses. *Journal of Experimental Botany* **60**, 2481-2492.

Coppin A, Varre JS, Lienard L, Dauvillee D, Guerardel Y, Soyer-Gobillard MO, Buleon A, Ball S, Tomavo S. 2005. Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite *Toxoplasma gondii* argues for a red alga ancestry. *Journal of Molecular Evolution* **60**, 257-267.

Dauvillée D, Colleoni C, Shaw E, Mouille G, D'Hulst C, Morell M, Samuel MS, Bouchet B, Gallant DJ, Sinskey A, Ball S. 1999. Novel starch-like polysaccharides are synthesized by a soluble form of granule-bound starch synthase in glycogen accumulating mutants of *Chlamydomonas reinhardtii. Plant Physiology* **119**, 321-330.

Dauvillée D, Kinderf IS, Li Z, Kosar-Hashemi B, Samuel MS, Rampling L, Ball S, Matthew MK. 2005. Role of the E. coli glgX gene in glycogen metabolism. *Journal of Bacteriology* **187**, 1465-1473.

Dauvillée D, Chochois V, Steup M, Haebel S, Eckermann N, Ritte G, Ral JP, Colleoni C, Hicks G, Wattebled F, Deschamps P, d'Hulst C, Liénard L, Buléon A , Puteaux JL, Ball S. 2006. Plastidial phosphorylase is required for normal starch granule biogenesis in the monocellular alga *Chlamydomonas reinhardtii.The Plant Journal* **48**, 274-285.

Dauvillée D, Deschamps P, Ral JP, Plancke C, Putaux JL, Devassine J, Durand-Terrasson A, Devin A, Ball SG. 2009. Genetic dissection of floridean starch synthesis in the cytosol of the model dinoflagellate *Crypthecodinium cohnii*. *Proceedings of the National Academy of Sciences* **106**, 21126-21130

De Koning AP, Keeling PJ. 2006. The complete plastid genome sequence of the parasitic green alga *Helicosporidium* sp. is highly reduced and structured. *BMC Biology* **4**, 12

Derelle E, Ferraz C, Rombault S et al. 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences* **103**, 11647-11652.

Deschamps P, Haferkamp I, Dauvillee D, Haebel S, Steup M, Buleon A, Putaux JL, Colleoni C, d'Hulst C, Plancke C, Gould S, Maier U, Neuhaus HE, Ball S. 2006. Nature of the periplastidial pathway of starch synthesis in the cryptophyte *Guillardia theta*. Eukaryotic Cell **5**, 954-963.

Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux JL, Buléon A, Haebel S, Ritte G, Steup M, Falcon LI, Moreira D, Loeffelhardt W, Nirmal Raj J, Plancke C, d'Hulst C, Dauvillée D, Ball S. 2008*a*. Metabolic symbiosis and the birth of the plant kingdom. *Molecular Biology and Evolution* **25**, 536-548.

Deschamps P, Moreau H, Worden AZ, Dauvillée D, Ball SG. 2008*b*. Early gene duplication within Chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. *Genetics* **178**, 2373-2387.

Deschamps P, Haferkamp I, d'Hulst C, Neuhaus E, Ball S. 2008*c*. The relocation of starch metabolism to chloroplasts: when, why and how. Trends in Plant Science **13**, 1802-1816.

Deschamps P, Guillebeault D, Devassine J, Dauvillée D, Haebel S, Steup M, Buléon A, Putaux JL, Slomianny MC, Colleoni C, Devin A, Plancke C, Tomavo S, Derelle E, Moreau H, Ball S. 2008*d*. The heterotrophic dinoflagellate *Crypthecodinium cohnii* defines a model genetic system to investigate cytoplasmic starch synthesis. *Eukaryotic Cell* **7**, 247-257.

Edner C, Li J, Albrecht T, Mahlow S, Hejazi M, Hussain H, Kaplan F, Guy C, Smith SM, Steup M, Ritte G. 2007. Glucan, water dikinase activity stimulates breakdown of starch granules by plastidial beta-amylases. *Plant Physiology* **145**, 17-28.

Eichinger L, Pachebat JA, Glöckner G et al. 2005. The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435**, 43–57.

Eisen et al. 2006. Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. PLoS Biology **4**, 1620-1642.

Falcon L, Lindwall I, Bauer K, Bergman B, Carpenter E J. 2004. Ultrastructure of unicellular N-2 fixing cyanobacteria from the tropical North Atlantic and subtropical North Pacific Oceans. *Journal of Phycology* **40**, 1074-1078.

Fettke J, Hejazi M, Smirnova J, Höchel E, Stage M, Steup M. 2009. Eukaryotic starch degradation: integration of plastidial and cytosolic pathways. *Journal of Experimental Botany* **60**, 2907-2922.

Gentry MS, Dowen RH, Worby CA, Mattoo S, Ecker JR, Dixon JE. 2007. The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. *The Journal of Cell Biology* **178**, 477–488.

Gil R, Latorre A, Moya A. 2004. Bacterial endosymbionts of insects: insights from comparative genomics. *Environmental Microbiology* **6**, 1109-1122.

Hamana K, Aizaki T, Arai E, Saito A, Uchikata K, Ohnishi H. 2004. Distribution of norspermidine as a cellular polyamine within microgreen algae including the non photosynthetic achlorophyllous Polytoma, Polytomella, Prototheca, and Helicosporidium. *The Journal of General and Applied Microbiology* **50**, 289-295.

Handford M, Rodriguez-Furlán C, Orellana A. 2006. Nucleotide-sugar transporters: structure, function and roles *in vivo*. *Brazilian Journal of Medical and Biological Research* **39**, 1149-1158.

Hennen-Bierwagen TA, Liu F, Marsh RS, Kim S, Gan Q, Tetlow IJ, Emes MJ, James MG, Myers AM. 2008. Starch biosynthetic enzymes from developing *Zea mays* endosperm associate in multisubunit complexes. *Plant Physiology* **146**, 1892–1908.

Henrissat B, Deleury E, Coutinho PM. 2002. Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? *Trends in Genetics* **18**, 437-440.

Honda D, Yokota A, Sugiyama J. 1999. Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine Synechococcus strains. *Journal of Molecular Evolution* 48, 723–739.

Horn M, Collingro A, Schmitz-Esser S, Beier CL, Purkhold U, Fartmann B, Brandt P, Nyakatura GJ, Droege M, Frishman D, Rattei T, Mewes HW, Wagner M. 2004. Illuminating the evolutionary history of chlamydiae. *Science* **304**, 728-730.

Horn M. 2008. *Chlamydiae* as symbionts in eukaryotes. *Annual Review of Microbiology* **62**, 113-131.

Huang J, Gogarten JP. 2007. Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biology* **8**, R99.

James MG, Robertson DS, and Myers AM. 1995. Characterization of the maize gene sugary1, a determinant of starch composition in kernels. *Plant Cell* **7**, 417-429.

Keeling PJ. 2009. Chomalveolates and the evolution of plastids by secondary endosymbiosis. *Journal of Eukaryotic Microbiology* **56**, 1-8

Kotting, O., Santelia, D., Edner, C., Eicke, S., Marthaler, T., Gentry, M. S., Comparot-Moss, S., Chen, J., Smith, A. M., Steup, M., Ritte, G., Zeeman, S. C. 2009. STARCH-EXCESS4 Is a Laforin-Like Phosphoglucan Phosphatase Required for Starch Degradation in Arabidopsis thaliana. *Plant Cell* 21, 334-346.

Krebs EG. 1983. Historical Perspectives on Protein Phosphorylation and a Classification System for Protein Kinases. *Philosophical Transactions of the Royal Society B* **302**, 3-11.

Linka N, Hurka H, Lang BF, Burger G, Winkler HH, Stamme C, Urbany C, Seil I, Kusch J, Neuhaus HE. 2003. Phylogenetic relationship of nonmitochondrial nucleotide transport proteins in bacteria and eukaryotes. *Gene* **306**, 27–35.

Loddenkötter B, Kammerer B, Fischer K, Flügge UI. 1993. Expression of the functional mature chloroplast triose phosphate translocator in yeast internal membranes and purification of the histidine-tagged protein by a single metal-affinity chromatography step. *Proceedings of the National Academy of Sciences* **90**, 2155-2159.

Loftus B, Anderson I, Davies R et al. 2005. The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433, 865–868.

Lu, Y. and Sharkey, T.D. 2006. The importance of maltose in transitory starch breakdown. *Plant Cell and Environment* **29**, 353-366.

Martin W. 2010. Evolutinary origins of metabolic compartmentalization in eukaryotes. *Philosophical Transactions of the Royal Society B* **365**, 847-855.

Martinez-Duncker I, Mollicone R, Codogno P, Oriol R. 2003. The nucleotide-sugar transporter family: a phylogenetic approach. *Biochimie 85*, 245-260.

Matsuzaki M, Misumi O, Shin-i T, Maruyama S, Takahara M, Miyagishima S, Mori T, Nishida K, Yagisawa F, Nishida K, et al. 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. *Nature* **428**, 653–657.

McFadden GI, van Dooren GG. 2004. Evolution: red algal genome affirms a common origin of all plastids. *Current Biology* **14**, 514–516.

Miao X, Wu Q, Wu G and Zhao N. 2003. Changes in photosynthesis and pigmentation in an agp deletion mutant of the cyanobacterium Synechocystis sp. *Biotechnology Letters* **25**, 391-396.

Miao X, Wu Q, Wu G, Zhao N. 2006. Sucrose accumulation in salt-stressed cells of agp gene deletion-mutant in cyanobacterium Synechocystis sp PCC 6803. *FEMS Microbiology Letters* **218**, 71-77.

Michels PAM, Opperdoes FR. 1991. The evolutionary origin of glycosomes. *Parasitology Today* **7**, 105-109.

Morell MK, Myers AM. 2005. Towards the rational design of cereal starches. *Current Opinion in Plant Biology* **8**, 204-210.

Mouille G, Maddelein ML, Libessart N, Talaga P, Decq A, Delrue B, Ball S. 1996. Preamylopectin Processing: A Mandatory Step for Starch Biosynthesis in Plants. *Plant Cell* **8**, 1353-1366.

Moustafa A, Reyes-Prieto A, Bhattacharya D. 2008. Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. PLoS ONE **3**, e2205.

Myers AM, Morell MK, James MG, Ball SG. 2000. Recent progress toward understanding biosynthesis of the amylopectin crystal. *Plant Physiology* 122, 989-998.

Nakamura Y, Kubo A, Shimamune T, Matsuda T, Harada K, Satoh H. 1997. Correlation between activities of starch debranching enzymes and α-polyglucan structure in endosperms of sugary-1 mutants of rice. *The Plant Journal* **12**, 143-153.

Nakamura Y, Takahashi J, Sakurai A, Inaba Y, Suzuki E, Nihei S, Fujiwara S, Tsuzuki M, Miyashita H, Ikemoto H, Kawachi M, Sekiguchi H, Kurano N. 2005. Some cyanobacteria synthesize semi-amylopectin type α-polyglucans instead of glycogen. *Plant Cell Physiology* **46**, 539-545.

Niittylä T, Messerli G, Trevisan M, Chen J, Smith AM, Zeeman SC. 2004. A previously unknown maltose transporter essential for starch degradation in leaves. *Science* **303**, 87–89.

Nowack ECM, Melkonian M, Glöckner G. 2008. Chromatophore genome sequence of Paulinella sheds light on acquisition of photosynthesis by eukaryotes. *Current Biology* 18, 410-418.

Patron NJ, Keeling PK.2005. Common evolutionary origin of starch biosynthetic enzymes in green and red algae. *Journal of Phycology* **41**, 1131-1141.

Plancke C, Colleoni C, Deschamps P, Dauvillée D, Nakamura Y, Haebel S, Ritte G, Steup M, Buléon A, Putaux JL, Dupeyre D, d'Hulst C, Ral JP, Loffelhardt W, Ball SG. 2008. The pathway of starch synthesis in the model glaucophyte *Cyanophora paradoxa*. *Eukaryotic Cell* **7**, 247-257.

Pombert J-F, Keeling PJ. 2010. The Mitochondrial Genome of the Entomoparasitic Green Alga *Helicosporidium*. *PLoS ONE* **5(1)**, e8954.

Preiss J. 1984. Bacterial glycogen synthesis and its regulation. *Annual Review of Microbiology* **38**, 419-458.

Ral JP, Derelle E, Ferraz C, Wattebled F, Farinas B, Corellou F, Buléon A, Slomianny MC, Delvalle D, d'Hulst C, Rombauts S, Moreau H, Ball S. 2004. Starch division and partitioning a mechanism for granule propagation and maintenance in the picophytoplanktonic green alga *Ostreococcus tauri*. *Plant Physiology* **136**, 3333-3340.

Reddy KJ, Benjamin Haskell J, Sherman DM, Sherman LA. 1993. Unicellular, aerobic nitrogen-fixing cyanobacteria of the Genus *Cyanothece*. Journal of Bacteriology **175**, 1284-1292.

Reinhold T, Alawady A, Grimm B, Beran KC, Jahns P, Conrath U, Bauer J, Reiser J, Melzer M, Jeblick W, Neuhaus HE. 2007. Limitation of nocturnal import of ATP into Arabidopsis chloroplasts leads to photooxidative damage. *The Plant Journal* **50**, 293-304.

Reyes-Sosa FM, Molina-Heredia FP, De la Rosa MA. 2010. A novel α-amylase from the cyanobacterium *Nostoc* sp. PCC7119. Applied Microbiology and Biotechnology **86**, 131-141.

Richards T, Cavalier-Smith T. 2005. Myosin domain evolution and the primary divergence of eukaryotes. *Nature* **436**, 1113-1118.

Ritte G, Lloyd JR, Eckermann N, Rottmann A, Kossmann J, Steup M. 2002.The starch-related R1 protein is an α-glucan, water dikinase. *Proceedings of the National Academy of Sciences* **99**, 7166-7171.

Roach P J. 2002. Glycogen and its metabolism. *Current Molecular Medicine* **2**, 101-120.

Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF. 2005. Monophyly of primary photosynthetic eukaryotes: Green plants, red algae, and glaucophytes. *Current Biology* **15**, 1325-1330.

Rollwitz I, Santaella M, Hille D, Flügge U-I, Fischer K. 2006. Characterization of AtNST-KT1, a novel UDP-galactose transporter from *Arabidopsis thaliana*. *FEBS Letters* **580**, 4246-4251.

Shadan S. 2007. Evolutionary genetics: you are what you ate. *Nature* **449**, 155.

Scheidig A, Frèohlich A, Schulze S, Lloyd JR, Kossmann J. 2002. Downregulation of a chloroplast-targeted beta-amylase leads to a starch-excess phenotype in leaves. *The Plant Journal* **30**, 581–591.

Shimonaga T, Fujiwara S, Kaneko M, Izumo A, Nihei S, Francisco BP, Satoh A, Fujita N, Nakamura Y, Tsuzuki M. 2007. Variation in storage alpha-polyglucans of red algae:

amylose and semi-amylopectin types in porphyridium and glycogen type in cyanidium. *Marine Biotechnology* **9**, 192-202.

Shimonaga T, Konishi M, Oyama Y, Fujiwara S, Satoh A, Fujita N, Colleoni C, Puteaux JL, Ball SG, Yokoyama A, Hara Y, Nakamura Y, and Tsuzuki M. 2008. Variations in storage α-polyglucans of the Porphyridiales (Bangiophycidae, Rhodophyta). *Plant Cell Physiology* **49**, 103-116.

Schneegurt MA, Sherman DM, Nayar S, Sherman LA. 1994. Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium Cyanothece sp. strain ATCC 51142. *The Journal of Bacteriology* **176**, 1586–1597.

Schneegurt MA, Sherman DM, Sherman LA. 1997. Composition of carbohydrate granules of the cyanobacterium, Cyanothece sp. Strain ATCC51142. *Archives of Microbiology* **167**, 89-98.

Shearer J, Graham TE. 2002. New perspectives on the storage and organization of muscle glycogen. *Journal of Applied Physiology* **27**, 179-203.

Song J, Xu Q, Olsen R, Loomis WF, Shaulsky G, Kuspa A, Sucgang R. 2005. Comparing the Dictyostelium and Entamoeba genomes reveals an ancient split in the Conosa lineage. *PLoS Computational Biology* **1**, 579-584.

Stechmann A, Cavalier-Smith T. 2003. The root of the eukaryote tree pinpointed. *Current Biology* **13**, R665-R666.

Stegemann S, Hartmann S, Ruf S, Bock R. 2003. High-frequency gene transfer from the chloroplast genome to the nucleus. *Proceedings of the National Academy of Sciences* **100**, 8828–8833.

Suzuki E, Umeda K, Nihei S, Moriya K, Ohkawa H, Fujiwara S, Tsuzuki M, Nakamura Y. 2007. Role of GlgX protein in glycogen metabolism of the cyanobacterium, Synechococcus elongatus PCC 7942. *Biochimica et Biophysica Acta* **1770**, 763-773.

Tagliabracci VS, Girard JM, Segvich D, Meyer C, Turnbull J, Zhao X, Minassian BA, Depaoli-Roach AA, Roach PJ. 2008. Abnormal metabolism of glycogen phosphate as a cause for lafora disease. *Journal of Biological Chemistry* **283**, 33816–33825.

Teste MA, Enjalbert B, Parrou JL, Francois JM. 2000. The *Saccharomyces cerevisiae YPR184w* gene encodes the glycogen debranching enzyme. *FEMS Microbiology Letters* **193**, 105-110.

Tetlow IJ, Morell MK, Emes MJ. 2004*a*. Recent developments in understanding the regulation of starch metabolism in higher plants. *Journal of Experimental Botany* **55**, 2131–2145.

Tetlow IJ, Wait R, Lu ZX, Akkasaeng R, Bowsher CG, Esposito S, Kosar-Hashemi B, Morell MK, Emes MJ. 2004*b*. Protein phosphorylation in amyloplasts regulates starch branching enzyme activity and protein–protein interactions. *The Plant Cell* 16, 694–708.

Thorsness PE, Fox TD. 1990. Escape of DNA from mitochondria to the nucleus in *Saccharomyces cerevisiae. Nature* **346**, 376-379.

Tomlinson K, Denyer K. 2003. Starch synthesis in cereal grains. *Advances in Botanical Research* **40**, 1-61.

Trentmann O, Horn M, van Scheltinga ACT, Neuhaus HE, Haferkamp I. 2007. Enlightening energy parasitism by analysis of an ATP/ADP transporter from chlamydiae. *PLos Biology* **5**, e231.

Ugalde JE, Parodi AJ, Ugalde RA. 2003. *De novo* synthesis of bacterial glycogen: *Agrobacterium tumefaciens* glycogen synthase is involved in glucan initiation and elongation *Proceedings of the National Academy of Sciences* **100**, 10659–10663.

Viola R, Nyvall P, Pedersen M . 2001. The unique features of starch metabolism in red algae. *Proceedings of the Royal Society B: Biological Sciences* **268**, 1417-1422.

Wang Z, Wilson WA, Fujino MA, Roach PJ. 2001. Antagonistic Controls of Autophagy and Glycogen Accumulation by Snf1p, the Yeast Homolog of AMP-Activated Protein Kinase, and the Cyclin-Dependent Kinase Pho85p. *Molecular and Cellular Biology* **21**, 5742-5752.

Wattebled F, Dong Y, Dumez S, Delvallé D, Planchot V, Berbezy P, Vyas D, Colonna P, Chatterjee M, Ball S, d'Hulst C. 2005. Mutants of Arabidopsis lacking a chloroplastic isoamylase accumulate phytoglycogen and an abnormal form of amylopectin. *Plant Physiology* **138**, 184–195.

Weber APM, Linka M, Bhattacharya D. 2006. Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryotic Cell* **5**, 609–612**.**

Wing-Ming C, Hsueh-Mei C, Hso-Freng Y, Jei-Fu S, Tan-Chi H. 1994. The aerobic Nitrogen-fixing Synechococcus RF-1 containing uncommon polyglucan granules and multiple forms of a-amylase. *Current Microbiology* **29**, 201-205.

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol **21,** 809–818

Zabawinski C, Van Den Koornhuyse N, D'Hulst C, Schlichting R, Giersch C, Delrue B, Lacroix JM, Preiss J, Ball S. 2001. Starchless mutants of *Chlamydomonas reinhardtii* lack the small subunit of an heterotetrameric ADP-glucose pyrophosphorylase. *The Journal of Bacteriology* **183**, 1069-1077.

Zeeman SC, Umemoto T, Lue W-L, Au-Yeung P, Martin C, Smith AM, Chen J. 1998. A mutant of Arabidopsis lacking a chloroplastic isoamylase accumulates both starch and phytoglycogen. *Plant Cell* **10**, 1699-1712.

Zeeman SC, Smith SM, Smith AM. 2007. The diurnal metabolism of leaf starch. *Biochemical Journal* **401**, 13-28.

Materials and Methods

Materials Required

1. Strain and culture conditions

The axenic nitrogen-fixing cyanobacteria, strain Clg1 was recently isolated from the Sub-tropical North Atlantic (Falcón et al., 2004) and axenized through repeated cloning on solid ASNIII medium (0.9% agar), Then the strain was cultured in ASNIII medium (See Table 1 & 2) (Rippka et al., 1979) at 25°C at a day/night cycle (12H/12H) under a photosynthetic photon flux density of approximately 30-µmol photons.m^{-1}.s^{-1} for developing more cultures.

Ingredients	g L^{-1}	mM
NaCl	25.0	427
$MgCl_2.6H_2O$	2.0	9.8
KCl	0.5	6.7
$NaNO_3$	0.075	0.88
$K_2HPO_4.3H_2O$	0.02	0.09
$MgSO_4.7H_2O$	3.5	14.2
$CaCl_2.2H_2O$	0.5	3.4
Citric acid	0.003	0.015
Ferric ammonium citrate	0.003	0.015
EDTA (disodium magnesium)	0.0005	0.0015
Na_2CO_3	0.02	0.19
Trace metal mix A5+Co	1 ml	
Deionized water	Up to 1 l	
pH after autoclaving and cooling	7.5	

Table 1: Media ASN – III Composition

Ingredient	g l^{-1}
H_3BO_3	2.86
$MnCl_2.4H_2O$	1.81
$ZnSO_4.7H_2O$	0.222
$Na_2MoO_4.2H_2O$	0.390
$CuSO_4.5H_2O$	0.079
$Co(NO_3)_2.6H_2O$	0.049

Table 2: Trace Metal Mix A5 + Co Composition

2. Mutagenesis Treatments

2.1. Ultraviolet Irradiation

Ultraviolet mutagenesis irradiation was performed at room temperature on cell concentrations of different dilutions 0, 1/5 and 1/10. This process was performed with a standard Transilluminator TS-15 (254 nm, Ultra-violet Products, Inc., San Gabriel, Calif.) with a peak intensity of 7.0mW.$cm^{-2.}$ Exposures were made at three different heights of 0.5 cm, 3cm and 6cm at time intervals of 30", 1' and 2' (See Table 3).

Treatment of the culture suspension with UV irradiation was carried out in complete darkness in a 90mm sterile Petri-dishes with an open lid, each containing 1ml of homogenous cell suspension and held under UV irradiation to facilitate uniform exposure. After the exposure the plates were left in the dark for 48 hours to avoid any photo damage reparation. The cells thus obtained were gently scratched by adding ASN III over the agar slants and then transferred to a liquid culture. Expression and segregation of the induced mutations was achieved by allowing the cells to grow for 4 to 10 generations. The grown cells were then

diluted and sprayed in agar plates to get an individual colony, which increased the possibility of mutation frequency.

Exposure Height	Dilution	Exposure Time		
.5 cm	0	30”	1’	2’
	1/5			
	1/10			
3 cm	0	30”	1’	2’
	1/5			
	1/10			
6 cm	0	30”	1’	2’
	1/5			
	1/10			

Table 3: Ultraviolet Mutagenesis Table3

3. Preparation of Agar Plates

Solid media was prepared using Difco Bacto-Agar that was further purified using the following protocol: 100 g of agar was washed by stirring with 3 liters of double-distilled water in a beaker. After 30 minutes of stirring, the agar was allowed to settle, the wash water was siphoned off, and the agar was filtered onto Whatman F4 filter paper in a Buchner funnel. This procedure was repeated until the filtrate was clear. The agar was then washed with 3 liters of 95% ethanol followed by a final 3-liter wash with analytical grade acetone. The agar is then dried at 50°C in glass baking dishes for 2 to 3 days and stored in a tightly covered container. Solid media thus prepared had a final concentration of 0.6% and was sufficiently stable for growing cells and the medium of Agar was poured in the plates with the composition of ASN III.

4. Determination of Survival Percentage

For determining the survival percentage, three different dilutions were withdrawn at known intervals of time after UV irradiation treatment. After a 48-h dark incubation, plates were transferred to light in the culture room. Percentage survival was scored with the help of colony counter and it was plotted semi logarithmically.

5. Screening and Harvesting Mutants

Single colonies were transferred in 96 Well plates. From the same cell suspension and the patches were duplicated and performed in two different plates. One plate was used for screening with iodine vapors. Selection of mutant strain by observing difference in color variations, potential mutants were picked from the master plate and inoculate in fresh ASNIII Medium.

6. Starch Extraction, Purification and Quantification

Cells were harvested after 7–10 days of culture by centrifugation 10 min at 3,000 g at 4 °C in sterile 50ml Falcon tube. The pellets were re-suspended in water and centrifuged again for 15 min at 3,000 g. The pellets were finally resuspended in 10ml of TpA buffer (25 mM Tris/Acetate pH 7.5, DTT 10 mM). Cells were disrupted using a French press (10,000 Psi). Starch and cell debris were collected by centrifugation (10,000 g, 15 min) and re-suspended in extraction buffer (Hepes pH 8 500mM, DTT 5mM, EDTA 1mM, 125µLTriton 100%). Due to density differences, starch can be pelleted by centrifugation (10,000 g, 30 min) away from the bulk of cell debris. The gradient step for starch purification was repeated once to ensure a complete removal of cell debris from the purified starch granules then it was stored at 4°C. Starch amount were assayed using Diffchamb Enzyplus Starch kit (Diffchamb, Lyon, France).

7. Starch and WSPs Incubation (+/-) Phosphate

Incubation experiments were carried out by incubating 2mg of starch at different temperatures (0°C, 30°C and 40°C) and in the presence or in the absence of inorganic phosphate (25mM) under continuous mixing (**~1400 rpm**). After overnight incubation, insoluble and soluble materials were separated by centrifugation (10 min at 13.2 rpm). Insoluble material and WSP were quantified by amyloglucosidase assay and loaded on size exclusion chromatography CL-2B and TSKHW 50, respectively. The quantity of equivalent glucose and the absorbency of iodine-polysaccharide complex were determined for each fraction. G-1-P assay were separated by centrifugation like Water soluble polysaccharide and enzymatic reaction stopped by boiling supernatant then the samples were transformed to 10µg of phosphoglucomutase and latter 2µg of glucose-6-phosphate dehydrogenase. Thus the quantity of G-1-P was observed by reaction coupled to NAPD reduction.

8. Production and Preparation of Crude extract

Crude extract has to be collected at the time of harvesting of starch while after using French press (10,000 Psi). The cells were disrupted and the supernatant were collected by centrifugation (10,000 x g, 15 min), remaining starch and cell debris was used for assay. This crude extract was kept under -80°C for enzymological studies. The determination of continuous protein in the supernatant was realized by using Bio-Rad protein Kit.

9. Fracinations of Starch Though Size Exclusion Chromatography

Size exclusion chromatography (SEC) is a chromatographic method in which molecules in solution were separated depending on their size (more correctly, their hydrodynamic volume).

9.1. CL-2B

2 to 4 mg of Starch was dispersed in 100 µL of DMSO (100%). It was then agitated with a micropipette and left at 100°C for about 10min. It was verified if the resultant sample was clear. 1000 µL of ethanol (100%) was added and allowed the polysaccharides to precipitate using absolute ethanol. The sample was then stored at -20°C for about 30 min. It was then centrifuged (4000 g) for 5 min. The supernatant was removed and the tube was dried upside down over a Whatman F4 filter paper, then the tube was suspended by 500 µL of NaOH 10mM and subsequently loaded the sample in CL-2B (0.5 cm X 65 cm) column containing sepharose gel and the elution was equally supported by buffer NaOH 10mM. Fractions were eluted and collected in the tubes by 300 ml per 90 seconds of each fraction. The fractions containing glucans were mixed with an Iodine solution (I2 0.1%, KI 1%) and was measured for the absorbance of each fraction with spectrophotometer to find the λ_{max} and the wavelength of the maximum absorbance of the iodine-polysaccharide complex (Delrue et al., 1992). Amylopectin and amylase were separated through a CL2B gel permeation chromatography (Pharmacia Biotech, Piscataway, NJ) equilibrated in 10mM Na0H as described in (Delrue et al., 1992).

9.2. TSK-HW-50[F]

Water soluble polysaccharide (WSP) was dispersed in 100 µL of DMSO (100%). The sample was agitated with a micro pipette and was kept at 100°C for 10min. After a while, it was verified if it was clear. Once the sample was clear, it was loaded in the TSK-HW-50 [F], Merck, Darmstadt, Germany (0.8 cm X 65 cm) column containing sephadex gel eluted in 10% (W/V) DMSO (dimethyl sulfoxide). A detailed description of the method could be found in (Maddelein et al., 1994). The method for WSP fraction purification could be found in (Dauvillée et al., 1999). Elutions were equally supported by buffer DMSO 10%. Fractions

were eluted and collected in the tubes. The fractions contained Oligosaccharides and WSP were assayed with Phenol-Sulphuric acid to find the total amount of sugar present in the sample and for controlling 10 µL glucose solution 1 mg/ml + 10 µL H2O and NaOH 10mM used and absorbance of each fraction measured using micro titration plate with the spectrophotometer in 490nM.

10. Materials

Glucose-1-phosphate (glucose-1-P) and ADP-glucose was purchased from Amersham (Little Chalfort, Buckinghamshire, UK). Rabbit muscle phosphorylase a, glucose-6-phosphate (glucose-6-P) dehydrogenase, rabbit-muscle phosphoglucomutase, and the starch-determination kit were purchased from Boehringer (Mannheim, FRG). Rabbit-liver glycogen was supplied by Sigma Chemical (St. Louis, Mo., USA).

11. Zymogram

Zymography is an electrophoretic technique which separates protein according to their size and charge properties. Proteins that are highly charged at the pH of the separating gel migrate faster than those with less charged molecules. Many enzymes still have biological activities after running PAGE. These include, for example, esterase, dehydrogenase alkaline phosphatase, α-amylase, transferase, hydrolases, lyases, β-fructosidase (invertase), and isomerases. We utilize zymogram technique well develop and adapted method by (Lacks and springhorn., 1980) Who used an enzymatic model as α-amylase by studying their renaturation techniques and detection can be found in (Gabriel and gerstern., 1992). Crude Extracts are harvested at the middle of their exponential phase growth 10 to 14 days of culture. Cultivation was carried out (every 3 hours) under different conditions of light (i.e) day, night and continuous light. Crude extract was collected by pressing cells under French press (10,000

Psi). The cells had disrupted and the supernatant was collected by centrifugation (10,000 x g, 15 min at 4°C). This crude extract was kept under -80°C for enzymological studies. The quantity of continuous protein in the supernatant was realized using Bio-Rad protein Kit. Preparation of buffer for migration: For 1 lit of Tris glycine buffer (100 ml of Tris Glycine 10 X), 900 ml of water, and 1mM of DTT). Different fraction of protein samples was prepared with 10 µL glycerol 60% and 40 µL bromophenol blue solution. Assemble the vertical slab gel unit, Such as the SE 600 vertical slab unit (Hoeffer).

Protein Gel composition – Zymogram

	For 10 mL	2 gels at 7 mL
Acrylamide/bisacrylamide 40%	1.875 mL	2.625 mL
1.5 M Tris/HCl pH=8.8	2.5 mL	3.75 mL
H_2O or Glycogen 1%	5.67 mL	8.5 mL
APS 10%	50 µL	75 µL
TEMED	5 µL	8 µL

Table 4: Separation Gel 7.5%

A/B	1.0 mL
0.5 M Tris/HCl pH 6.8	2.5 mL
H_2O	6.450 mL
APS	50 µL
TEMED	10 µL

Table 5: Concentration or Stalking Gel 4 % (for 2 gels)

	For 10 mL (1 gel)	For 20 mL (2 gels)
Acrylamide/bisacrylamide 99%	1.875 mL	3.75 mL
1.5 M Tris/HCl pH=8.8	2.5 mL	5 mL
Starch, Glycogen or β-limit dextrin 1 %	5.67 mL	11.34 mL
APS 10%	50 µL	100 µL
TEMED	5 µL	10 µL

Table 6: Starch, Glycogen or β-limit dextrin Gel

11.1 Electrophoresis

After polymerization of gels, the electrophoresis apparatus was setup (Mini protean II Bio-Rad) Richmond, CA, USA. Samples were loaded into each well with quantity less than 40µL, it was at 4°C at 15 mA (constant) for 3 hrs. After the migration, the gel was let to incubate ovenight with different buffer mix and stain with Iodine solution (0.25% KI, 0.025% I_2). The zymogram was immediately photographed. The molecular mass of each enzyme activity detected on zymogram and measured on the same polyacrylamide gel. Branching enzyme and D-enzyme activies were detected on native glycogen containing Zymograms.

Proteins were subjected to separation on a native 7.5% acrylamide gel containing 1% of rabbit liver glycogen (Sigma) and stacking gel contained 2.5 % acrylamide (See Table 4 and Table 5) then electrophoresis was performed which incubate gel overnight in the buffer branching enzyme (G-1-P 40mM, AMP 2mM, Tampon A 20ml, Phosphorylase a 0.5 mg/ml) and Buffer D-enzyme (50 mM Tris, 5 mm EDTA, 10mM DTT, and 2mM maltoheptaose) at 25°C and revealed the activities by staining the gels.

11.2. Electroblotting

Proteins were separated by electrophoresis on native gel and then transferred onto another PAGE containing 0.3% of polysaccharide (i.e. glycogen) (See figure: 1) for 1 hr 20V and 110 mA and incubate overnight in the phosphorylase mix (G-1-P 40mM and Tris acetate DTT 10mM) for detect phosphorylase activity, Starch synthase mix (glycyl glycine 1M pH 9, NH_4S0_4 2M, $MgCL_2$ 1.2M, BSA 10µg/ml, β–mercaptoethanol, ADP glucose 360mM).

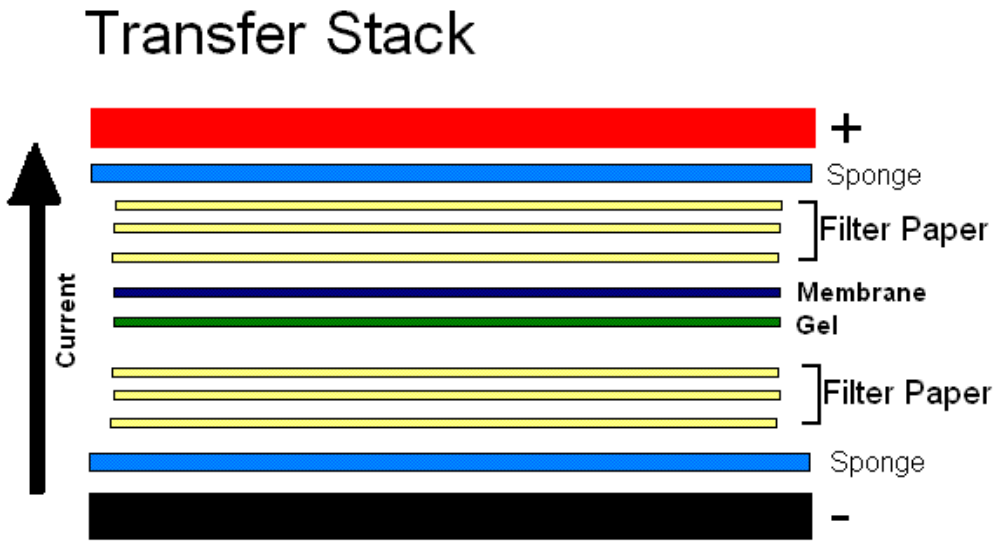

Figure 1: Sandwich preparation for transfer proteins onto a membrane gel by Electro blotting technique.

Moreover, the setup was run overnight at 20V and 110mA with β-limit dextrin or starch gel and staining it with iodine directly after 12 hr of electroblotting enzyme transport.

12.ADP Glucose Pyrophosphorylase Assay

Assay conditions were chosen so that nearly saturating concentrations of the substrates were present in the absence and in the presence of the activator 3-PGA and Pi (pH 7.5) as inhibitor. 50µL of Crude extract (10^6 10^7 cells) (Lowest conc. Samples – 2.5 µg/ml present in 200µL) were suspended in 0.2 ml of 80 mM glycylglycine-NaOH, (pH 7.5), 60mM $Mgcl_2$, 10mM

NaF, 1 mM sodium pyrophosphate, 0.05mM glucose-I,6-biphosphate, 0.25mM NADP, 1 mM ADP-glucose in the presence or absence of 5 mM 3-PGA and 0.0625mM of Pi. After 30 min incubation at 37°C, the reaction mixture was stopped by boiling for 10 min in a water bath and cleared by a 10-min spin at 10000.g.

Glucose-1-phosphate produced by the reverse reaction catalyzed by ADP-glucose pyrophosphorylase was converted to glucose-6-phosphate by the addition of 10µg of phosphoglucomutase. After 5-min incubation at room temperature, glucose-6-phosphate levels were measured by adding 2µg of glucose-6-phosphate dehydrogenase. Reduction of NADP to NADPH was monitored by reading the absorbance at 365nm using an extinction coefficient of 3.5 (1 $nmol^{-1}.cm^{-1}$). The data are presented as averages of three independent assays with one standard deviation.

Results

I

Diurnal rhythm of starch metabolism In Clg1 strain

Circadian control of starch metabolism in Clg1 strain

Circadian (Latin *circa* "about" *diem* "one day") clocks have been found in a wide spectrum of organisms from cyanobacteria to eukaryotic cells. The circadian rhythm responds to an "internal clock" which allows organisms to adapt efficiently biological processes to day-night cycles. For instance, most photosynthetic genes are under circadian control in photosynthetic organisms. In cyanobacteria, the circadian clock controls the expression of genes responsible for photosynthesis and nitrogen fixation. Nitrogen fixation occurs through the activity of nitrogenase, which requires a large amount of ATP and the absence of oxygen. The filamentous diazotrophic cyanobacteria have specialized nonphotosynthetic cells called heterocysts, in which nitrogen fixation activity takes place. Heterocyst cells are characterized by a thick cell wall and by the absence of the oxygen evolving PSII photosynthetic complex. This spatial separation of both processes allows nitrogen fixation to occur during the light phase. Nitrogen fixing unicellular cyanobacteria have solved this dilemma by performing a temporal separation of both mutually exclusive processes. Thus, photosynthesis occurs during the light period and nitrogen fixation takes place at night when the oxygen level is low within the cell (Reddy et al., 1992; Mohr et al., 2009; Compaoré et al., 2010). In this process, the energy is temporally stored in a form of carbohydrate granules during the day in order to fuel the nitrogenase activity at night. Interestingly, the opposite oscillation of nitrogenase activity and carbohydrate content is maintained when unicellular diazotroph Cyanothece ATCC51142 are placed in continuous light (Schneegurt et al. 1994). Hence, this result suggests a strong coordination of gene expressions involved in nitrogen and carbon metabolisms by the circadian clock. Because Clg1 strain has been initially described as a unicellular diazotrophic cyanobacterium, we examine if the starch metabolism is under circadian clock control. The carbohydrate content and starch metabolism enzymes were analyzed from Clg1 cultures synchronized in day/ night cycles (12/12) and then placed in continuous light for 72 hours.

Circadian control of starch content in Clg1 strain

In contrast to a simple diurnal control, a biological process following a circadian rhythm persists under constant conditions with a period of about 12h/12h for several cycles. Clg1 cultures were synchronized for 14 days under a 12-light/12-h dark cycle in ASNIII medium. Two flasks of Clg1 cells were harvested every 3 hours. Cell layers at the bottom of flask were re-suspended and disrupted by sonication. The carbohydrate granules and water-soluble polysaccharide (WSP) material were separated by centrifugation. Both amounts of carbohydrate granules in the pellet and WSP in the supernatant were measured by the amyloglucosidase assay. Free glucose in the supernatant was estimated without adding amyloglucosidase. Results are expressed as milligram of glucose equivalent per milligram of protein. Polysaccharides and free glucose were measured on synchronized cells under 12h-day/12h-dark (figures 1A; 1C and 1E) and after 72 hours in constant light during 72 hours (figure 1B; 1D and 1F).

In day-night cycles, Clg1 cells display a diurnal pattern of carbohydrate granules, WSP and free glucose contents (figures 1A, 1C and 1E). The starch granules are synthesized until the onset of dark and are mobilized quickly in the dark. More than 50 % of starch amount is degraded in three hours. On the contrary, the production of WSP increases slowly to reach a peak during the middle of the night (figure 1C). In addition, the concentration of free glucose is stable (0.5 mg of glucose/mg of total protein) during the 12 hours of light period and increases at the night transition to reach 0.7 mg of glucose/mg of total protein before decreasing up to 0.6 mg of glucose/mg of protein during the rest of night.

Diurnal patterns persist with the same amplitudes for both starch and WSP content in continuous light after 72 hours. Nevertheless, both peaks of starch and WSP contents shifted of 6 and 12 hours respectively (figures 1B and 1D). Only three hours separate both starch and WSP peaks instead of 12 hours in day-night cycles. Although diurnal pattern of free glucose

does not clearly persist, there seems to be a boost of free glucose at the supposed subjective day-night transition (Figure 1F).

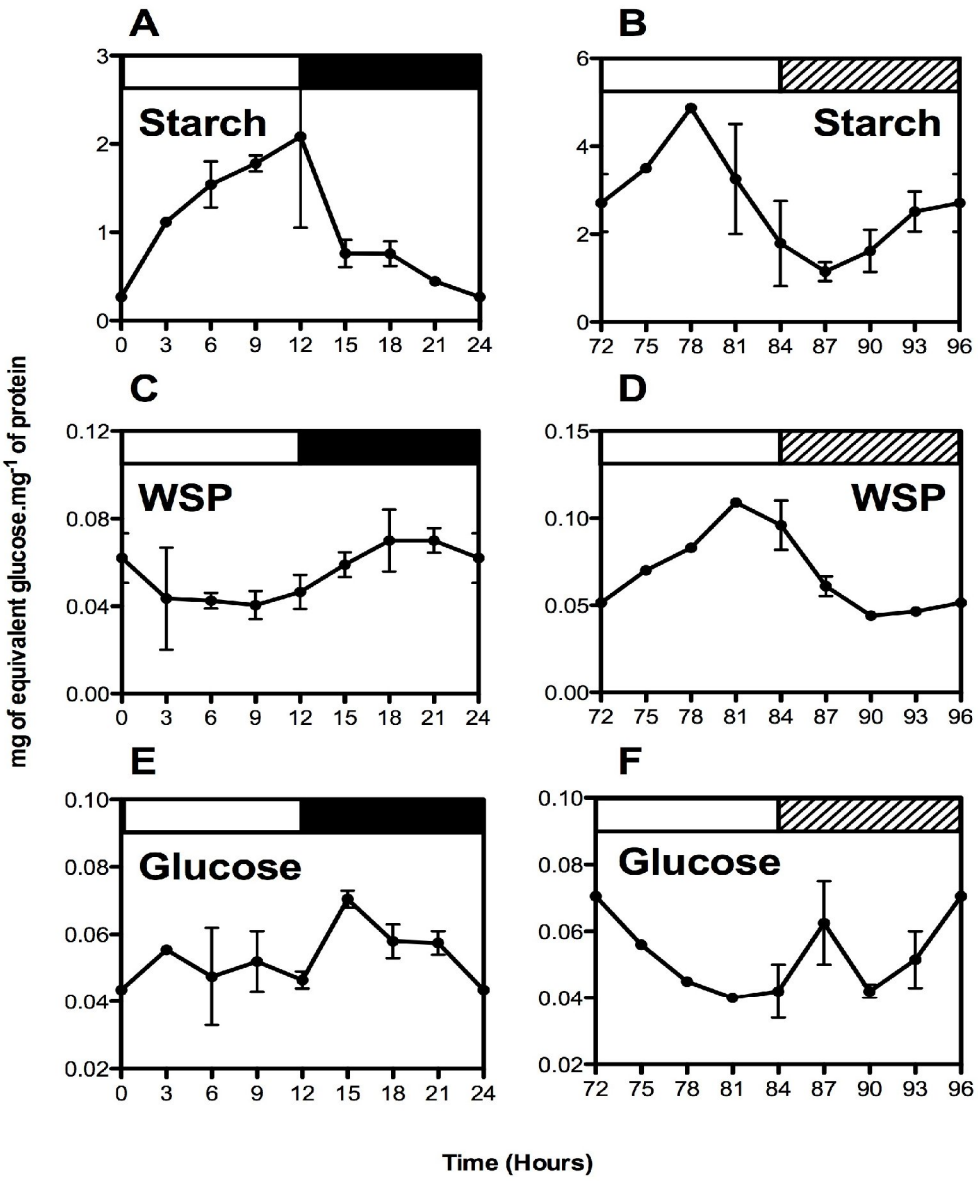

Figure 1: Oscillation of starch and water-soluble polysaccharide content in Clg1 cells placed under a 12Hours-light/12Hours-night cycle and under continuous light. Cells were synchronized for 14 days under a 12Hours-light/12Hours-night cycle and then, placed under continuous illumination. After disrupting cells by sonication, the carbohydrate granules and water-soluble polysaccharide (WSP) were separated by centrifugation. The amount of carbohydrate granules (pellet) and WSP (supernatant) were determined by the amyloglucosidase assay. The graphs show the mean of triplicate assays upon two independent cultures. Data are expressed as mg of equivalent glucose per milligram of protein. The oscillation of polysaccharide content was determined after three days of continuous illumination. The bar at the top of the graph indicates the corresponding day (white box)/night (black box) or subjective night (zebra box).

Interestingly, if we assume that WSP is produced during the starch degradation, these continuous light experiments suggest that gene expression of enzymes involved in both catabolism and anabolism pathways are expressed almost at the same moment. To investigate the nature of the major steps responsible for the oscillation of polysaccharide content in Clg1 cultures, we assayed ADP-glucose pyrophosphorylase activity, a key enzyme bacterial and plant starch synthesis. This enzyme controls the synthesis of ADP-glucose, the nucleotide-sugar used as substrate by starch synthases activities. In plants and cyanobacteria, ADP-glucose pyrophosphorylase is tightly regulated by two effectors: inorganic phosphate as inhibitor and 3-PGA as activator (Frueauf et al., 2002; Gosh and Preiss, 1966). As reported in *Chlamydomonas reinhardtii*, the oscillation of ADP-glucose pyrophosphorylase activity is under circadian clock control (Ral et al., 2006).

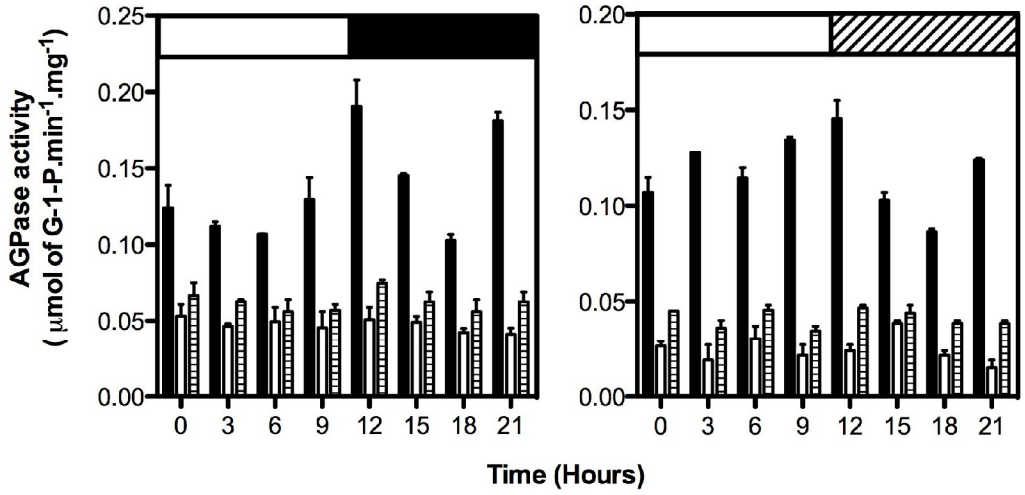

Figure 2: ADP-Glucose pyrophosphorylase (AGPase) activity measured in day-night cycle culture (A) and under continuous light (B). AGPAse activities were measured in the ADP-glucose pyrophosphorolysis direction in the absence (zebra boxes) or in the presence of 3-PGA (black boxes) or inorganic phosphate (white boxes). The results represent the mean of triplicate assays upon two independent experiments. The results are expressed as μmol of G-1-P produced. min^{-1}. $mg^{-1}$of protein. The bar at the top of the graph indicates the corresponding day (white box)/night (black box) or subjective night (zebra box).

ADP-glucose pyrophosphorylase activity was measured in the presence or in the absence of both effectors on synchronized Clg1 cells harvested during 12H-day/12h-night cycle and

under continuous light (figure2). In day-night cycle, two peaks of ADP-glucose pyrophosphorylase activities are observed, one peak at the day-night transition and another one three hours before the sun rise. Only the first peak is correlated to the peak of starch accumulation observed on figure 1A. These two peaks appear significantly when the AGPase activity is measured in the presence of 3PGA and in the absence of both effectors. In continuous light, the AGPase activity is higher in all samples harvested during the light period than in those harvested in the subjective night period. Again, a correlation between the level of AGPase activity and the starch granules content is observed in continuous light (figures 1B and 2B). The sensitivity of ADP-glucose pyrophosphorylase to inorganic phosphate inhibition seems more important in continuous illumination samples than in day-night cycle samples. Indeed, the average values of AGPase activities were estimated at 0.041±0.005 and 0.062±0.008 µmol of G-1-P.min^{-1}.mg^{-1} of protein in continuous light and day-night cycle, respectively.

Enzyme	Locus	Amino acid	Number of isoform in *C.watsonii*	Number of isoform observedin Clg1
ADP-glucose pyrophosphorylase	ZP_00518084	429	1	**ND**
Isoamylase	ZP_00514207	705	1	**1**
Branching enzyme	ZP_00518814	773	3	**1**
	ZP_00517916	651		
	ZP_00515366	644		
Starch synthases	ZP_00517910	526	3	**1**
	ZP_00518684	591		
	ZP_00514159	490		
Phosphorylase	ZP_00517515	298	2	**1**
	ZP_00515288	848		
D-enzyme	ZP_00514038	502	1	**1**
Amylases	ZP_00514888	969	3	**1**
	ZP_00516229	453		
	ZP_00515884	486		

Table 1: Number of isoforms predicted according to genome of *Crocosphaera watsonii* and visualized on zymograms in CLg1 strain.

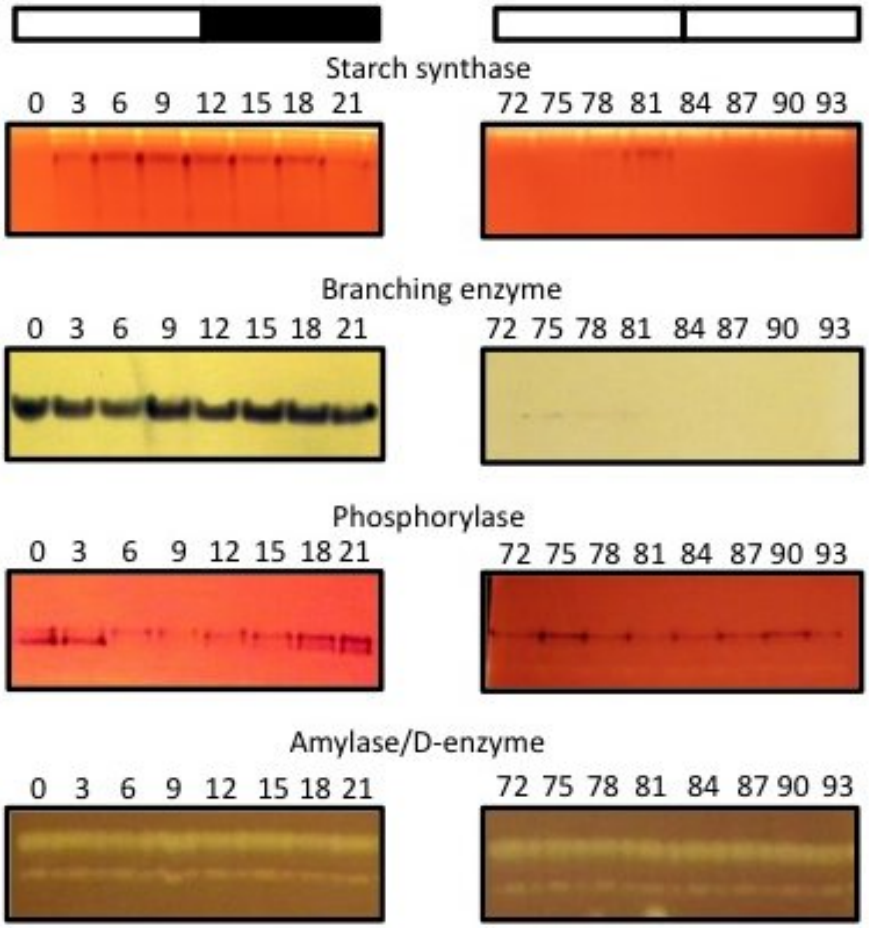

Figure 3: Zymogram analyses of the major starch-metabolizing enzymes from Clg1 cultures grown under day-night cycle and after 72 hours of continuous light. Protein extracts (50 µg of total protein) were loaded and separated on native PAGE. The specific branching enzyme zymograms were incubated overnight in a buffer containing G-1-P and phosphorylase A from rabbit. Phosphorylase, starch synthase, D-enzyme-and amylase activities were revealed after transferring onto polyacrylamide gel containing either 0.3% of glycogen or 0.3% of beta-limit dextrin. After one hour of electroblotting, the glycogen-containing gels were incubated overnight in starch synthase or phosphorylase reaction buffers. D-enzyme and amylase activities were electroblotted onto beta-limit dextrin gel overnight. After incubation, gels were stained with iodine solution.

To further explore the oscillation of starch content, we have followed the other starch metabolism activities e.g. starch synthases, phosphorylase, branching enzymes, D-enzyme and isoamylase-type debranching enzymes by zymogram analyses. Thanks to phylogenetic trees based on 16S RNA sequences and nifH genes, we know that Clg1 and *Crocosphaera watsonii* strains belong to the same group of unicellular diazotroph cyanobacteria (Falcòn et al. 2002). The genome of *Crocosphaera watsonii* having been completely sequenced, we can use it as

model to have an estimate on the possible number of enzymatic isoforms involved in carbohydrate metabolism (table I).

In day-night cycles or continuous light, there is no significant diurnal pattern of starch metabolism enzymes with the noticeable exception of phosphorylase. Indeed, two different phosphorylase isoforms appear during the day-night cycle. The lower phosphorylase isoform is clearly visible in the middle of the night until 6 hours of day and the upper phosphorylase isoform appears alone during the day. In addition, the number of isoforms is reduced for each type of activity. In this case, we expected three branching enzyme activities on a specific branching enzyme zymogram and at least two, starch synthase activities.

Discussion

Both carbohydrate granules and water-soluble polysaccharide contents show an opposite pattern when the Clg1 cells are maintained in day-night cycles. The peak of carbohydrate granules synthesis is reached at the end of the day and the amount of WSP increases until the middle of the night. The synthesis of carbohydrate granules was also correlated to an increasing activity of ADP-glucose pyrophosphorylase. Furthermore, oscillations in carbohydrate content in Clg1 cells persisted under continuous light thereby suggesting a circadian-clock control of starch metabolism. Interestingly, when placed under continuous light, Clg1 cells display both catabolism and anabolism pathways at the same time. If we presume that water-soluble polysaccharides may define a transitory degradation form of carbohydrate granules during the catabolism pathway, this suggests also that the expression of genes involved in the catabolism pathway has shifted of 12 hours. On the other hand these WSP might be byproducts of starch synthesis. No significant diurnal pattern of starch metabolism enzymes is visualized on zymogram gels. One explanation is that the level of gene expression does not fluctuate enough to be visualized on zymogram analysis through a

day-night cycle. Indeed, a recent study on diurnal expression of genes involved in the carbon and nitrogen metabolisms in *Crocosphaera watsonii* show that both starch synthase and phosphorylase genes display more expression in the light or in darkness, respectively (Mohr et al., 2009). However, if we assume a direct correlation between transcript and protein levels, the relative abundance variations of transcripts during a day-night cycle may be not significant enough to be visualized on zymogram. We do not know if the significant variations of the upper and lower phosphorylase isoformsreflects changes in mRNA abundance of the two phosphorylase genes or post-translational modifications of the phosphorylase protein.

II

In vitro incubation of native starch granules

In Vitro Incubation of Starch

In our previous experiments, four major polypeptides (two starch synthases, one phosphorylase and one branching enzyme) associated to carbohydrate granules of Clg1 have been identified after trypsic digestion followed by mass spectrometry analysis. Starch synthase activities were only characterized by following the incorporation of ^{14}C-glucose in to native carbohydrate granules (Deschamps et al., 2008). In order to investigate if the phosphorylase activity is also active on the native carbohydrate granules, we performed in vitro incubation experiments conducted in the presence or absence of inorganic phosphate. Because of the presence of unidentified polypeptides attached to carbohydrate granules, we washed carbohydrate granules with cold detergent (5% SDS at 20°C) (Mu-forster et al., 1998).

Loosely attached polypeptides are removed by this washing step while tightly attached polypeptides were released by boiling carbohydrate granules in SDS buffer. The patterns of tightly or loosely attached polypeptides were appreciated by SDS-PAGE analysis (figure 4). Most of polypeptides associated to carbohydrate granules are found in the upper phase of washing steps (W1 and W1), including the four major polypeptides previously identified.

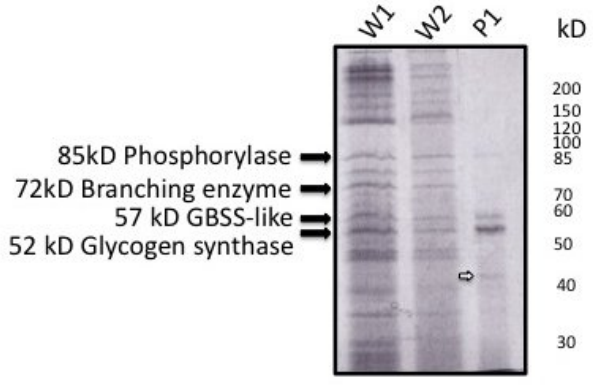

Figure 4: SDS-PAGE analysis of polypeptide associated to carbohydrate granules of Clg1. 10 mg (dry weight) of carbohydrate granules were resuspended in SDS-buffer at 20°C. In order to remove the proteins loosely bound, carbohydrate granules were washed twice in the presence of detergent (10% SDS-buffer). After

recovering the granules by centrifugation, proteins were released from carbohydrate granules by boiling the sample in the loading buffer. After centrifugation, proteins in the supernatant (P1) and in both upper phases of washing step (W1 and W2) were separated on SDS-PAGE (7.5%). Following the electrophoresis, proteins were visualized by Coomassie blue. In the P1 line, the white arrow points an unidentified polypeptide of ~43 kD tightly attached to carbohydrate granules.

The GBSS-like and "soluble" starch synthase polypeptides are clearly tightly attached to carbohydrate granules as well, a weak band corresponding to 85 kD phosphorylase and an unknown polypeptide of ~43 kD (white arrow). In contrast, branching enzyme proteins are not observed in the P1 fraction P1 suggesting that this protein is loosely attached to carbohydrate granules. Carbohydrate granules were collected by centrifugation after disrupting Clg1 cells by French press. The following purification steps were then performed in HEPES buffer containing Triton-X100 (Ritte et al., 2000). The total amount of carbohydrate granules was estimated by the amyloglucosidase assay. Two milligrams of carbohydrate granules were used for each condition. Amounts of carbohydrate granules, WSP and G-1-P were estimated after overnight incubation at 0°C (control), 30°C and 40°C in the absence or in the presence of inorganic phosphate (figure 5).

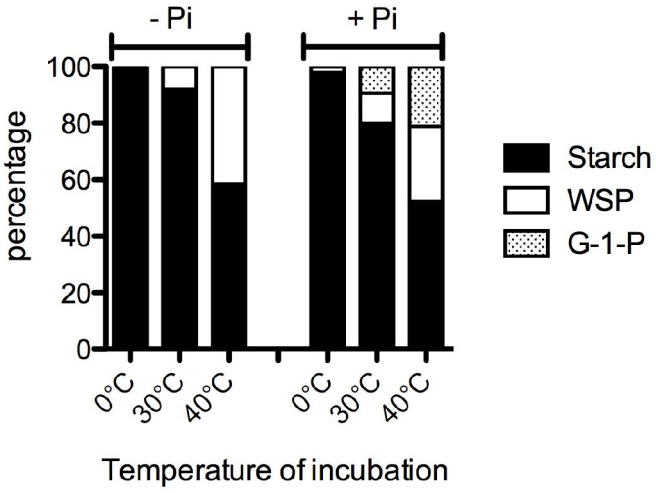

Figure 5: Incubation experiments of purified carbohydrate granules. Carbohydrate granules were purified from Clg1 strain in presence of Triton-X100 buffer. Two milligrams of carbohydrate granules were incubated at 0°C, 30°C and 40°C in the absence or in the presence of 0.0625mM inorganic orthophosphate. After overnight incubation, starch granules and water-soluble polysaccharide were separated by centrifugation. Both polysaccharides were measured by amyloglucosidase assay. The glucose-1-phoshphate (G-1-P) was quantified

by using a mix of phosphoglucomutase-glucose-6-phosphate dehydrogenase enzymes. The sum of carbohydrate granules, WSP and G-1-P was established at 100% for each condition.

Both water-soluble polysaccharide, G-1-P and carbohydrate granules were separated by centrifugation. The enzymatic reactions are stopped by boiling supernatant and the pellet containing starch granules. After cooling, the glucose-1-Phosphate produced is transformed to 6-phospho-gluconate by using a mix of Phosphoglucomutase and glucose-6-phosphate dehydrogenase enzymes coupled to NAPD reduction. Both starch and WSP were quantified by amyloglucosidase assay. At 0°C, G-1-P is not produced and 2% of WSP is observed when the incubation experiment contains 0.0625mM of Pi. As expected, the production of G-1-P is measured only in the samples incubated in presence of Pi.

The amounts of G-1-P increase from 9 to 21% at 30°C and 40°C, respectively. In regards to the results, 85 kD phosphorylase appears to be active on native carbohydrate granules. Interestingly, at 40°C and in the absence of Pi, 40% of the carbohydrate granule seems solubilized. This strongly suggests the presence of starch hydrolase attached to the carbohydrate granules. At 40°C, the undigested starch decreases from 60% to 50% in presence of Pi suggesting that both hydrolytic enzymes, phosphorylase and amylase, act synergistically on the carbohydrate granules catabolism.

Residual carbohydrate granules and water-soluble polysaccharide were subjected to size exclusion chromatography analyses (Figure 6).

In control experiments (O°C +Pi and 0°C -P), both amylopectin and amylose polysaccharides were clearly separated on Cl2B gel filtration. The λmax value for amylopectin (fractions 38 to 60) was estimated at 510nm in both experiments (figures 6A and 6B). At 40°C, the amylose fraction was completely degraded when both phosphorylase and amylase are active (figure 6D). This can be either explained by a higher of amylose resistance to unknown hydrolase or a higher sensitivity to phosphorylase (Figure 6C).

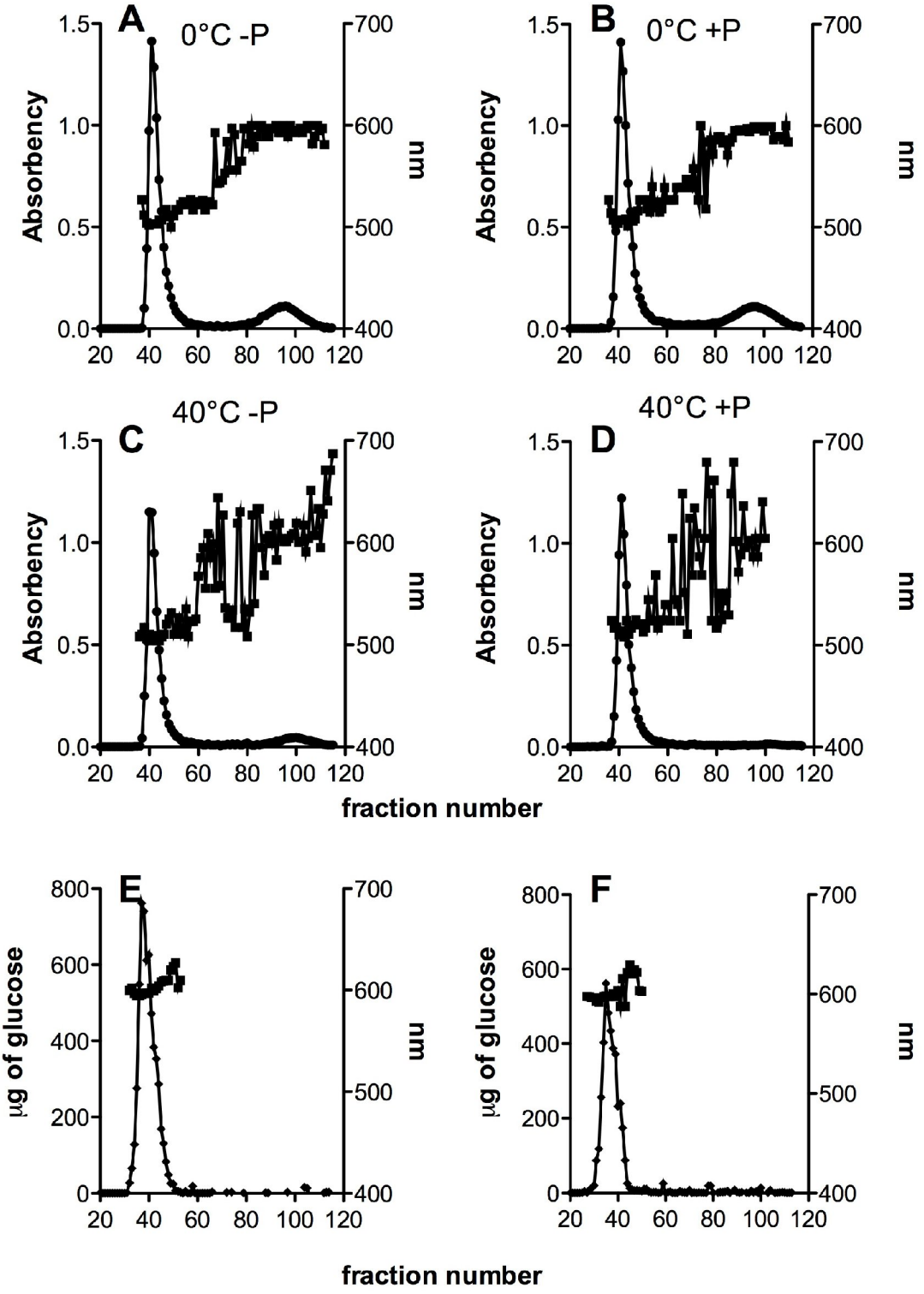

Figure 6: Size exclusion chromatographies of residual carbohydrate granules (A to D) and water soluble polysaccharide produced (E and F) after incubation at 0°C or 40°C in the absence or in the presence of inorganic phosphate. After incubation, residual starch granules were prepared as described in the materials and methods part. Both amylopectin and amylose polysaccharides were separated on size exclusion chromatography CL-2B.

The absorbency between 400 and 700 nm was determined for each fraction. Maximum absorbency values (black circle) and the wavelenght at the maximum absorbency, λmax, (black square) were reported on the graphs. Amylopectin and amylose are present in 38 to 60 fractions and 80 to 110 fractions respectively. (E and F) the molecular weight of WSP produced in –Pi at 40°C (E) and in +Pi at 40°C (F) were analyzed on size exclusion chromatography TSKHW 50. Polysaccharides present in each fraction were detected and quantified by phenol-sulfuric assay. The results are expressed as µg of equivalent glucose per fraction (black diamond). As above, the λmax values were defined for high molecular weight polysaccharides (black square).

During the whole process the apparent molecular weight and λmax values of amylopectin remained unchanged. Interestingly, the water-soluble polysaccharides produced in all experimental conditions display a λmax value estimated at 525 nm and are found in the exclusion volume of column (fraction 30 to 50) like glycogen particules and not distributed throughout column fraction. These results are in agreement with the presence of an endo-type of hydrolase releasing large size branched dextrins. We do not know if these dextrins outer chains are attacked by phosphorylase, if the latter restricts its action on amorphous-amylose-like material or if phosphorylase is immediately able to attach the outer chains of the amylopectin material on solid starch.

Discussion

Detergent treatment indicates that a large number of protein are loosely attached to the carbohydrate granules. Only phosphorylase, GBSS-like, Starch/glycogen synthases and an unidentified polypeptide of 43 kD remained tightly attached. In addition to the identification these polypeptides, a complementary experiment based on their localization (inside/surface) instead of affinity could be carried out based on the work of Mu-forster et al. (1998). By incubating carbohydrate granules with a protease (e.g thermolysin), only proteins on the surface will be hydrolyzed, and consequently we may observe a different pattern on SDS-PAGE which may reflect the proteins trapped during the biosynthesis pathway. The *in vitro* incubation experiments suggests that the presence of an "endo" type of starch hydrolase tightly bound to the granules which has escaped our first round of protein identification. This could be explained by the fact that we only focused on the major polypeptides. However, lines

of evidence suggest that some kind of amylase activity is present in Clg1 granules : (i) production of high mass branched WSP during the incubation experiments (ii) the identification of a 115kD amylase attached to carbohydrate granules produced by Cyanothece ATCC51142 (iii) the presence of gene coding for the corresponding 110 kD amylase in the genome of *Crocosphaera watsonii*. Interestingly, the digestion products of this amylase type activity, which represents 50% of the total amount of glucose after overnight incubation at 40°C, is not composed by a population of maltooligosaccharide but only by a large polysaccharide found in the exclusion volume. In addition, by combining both starch bound phosphorylase and amylase, 25% of WSP is converted into glucose-1-phosphate and the amylose fraction is completely hydrolized. This may suggest that phosphorylase works prefentially on the soluble polysaccharide and not on solid starch. To confirm this hypothesis, an incubation experiment with ^{14}C-Glucose-1-Phosphate could be carried out and by following the incorporation of ^{14}C-glucose into either the soluble fraction (WSP) or into carbohydrate granules, we will find out if the phosphorylase activity is able to work directly on semi-crystalline granules or requires a partially digested polysaccharide. Some hydrolylitic activities have to be able to hydrolyse direcly the semicrystalline polysaccharide. Indeed, unlike in plants, there is no gene coding for Glucan Water Dikinase or Phospho Water Dikinase in the genome of cyanobacteria and cyanobacterial starch does not appear significantly phosphorylated (M Steup: personal communication). We did not detect any glucose phosphorylated in position 3 and 6 (results not shown and M Steup personal communication). This suggest that the cyanobacteria accumulating such carbohydrate granules have circumvented the problems linked to the semi-crystallinite physical state of starch. interestingly, this high molecular weight “amylase” attached to the carbohydrate granules appears to be present in all cyanobacteria which accumulate insoluble semi-crystalline granules.

III

UV MUTAGENESIS ON CLg1 strain

Introduction

Most of mutagenesis protocols, based on the random integration of a transposon into the host chromosome, have been successfully performed in different cyanobacterial strains including *Nostoc ponctiform* and *Anabaena sp.*, two filamentous diazotroph cyanobacteria (Cohen et al., 1994; Maldener et al., 2003). When combined with an appropriate selection, transposon mutagenesis is indeed a powerful tool for genetic dissection and functional analysis. Such an approach was performed on Clg1 strain by using a highly efficient method of transposon mutagenesis developed for *Xanthobacter autotrophicus* and efficient for others prokaryotes (Larsen et al., 2002). Unfortunately, we did not get any results and so far, no insertional mutagenesis has ever been reported in the literature for *Crocosphaera* and *Cyanothece* species despite several attempts from at least three distinct groups. We don't understand clearly the reasons underlying this lack of success clearly depending on the cyanobacterial species concerned. For this reason, we turned to classical UV mutagenesis. This method was used essentially in the past with a variety of cyanobacteria (Astier et al., 1979).

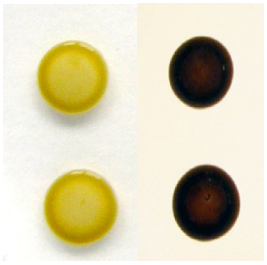

Figure 7: Forty microliters of Clg1 suspensions were loaded on ASNIII agar medium plate. The plate was incubated at room temperature during a month and a half under day/night cycles (12/12). The picture shows yellow-greenish color of cell patches (left) stain in dark-blue color (right) after iodine vaporization.

This strategy offers the advantage to circumvent all problems linked to the establishment of transgenesis techniques. Thus, we designed an experimental mutagenesis protocol for CLg1 strain based on the the one previously described by Golden (1988). Clg1 strain was used in the

UV-mutagenesis campaign instead of Cyanothece ATCC51142 for two reasons: on the one hand, Clg1 synthesizes true large size starch granules composed of both amylopectin and amylose and on the other hand, cell patches of wild-type Clg1 stain in dark-blue color after iodine vaporization (Figure 7). The dark-blue color observed with the Clg1 strain is probably due to the presence of both amylose and amylopectin in the starch granules. Because of the absence of amylose fraction in the granules of Cyanothece ATCC51142, cell patches do not color at all, making the screening tedious or impossible.

We have mutagenized Clg1 cells with ultraviolet irradiation at 254 nm and sub-cultured the surviving cells repeatedly in liquid medium for several generations. Because Clg1 cells might contain multiple of copies of a single chromosome, we allowed many cycles of cell division to segregate any mutations that might be recessive. Classical mutagenesis protocols are usually followed by an enrichment step which according to the desired mutation (*e.g* auxotroph mutants, unable to fix nitrogen), consists on the use of a specific medium preventing the growth of mutants, while dividing wild type cells are sensitive to a drug. In our particular case, this enrichment step could not be included in the mutagenesis protocol (figure 8).

Previous works emphasized the resistance of cyanobacteria to ultraviolet light (Thiel and Leone, 1986). Indeed, after several seconds of exposure, an efficient damage reparation system depending on light is activated. To avoid this photoreactivation, the irradiated cells have to be placed in darkness for a period of time (36 to 48 hours) before growing them on day/night cycles. In both parameters, time of exposure (15; 30; 60 and 120 seconds) and distance from our UV source (0.5; 3 and 6 cm) were tested and survival of Clg1 under each experimental condition was estimated. As expected, we observe a significant decrease of the number of single colonies on agar ASNIII medium plate in the case of long exposures and small distances from the UV light source.(The percentage of survival was estimated at 0.1% under 3cm and over 60 seconds).

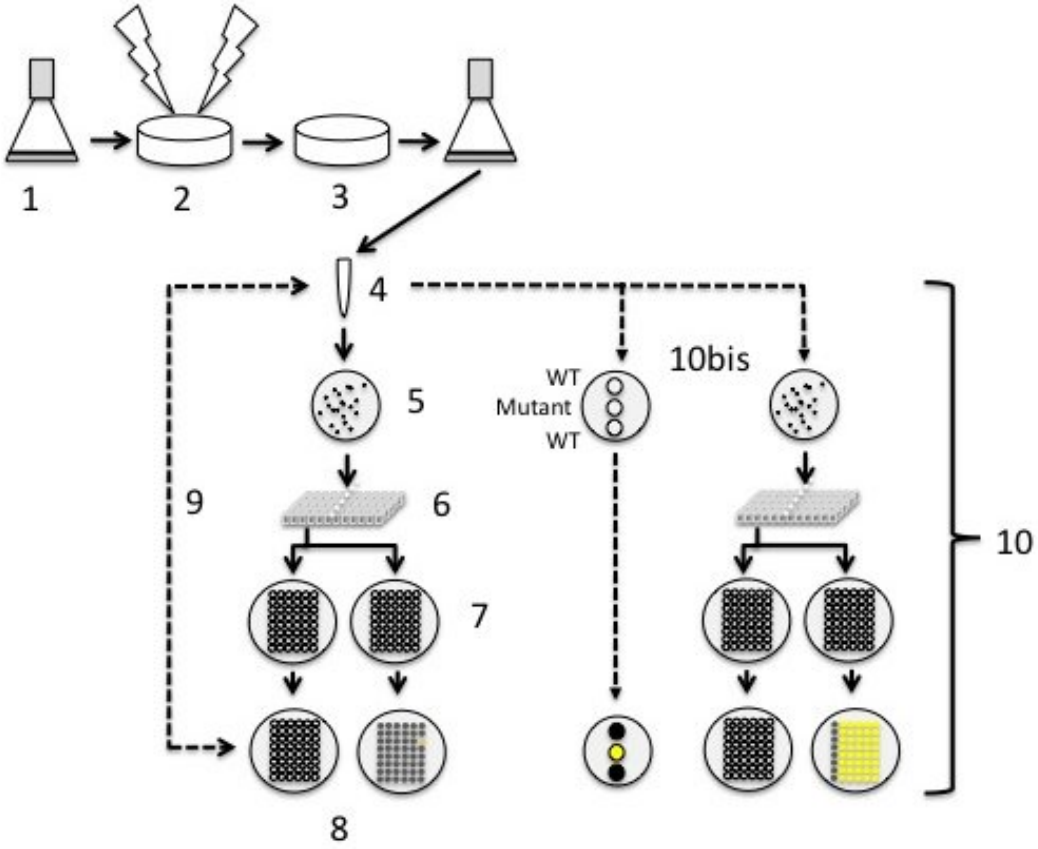

Figure 8: Mutagenesis protocol used on Clg1 strain. (1) Fresh cultures of Clg1 were harvested in the middle of the night and then resuspended in ASNII liquid medium. Cell suspensions were spread on agar-medium plates. (2) After drying, the cells were irradiated under UV light source for different time of exposure (15"; 30"; 60"; 120") and height. (0.5; 3; and 6 cm)(3) After 48 hours in the dark to avoid photoreactivation, the plates were incubated under day-night cycles (12hours/12 hours) (4) Surviving colonies were collected and grown in liquid medium in order to allow chromosome segregation. (5) Cells were suspended in ASNIII medium and spread on agar-medium plates. (6) Individual colonies were transferred in sterile 96 wells plates. (7) Duplicate cell patches were performed on agar medium plates (8) one agar-medium plate was screened with iodine vapors. The sister plate was used to collect the mutants. (9) These mutants were grown in liquid medium and screened once more with iodine vapors in order to confirm their phenotype. (10 and 10bis) Before completing the sub-cloning process, we confirmed the abnormal phenotype of mutant cell patches in comparison to wild-type cell patches by iodine vaporization (WT) (10bis). We then performed sub-cloning on each mutant to be sure that segregation was completed. In order to achieve this liquid diluted suspension has been sprayed on agar-medium plates and single colonies were selected (as in steps 5 8)). Among the latter a single cell patch displaying abnormal iodine staining has been picked up again. Sub-cloning was repeated until a uniform abnormal staining is obtained for each mutant.

For each condition (e.g : 6cm /1' exposure), colonies were pooled and cultivated in ASNIII liquid medium, then diluted and sprayed onto agar-medium in order to obtain single colonies In total, we screened roughly 20,000 clones and so far isolated 55 clones that display confirmed abnormal iodine staining of cell patches as shown on figure 3. Sub-cloning has been performed for 22 mutants so far. It must be stressed that a single phenotype test, made

after inoculation of a cell patch requires a minimum of 6 weeks. Growth of single colonies also require from 1 to 2 months. It thus took us 3 full years to complete this mutagenesis program with the full assistance of one technician.

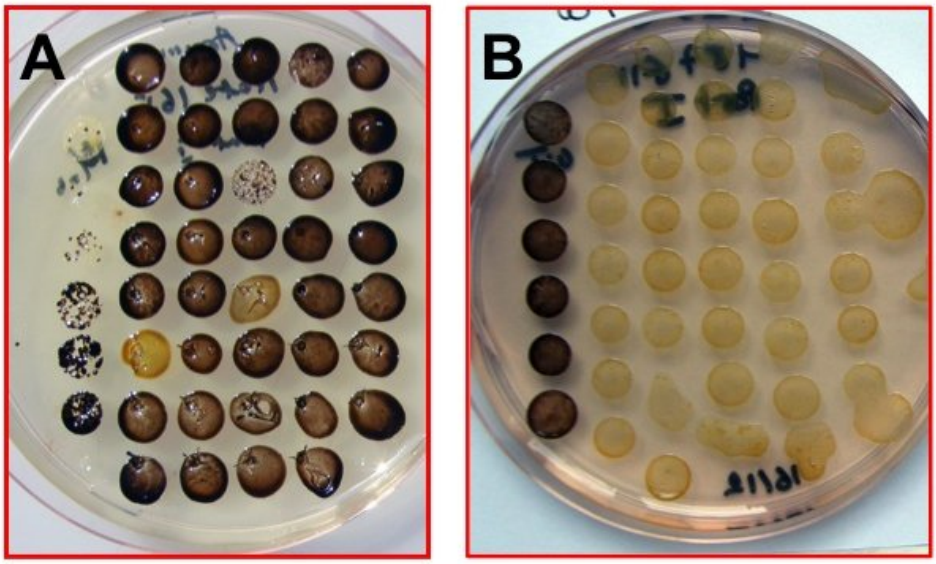

Figure 9: Examples of an iodine screening step (A) and a sub-cloning step (B).

Two major phenotypes, yellow-orange and brown, were observed during the screening step and confirmed at the sub-cloning step. Yellow-orange phenotypes could reflect a decrease in starch amount or a structural modification resulting in a modification of the iodine polysaccharide interaction. Further characterizations are required to identify the defect for each mutant.

Preliminary characterization of 6 mutants impaired in starch metabolism over 55.

The preliminary characterization is made to sort out the different mutant phenotype classes. This step is required before starting a detailed characterization for each mutant which requires in some instances additional subcloning steps to check for complete segregation. Nevertheless, mutagenized cells have grown for a sufficient number of generations to

segregate their chromosomes and consequently display any recessive mutation. Therefore, we carried out a quick determination of polysaccharide content as well as a variety of specific zymograms to visualize all starch metabolism enzymes.

Polysaccharide content: starch granules and Water-soluble polysaccharide.

In order to estimate the quantity of carbohydrate within granules or water-soluble polysaccharides, mutant cells were grown in liquid medium without nitrogen source and harvested at the middle of the day after one week of culture. Cell suspensions were disrupted by a French press. Carbohydrate granules and water-soluble polysaccharide were separated by centrifugation and quantified by the amyloglucosidase assay as described in the materials and methods. Results are expressed as milligram of equivalent glucose per milligram of protein. As shown in figure 10A and 10B, all the mutants accumulate less than 50% of the wild-type polysaccharide amount. Except for 1'C6 mutant, the starch contents were variable in the two independent cultures. However, we did not notice in both cultures the presence of water-soluble polysaccharide. In contrast to 118H9 and 174B6 mutants, they accumulate 1.5 or 3 times more WSP than the wild-type, a rather modest increase. 187G11, 15'E4 and 121B6 mutants accumulate an identical and two times less of WSP in comparison to wild-type cell, respectively.

Zymogram analysis

Native Page activity or zymogram analysis is a useful technique for detecting the activities of specific isoforms, and for making relative comparisons in a crude extract (Colleoni et al., 2003). Therefore, this technique was carried out on crude extracts of mutants and wild type cells cultivated in day/night cycles. All starch metabolism enzymes were visualized, except the debranching enzyme activity (GlgX), which was lost during the sample's preparation.

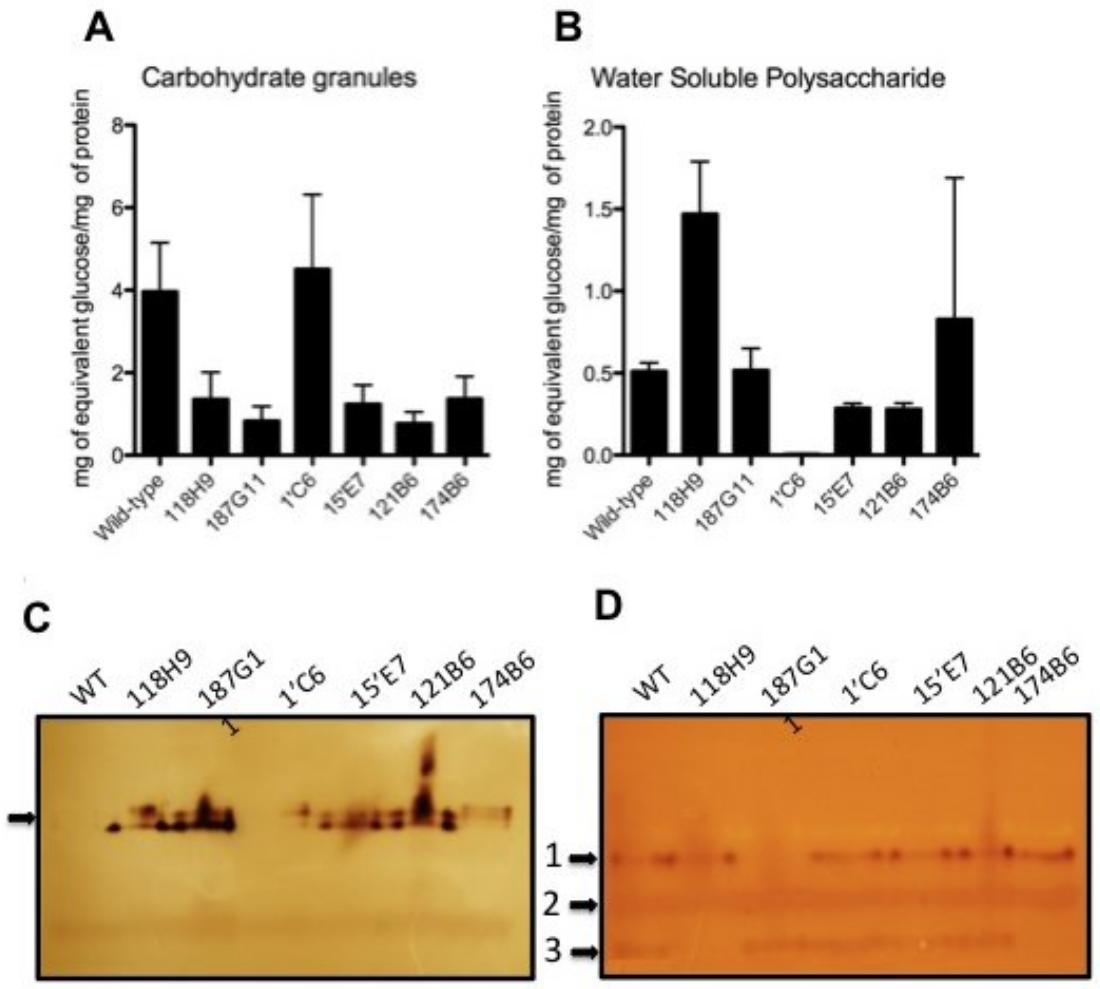

Figure 10: Preliminary characterization of Clg1 mutants. (A and B) carbohydrate granules and water-soluble polysaccharide contents were determined by the amyloglucosidase assay. Water-soluble polysaccharides represent branched or linear glucan chains made of two and more residues of glucose. The results are expressed as mg of equivalent glucose per milligram of protein. The data represents the mean of triplicate assays on two independent cultures. (C and D) phosphorylase and starch synthase zymograms were carried out on crude extract of Clg1 mutants and wild-type cells. 50 µg of total protein were separated on native-PAGE gels and then electroblotted one hour onto another native-PAGE containing 0.3% of glycogen. Glycogen gels were incubated either phosphorylase buffer containing glucose-1-phosphate or and starch synthase buffer containing ADP-glucose overnight at 30°C. After incubation, gels were stained with an iodine solution. Black arrow on phosphorylase points out the phosphorylase activity. Black arrows 1, 2 and 3 on starch synthase zymogram point out starch synthase activity, D-enzyme and an unknown activity, respectively.

For each zymogram, wild-type cells crude extracts were used as reference. Branching enzyme and starch zymogram results demonstrate that D-enzyme, amylases and branching enzyme activities were apparently unaltered in the mutants. However, phosphorylase activity, nearly undetectable in the wild type, is significantly more important in all mutant extracts except for 1'C6 mutant as seen on figure 10C (black arrow on the left).

Three mutants, 187G11, 118H9 and 174B6, display a reproducible modification of the starch synthase zymogram (black arrows 1 and 3 figure 10D). The upper brownish band missing in 187G11 Mutant is ADP-glucose dependant. In absence of ADP-glucose in the incubation buffer, the activity band is not visualized on glycogen gel. Consequently, this suggests that the 187G11 mutant may be impaired in a soluble starch/glycogen synthase activity. The second brownish band, pointed by black arrow 3, missing in both mutants, 118H9 and 174B6, is visualized in absence of ADP-glucose. Therefore, the brownish color results on a structural modification of the glycogen increasing the iodine interaction (like D-enzyme activity pointed by the black arrow 2). Interestingly, this activity shows up only with incubation buffer made for starch synthase activity and not with the regular buffer (Tris-acetate pH7.5 1 mM DTT) used to visualize starch metabolism enzymes.

Discussion

On six yellow mutants subjected to preliminary characterization, the phenotypes of 15'E7, 121B6 and 1'C6 mutant strains could not be correlated to a lack of activity. This could be due to punctual mutations, which reduce or modify enzyme specificity. Therefore, quantitative in vitro assays will be carried out on these mutants. Interestingly, two independent mutants (produced by two UV-mutagenesis conditions), 118H9 and 174B6 lack an unknown activity, which leads to some accumulation of water polysaccharide. The latter will be subjected to a structural characterization (e.g Size exclusion chromatography). The unknown activity will be purified and characterized. We were able to correlate the yellow phenotype of 187G11 to the low starch granule content and an absence of soluble starch synthase activity. Based on the genome of *Crocosphaera watsonii* (see table I in part II), two soluble starch/glycogen synthase genes and one GBSS-like gene are found. Because Clg1 belongs to the same sub-group of *C.watsonii*, we can reasonably suppose that Clg1 contains

also the same set of starch/glycogen genes. So far, we have identified two starch/glycogen starch proteins both tightly attached to carbohydrate granules; GBSS-like and starch/glycogen synthase (figure 4 in part II). To date, we don't know if the starch/glycogen synthase activity visualized on the zymogram corresponds to the starch/glycogen synthase associated to granules or to gene product of the third gene. SDS-PAGE analysis of proteins attached to carbohydrate granules performed on both 187G11 and wild-type strains could give us a clear answer. Furthermore, the absence of overlapping function suggests a specific role of starch/glycogen synthase activities in the starch metabolism pathway of Clg1.

General discussion and conclusion

General discussion and conclusion

The aim of our thesis was to characterize better the network of cyanophycean starch synthesis in the recently isolated Clg1 strain. This strain was successfully axenized and cultured in our laboratory on both solid and liquid media. Nevertheless the cells grew very slowly and took on average six weeks to generate cell patches from droplets inoculated for iodine screening. The growth of a small size colony took also between one to two months. Our first round of mutagenesis has failed to generate mutants (results not shown) and it took 3 full years to generate a second round of mutagenesis with exploitable results.

Despite these problems we were able to select out a minimum of one soluble starch synthase mutant and 3 other mutants with an identifiable enzymatic defect. Because a mere 20% of the certified mutant cell patches were subjected to preliminary characterization we can reasonably expect at least another 5 exploitable mutants making the whole project profitable. We believe it is of paramount importance to dissect genetically the starch metabolism network in cyanobacteria. Indeed the bacterial pathways are expected to differ fundamentally from those of starch metabolism both in Chloroplastida and Rhodophyceae or Glaucophyta. Cyanobacteria in fact do not appear to contain debranching enzymes with isoamylase specificity nor do they appear to phosphorylate their starch. We have been able to demonstrate that solid starch releases a substantial amount of dextrins and glucose-1-P by simple incubation of the granules in the presence of orthophosphate which is suggestive of entirely different pathway of starch mobilization in these organisms by comparison to green plants and algae. The function of an enzyme seems thus to be just as much dependent on the evolutionary history of the network to which this enzyme belongs as on its catalytic activity per se. If this is so we can predict that the starch metabolism network of cyanobacteria might function entirely differently.

A deep understanding of starch metabolism will require to understand the function of an enzyme through the entire evolutionary process: that is in cyanobacteria for those enzymes that originated from these organisms, in the cytosol of the heterotrophic eukaryotes for those of eukaryotic ancestry, after the merging of the pathways in the first archaeplastidal cell and within each of the major archaeplastidal lines.

Bibliography

Bibliography

Astier C, Joset-Espardellier F, Meyer I. 1979. Conditions for mutagenesis in the Cyanobacterium Aphanocapsa 6714. *Archives of Microbiology* **120**, 93-96.

Colleoni C, Myers A.M, James M.G. 2003. One- and two-dimensional native PAGE activity gel analyses of maize endosperm proteins reveal functional interactions between specific starch metabolizing enzymes. *The Japanese Society of Applied Glycoscience* **50**, 207-212.

Compaoré J, Stal L.J. 2010 Oxygen and the light –dark cycle of nitrogenase activity in two unicellular cyanobacteria. *Environmental Microbiology* **12**, 54-62.

Dauvillée D, Colleoni C, Shaw E, Mouille G, D'Hulst C, Morell M, Samuel MS, Bouchet B, Gallant DJ, Sinskey A, Ball S. 1999. Novel starch-like polysaccharides are synthesized by a soluble form of granule-bound starch synthase in glycogen accumulating mutants of *Chlamydomonas reinhardtii. Plant Physiology* **119**, 321-330.

Delrue B, Fontaine T, Routier F, Decq A, Wieruszeski J.M, Van den Koornhuyse N, Maddelein M.L, Fournet B, Ball S. 1992. *Waxy Chlamydomonas reinhardtii*: monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin. *The Journal of Bacteriology* **174**, 3612-3620.

Falcòn L, Lindvall S, Bauer K, Bergman B, Carpenter E.J. 2004. Ultrastructure of unicellular N2 fixing cyanobacteria from the tropical north atlantic and subtropical north pacific oceans. *Journal of phycology* **40**, 1074-1078.

Frueauf J.B, Ballicora M.A, Preiss J. 2002. Alteration of inhibitor selectivity by site-directed mutagenesis of Arg294 in the ADP-glucose pyrophosphorylase from *Anabaena* PCC7120. *Archives of Biochemistry and Biophysics* **400**, 208-214.

Gabriel O, Gersten D.M. 1992. Staining for enzymatic activity after gel electrophoresis, *Analytical Biochemisty* **203**, 1-21.

Ghosh H.P, Preiss J. 1966. Adenosine Diphosphate Glucose pyrophosphate: a regulatory enzyme in the biosynthesis of starch in spinach leaf chloroplasts. *The Journal of Biological Chemistry* **10**, 4491-4504.

Golden S.S. 1988. Mutagenesis of cyanobacteria by classical and gene-transfer-based methods. *Methods in Enzymology* **167**, 714-727.

Lacks S.A, Springhorn S.S. 1980. Renaturation of enzymes after polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. *The Journal of Biological Chemistry* **255,** 7467-7473.

Larsen R.A, Wilson M.M, Guss A.M, Metcalf W.W. 2002. Genetic analysis of pigment biosynthesis in Xanthobacter autotrophicus Py2 using a new, higly efficient transposon mutagenesis system that is functional in wide variety of bacteria. *Archives of Microbiology* **178**, 193-201.

Liu Y, Golden S.S, Kondo T, Ishiura M, Johnson C. H. 1995. Bacterial luciferase as a reporter of circadian gene expression in cyanobacteria . *The Journal of Bacteriology* **177**, 2080-2086.

Maddelein M.L, Libessart N, Bellanger F, Delrue B, D'Hulst C, Van Den Koornhuyse N, Fontaine T, Wieruszeski J.M, Decq A, Ball S.G. 1994. Toward an understanding of the biogenesis of the starch granule: determination of granule-bound and soluble starch synthase functions in amylopectin synthesis. *The Journal of Biological Chemistry* **269,** 25150-25157.

Mohr W, Intermaggio M.P, Laroche J. 2009. Diel rhythm of nitrogen and carbon metabolism in the unicellular, diazotrophic cyanobacteria Crocosphaera watsonii WH8501. *Environmental Microbiology* **12(2)**, 412-421.

Mu-Forster C, Wasserman B. P. 1998. Surface localization of zein storage proteins in starch granules from maize endosperm. Proteolytic removal by thermolysin and in vitro cross-linking of granule-associated polypeptides . *Plant physiology* **116**, 1563-1571.

Ral J.P, Colleoni C, Wattebled F, Dauvillée D, Nempont C, Deschamps P, Li Z, Morell M, Chibbar R, Purton P, D'hulst C, Ball S. 2006. Circadian clock regulation of starch metabolism establishes GBSSI as a major contributor to amylopectin synthesis in Chlamydomonas reinhardtii. *Plant physiology* **142**, 305-317.

Reddy K.J, Haskell B, Sherman D.M, Sherman L.A. 1993. Unicellular, Aerobic Nitrogen-fixing Cyanobacteria of the genus Cyanothece. *The Journal of Bacteriology* **175**, 1284-1292.

Reimels A.J, Reddy K.J. 1995. Analysis of Cyanothece sp.BH68K mutants defective in Aerobic nitrogen. *Current Microbiology* **32**, 174-175.

Rippka R, Deruelles J. 1979. Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Journal of General Microbiology* **111(1)**, 1-61.

Ritte G, Loberth R, Steup M. 2000. Reversible binding of the starch-related R1 protein to the surface of transitory starch granules. *The Plant Journal* **21**, 387-391.

Schneegurt M.A, Sherman D.M, Nayar S, Sherman L.A. 1994. Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the Cyanobacterium cyanothece sp. ATCC51142. *The Journal of Bacteriology* **176**, 1586-1597.

Thiel T, Leone M.J. 1986. Effect of glutamine on growth and heterocyst differentiation in the cyanobacterium *Anabaena* variabilis. *The Journal of Bacteriology* **168**, 769.