Journal of Experimental Botany, Vol. 62, No. 6, pp. 1775–1801, 2011
doi:10.1093/jxb/erq411 Advance Access publication 10 January, 2011

Journal of Experimental Botany
www.jxb.oxfordjournals.org

DARWIN REVIEW

The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis

Steven Ball*, Christophe Colleoni, Ugo Cenci, Jenifer Nirmal Raj and Catherine Tirtiaux

Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576 CNRS-USTL, Bâtiment C9, Cité Scientifique, F-59655 Villeneuve d'Ascq, France

* To whom correspondence should be addressed: E-mail: steven.ball@univ-lille1.fr

Received 10 September 2010; Revised 18 November 2010; Accepted 23 November 2010

Abstract

Solid semi-crystalline starch and hydrosoluble glycogen define two distinct physical states of the same type of storage polysaccharide. Appearance of semi-crystalline storage polysaccharides appears linked to the requirement of unicellular diazotrophic cyanobacteria to fuel nitrogenase and protect it from oxygen through respiration of vast amounts of stored carbon. Starch metabolism itself resulted from the merging of the bacterial and eukaryote pathways of storage polysaccharide metabolism after endosymbiosis of the plastid. This generated the three Archaeplastida lineages: the green algae and land plants (Chloroplastida), the red algae (Rhodophyceae), and the glaucophytes (Glaucophyta). Reconstruction of starch metabolism in the common ancestor of Archaeplastida suggests that polysaccharide synthesis was ancestrally cytosolic. In addition, the synthesis of cytosolic starch from the ADP-glucose exported from the cyanobacterial symbiont possibly defined the original metabolic flux by which the cyanobiont provided photosynthate to its host. Additional evidence supporting this scenario include the monophyletic origin of the major carbon translocators of the inner membrane of eukaryote plastids which are sisters to nucleotide-sugar transporters of the eukaryote endomembrane system. It also includes the extent of enzyme subfunctionalization that came as a consequence of the rewiring of this pathway to the chloroplasts in the green algae. Recent evidence suggests that, at the time of endosymbiosis, obligate intracellular energy parasites related to extant *Chlamydia* have donated important genes to the ancestral starch metabolism network.

Key words: Archaeplastida, *Chlamydia*, cyanobacteria, endosymbiosis, evolution of photosynthesis, glycogen, plastids, starch.

Introduction

Sometime between 0.7–1.5 billion years ago (Cavalier-Smith, 2006; Yoon *et al.*, 2004) an ancestor of present-day cyanobacteria was internalized, probably through phagocytosis (Raven *et al.*, 2009) 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 and van Dooren, 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 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 (Fig. 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). 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 two and most of the time by four 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 a review see Viola *et al.*, 2001). The term floridean starch was therefore 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 recently been reviewed (Comparat-Moss and Denyer, 2009). In this Darwin Review, the focus will be 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.

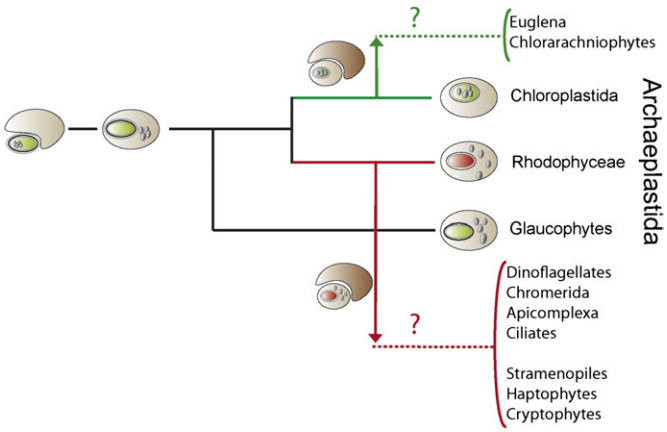

Fig. 1. Primary and secondary plastid endosymbiosis. Photoysynthetic eukaryotes are derived from a unique event involving phagoyctosis of a cyanobacterial ancestor by a heterotrophic eukaryotic host. The ancestral cyanobiont is depicted as a peptidoglycan-containing single cell cyanobacterium (in green) with both inner and outer membranes and no outer layer capsular polysaccharides as is presently the case for the *Paulinella* chromatophores or the glaucophyte cyanelles. The Archaeplastida define the extant photosynthetic eukaryotic lineages that have emerged from this unique ancestor. Among the Archaeplastida, the glaucophytes define single-cell freshwater algae containing a plastid called the cyanelle with phycobilisomes and other typical cyanobacterial-like features but displaying the same level of genome simplification and organization as other plastids. Starch is found in the cytosol of all glaucophytes. Red (Rhodophyceae) and green (Chloroplastida) algae contain plastids with no peptidoglycan called, respectively, rhodoplasts and chloroplasts. They can be distinguished by the structure and composition of their photosynthetic antennae which, in Chloroplastida, contain chlorophyll *b* while red algae still rely on bacterial phycobilins such as the red pigment phycoerythrin. Starch is found in the cytosol of red algae while it is found in plastids of all Chloroplastida including the land plants. It is presently thought that the cyanobacterial ancestor was a starch accumulator while the heterotrophic eukaryotic host partner synthesized glycogen in its cytosol. The ability to synthesize starch was transmitted to the archaeplastidal ancestor cytosol while it was lost by the cyanobiont. Upon evolution of the Chloroplastida, starch metabolism was rewired to the evolving chloroplasts. The Archaeplastida themselves became the substrate for 'secondary' endosymbiosis through phagocytosis by other heterotrophic eukaryotic lineages. The exact number of secondary endosymbiosis events is still debated (hence this is symbolized by dotted lines). Unlike primary endosymbiosis, the phagocytosis vacuole was either kept or fused with the ER, thus yielding, in most cases, four-membrane 'secondary' plastids. Starch was not universally transmitted to secondary endosymbiosis lines which mostly accumulate β-glucans. Starch, however, is still found in the cytosol of most dinoflagellates and some apicomplexa parasites and between the 2nd and 3rd membrane of secondary cryptophyte plastids.

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 accounts for 7–10% of the linkages and are evenly 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 two branches. This branching pattern allows for spherical growth of the particle generating tiers (a tier corresponds to the spherical space separating two consecutive branches from all chains 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 (Fig. 2A).

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 55 000 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, therefore, define 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-crystalline 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 pre-existing amylopectin-containing granule for its formation (Dauvillée *et al.*, 1999). Mutants deprived of this fraction can be readily isolated in green plants and algae (for a review see Ball *et al.*, 1998). These mutants build

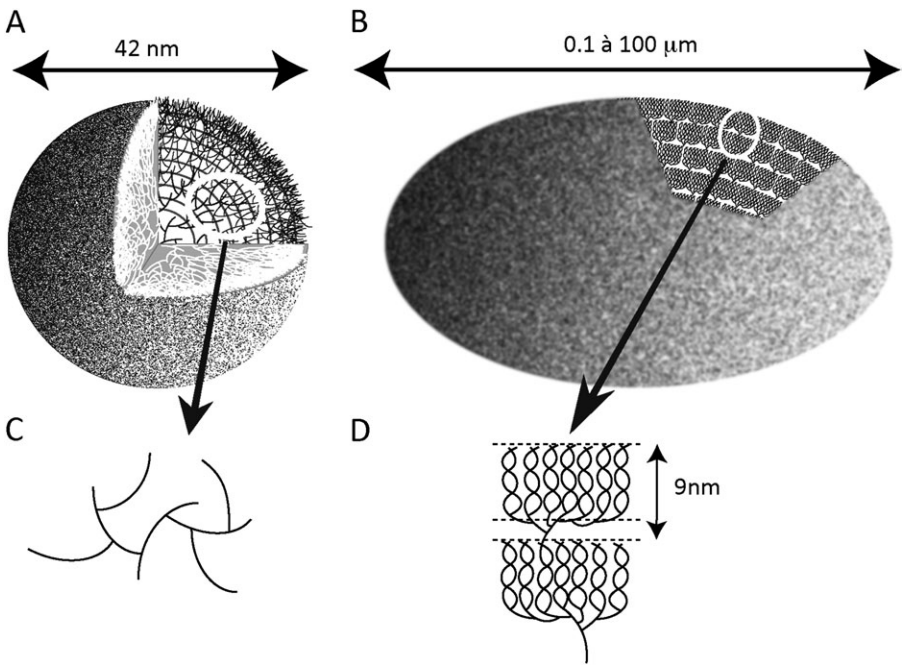

Fig. 2. Schematic representation of whole glycogen (A) and starch (B) granules. The lines represent α-1,4-linked glucan chains and the intersections of such lines symbolize the α-1,6 branches. (C, D) Enlarged views of the circled sections of the corresponding glycogen (C) and starch (D) granules. The distribution of branches exemplified in (C), with two α-1,6 linkages per glucan, leads to the exponential increase in the density of chains as one moves away from the centre of the particle. This leads to a predictable maximum of 42 nm for the glycogen granule displayed in (A). Indeed, further density increases will not accommodate the sizes of the glycogen metabolism enzymes active sites. (D) Two typical amylopectin clusters are displayed. The cluster structure is generated through the asymmetric distribution of the branches which are shown at the base of each of the two clusters. The small portion containing the branches is called the amorphous lamella of the unit cluster while the chains generated through the branches intertwine to form the double helical structures that define the unit crystalline lamella. The sum of one amorphous and one crystalline lamella amounts to 9 nm in all amylopectin clusters examined so far.

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.*, 2008*d*) 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 10^5–10^6 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 Fig. 2B, the branches are concentrated in sections of the amylopectin molecule leading to clusters of chains that allow for indefinite growth of the polysaccharide. Another major feature of the amylopectin cluster structure consists of 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 cluster and neighbouring clusters align and form sections of crystalline structures separated by sections of amorphous material (containing the branches) thereby generating the semi-crystalline 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 allows 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 able to extend amylose chains by synthesizing α-1,4 glucosyl linkages progressively within the polysaccharide matrix (reviewed in Ball *et al.*, 1998). Because no other enzyme is significantly 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. Consequently, starch can be seen as a very efficient intracellular sink immobilizing 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 humans 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.

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 (cyanobacterial and chlamydial) or eukaryotic glycogen metabolism (Coppin *et al.*, 2005; Patron and Keeling, 2005; Deschamps *et al.*, 2008*a*). 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, their common and distinctive features will be briefly outlined. Figure 3 summarizes the basic common pathway of storage polysaccharide synthesis in gram negative bacteria (and cyanobacteria) (for a review see Preiss, 1984) and opistokonts (for reviews see Roach, 2002; Wilson *et al.*, 2010).

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

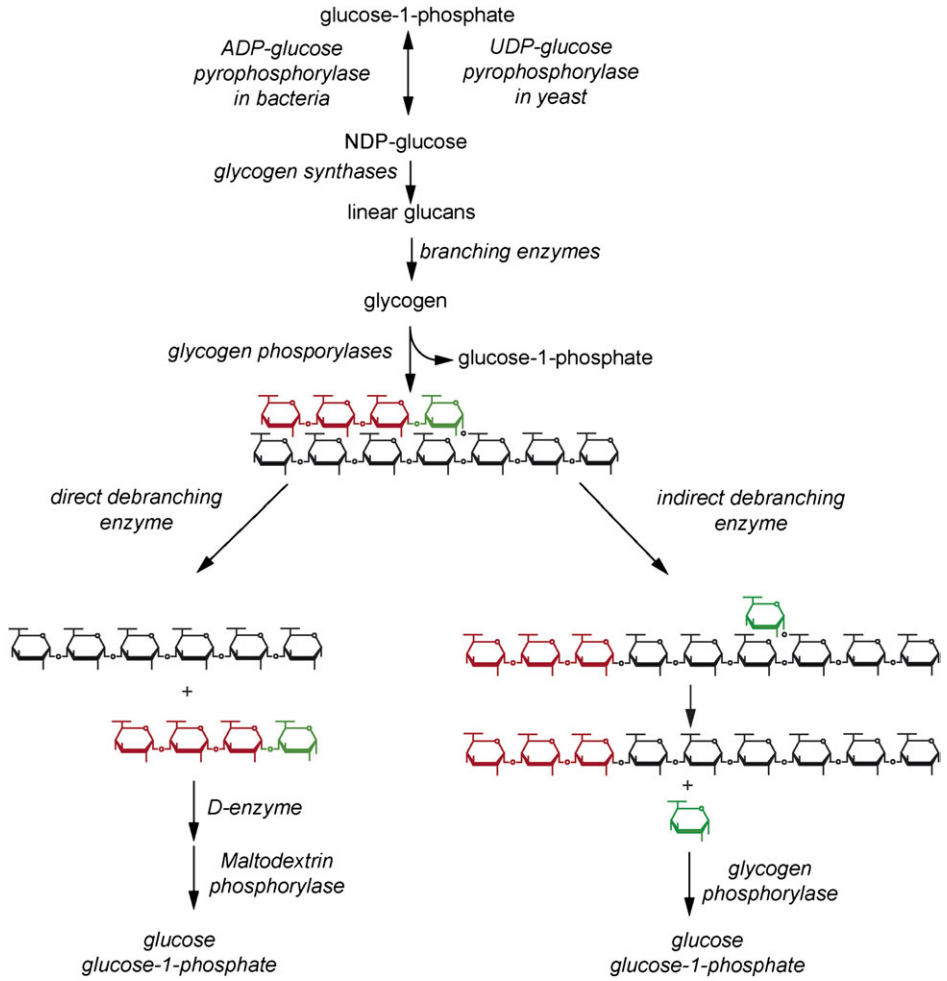

Fig. 3. Overview of glycogen metabolism in bacteria and eukaryotes. Both bacteria and eukaryotes synthesize glycogen from activated nucleotide-sugar substrates. The latter consist of ADP-glucose in bacteria and UDP-glucose in eukaryotes. The nucleotide sugar (NDP-glucose) is, in both cases, used for the transfer of glucose on the non-reducing end. Branching involves similar enzymes and reactions in both cases (see text). Glycogen breakdown is initiated in bacteria and eukaryotic pathways by glycogen phosphorylase which, in the presence of orthophosphate, generates glucose-1-P from the available non-reducing ends. In both cases, phosphorylase stops four glucose residues away from an α-1,6 branch generating the outer chain polysaccharide structure illustrated with the α-1,6 bound glucose in green and the novel chain generated by the branch symbolized by glucose residues in red. In bacteria (left column of the drawing), the direct debranching enzyme releases a maltotetraose (four Glc residues) and a glycogen particle with a longer outer chain (in this example, six Glc residues) that becomes a substrate for glycogen phosphorylase. In eukaryotes (right column of the drawing) indirect debranching enzyme hydrolyses the α-1,4 linkages next to the branch and transfers it on the neighbouring outer chain leading to further release of glucose-1-P by phosphorylase. The unmasked branch (in green) is then hydrolysed by a second active site within the same indirect DBE enzyme. Release of maltotetraose by the direct DBE of bacteria (left column) requires the presence of MOS metabolism enzymes such as D-enzyme and maltodextrin phosphorylase.

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 catalysed 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 the α-1,6 position. The branched polymers are subjected 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 four glucose residues away from the branch (Dauvillée *et al.*, 2005; Alonso-Casajús *et al.*, 2006). Therefore the short four-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 degrade this seven-glucoseresiduelong chain back to four while

the second catalytic site will hydrolyse the α-1,6 linkage from the residual unmasked glucose at the branch (for reviews see Roach, 2002; Wilson *et al.*, 2010). 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 four-glucose-residue-long malto-oligosaccharide (maltotetraose) (Dauvillée *et al.*, 2005). 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 maltotriose enabling maltodextrin phosphorylase to degrade the chains further. 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 is not found in opistokont genomes.

In addition to this phosphorolytic pathway of glycogen mobilization, there is good evidence for the presence of a hydrolytic pathway in opistokonts and circumstantial evidence for the presence of such a pathway in gram-negative bacteria. Fungi and animals indeed contain an enzyme able to hydrolyse both the α-1,4 and α-1,6 linkages responsible for the degradation of a significant pool of cellular glycogen (reviewed by François and Parrou, 2001; Roach, 2002; Wilson *et al.*, 2010). 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 the regulation of glycogen metabolism (Wang *et al.*, 2001). Undisputable functional proof of the importance of the glycogen hydrolysis pathway has been obtained, both in yeasts where it is triggered during sporulation or the 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 François and Parrou, 2001; Roach, 2002; Wilson *et al.*, 2010). 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 organisms, 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 specialization of the bacterial species (Preiss, 1984). 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 a 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; Wilson *et al.*, 2010). 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 6–12 enzymes of related function.

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; Roach 2002; Shearer and Graham, 2002; Wilson *et al.*, 2010). 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 (Fig. 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 polysaccharide 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) (Fig. 4). *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* harbour 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.*, 2008*a*; 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.*, 2008*a*) 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.*, 2008*a*). 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 (Coppin *et al.*, 2005; Aury *et al.*, 2006; Eisen *et al.*, 2006; Carlton *et al.*, 2007; Deschamps *et al.*, 2008*a*). Dpe2 and β-amylase were first reported in green plants and believed therefore to define green-lineage-specific genes. However, 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) argues that their presence cannot be explained by lateral gene transfer from Chloroplastida. The most logical explanation would consist of the existence of a richer suite of genes of glycogen metabolism in the eukaryotic ancestors 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 only have lost 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 fewer gene losses than other eukaryotes.

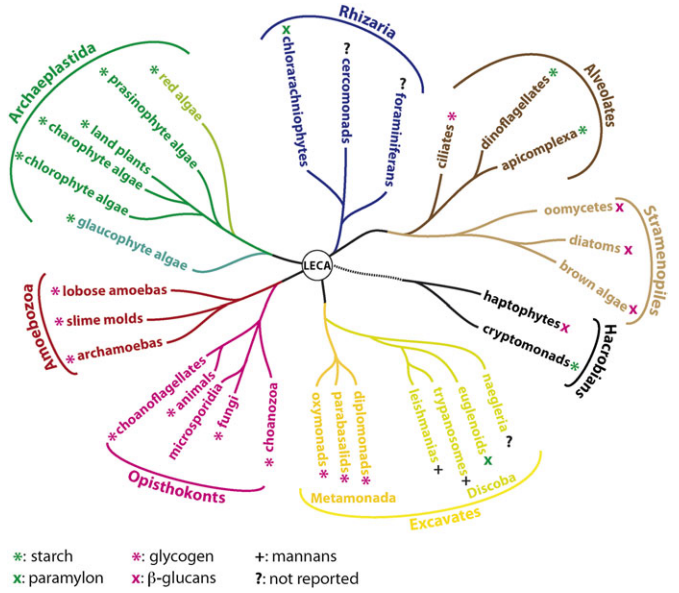

Fig. 4. Eukaryotic tree of life. The tree summarizing and simplifying our current understanding of eukaryotic phylogeny was inspired by those recently published by Baldauf (2003), Hampl *et al.* (2009), and Koonin (2010). The position of Hacrobians within the tree is problematic since they have been shown by Hampl *et al.* (2009) to disrupt the Archaeplastida monophyly. LECA illustrates the last eukaryotic common ancestor. When known, the nature of the storage polysaccharide present is reported; when unknown, the latter is displayed with a question mark (?). In short, eukaryotes can be subdivided into α and β glucan accumulators. In both cases, hydrosoluble oligo or polysaccharides and crystalline polysaccharides have been reported. Starch and paramylon define the solid insoluble crystalline α and β glucan storage polysaccharides, respectively, while maltoligosaccharides, glycogen, laminarin, chrysolaminarin (leucosin), and mycolaminarin define the hydrosoluble storage oligosaccharide and polysaccharide versions.

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 GT5 UDP-glucose utilizing glycogen synthase 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 *Entamoeba histolytica* has been chosen 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.

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 complex multicellular terrestrial plants such as maize or rice (Ral *et al.*, 2004; Derelle *et al.*, 2006; Deschamps *et al.*, 2008*b*). 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*a*; 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 fewer 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 catalyse each of the steps that have been outlined in the preceding sections. For instance, a minimum of five starch synthases participate in polymer elongation, three branching enzymes are reported to introduce the α-1,6 linkage, four 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 (Lin *et al.*, 1988; Zabawinski *et al.*, 2001). The enzymes of ADP-glucose production and those that elongate glucans with this substrate display a distinctive bacterial phylogeny which apparently correlates with the plastidial location of starch in the green lineage (Coppin *et al.*, 2005; Patron and Keeling, 2005, Deschamps *et al.*, 2008*a*). 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 directly to 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 loosen the tight crystal packing of glucans locally within the granule and to 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 degrade 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.* 2008*a*). 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, there is no indication that storage polysaccharides are phosphorylated in cyanobacteria and no equivalent to GWD, PWD, β-amylase, and dpe2 can be observed 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, 2005; Deschamps *et al.*, 2008*a*) (Table 1).

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 a 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.*, 2008*d*). Rhodophyceae are very poor biochemical and genetic models and no starch accumulating red alga can fulfil 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.*, 2008*d*; Dauvillée *et al.*, 2009). Severely impaired mutants of *C. cohnii* were demonstrated to have a decreased and modified UDP-glucose requiring starch synthase (Dauvillée *et al.*, 2009). The defect in starch amount and the 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.

Four Rhodophyceae genomes have recently been 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. Fewer 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 the 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; Shimonaga *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, 2005; Deschamps *et al.*, 2008*a*; Plancke *et al.*, 2008). The only major difference is defined by the absence of ADP-glucose

Table 1. The number of isoforms found for each class of glycogen/starch metabolism enzymes

Using phylogenetics, it was possible to 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 and those that were presumably transferred by Chlamydia are displayed in pink. Enzymes of uncertain origin are shaded in grey. The cyanobacteria, eukaryotes, and green plants display little variations in the corresponding sets of enzymes. *Crocosphaera watsonii*, *Entamoeba histolytica*, *Cyanidioschizon merolae*, and *Ostreococcus tauri* were chosen as paradigm genomes for, respectively, cyanobacteria, heterotrophic eukaryotes, red algae, and green plants. Reconstruction of starch metabolism in the common ancestor is emphasized within the black box and has been performed as explained in the text.

Activity	Cyanobacteria (*Crocosphaera watsonii*)	Eukaryotes (*Entamoeba histolytica*	Common ancestor	Green lineage (*Ostreococcus tauri*)	Red lineage (*Cyanidioschizon merolae*)
ADP-glucose pyrophosphorylase	1	0	1	2	0
Soluble starch synthase (ADPG) SSIII-SSIV	2	0	1	3	0
Soluble starch synthase (ADPG) SSI-SSII			1	2	
Soluble starch synthase (UDPG)	0	1	1	0	1
GBSS I	1	0	1	1	1
Branching enzyme	3	1	1	2	1
Isoamylase	1	0	2	3	2
Indirect debranching enzyme	0	1	0	0	0
Phosphorylase	2	2	2	2	1
Glucanotransferase	1	0	1	1	0
Transglucosidase	0	2	1	1	1
β-Amylase	0	4	1	2	1
Glucan water dikinase	0	?	1	3	1
Phosphoglucan water dikinase	0	?	1	2	1
Laforin or Sex4 type phosphatases	0	1	1	1	1

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, 2008*a, d*; Shimonaga *et al.*, 2007). In addition the floridean starch GBSS shows a marked preference for UDP-glucose while remaining capable of using 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.*, 2008*a*). 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 fulfil all functions which, in Chloroplastida, seem to require four different soluble starch synthases. This enzyme was initially thought by Patron and Keeling (2005) to descend from the cyanobacterial GT5 enzymes. However, at the time of their study, these authors were unable to realize that, 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 to that proposed for Chloroplastida. The absence in the rhodophycean genome of dpe1 (D-enzyme), an enzyme required for the 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.

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 a 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 harboured a cryptic plastid that resulted from the 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 represented 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 β-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) and Deschamps *et al.* (2008*a*). 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 exquisitely 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 lower the O_2 level further 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.* (2008*a*) proposed that this yielded a selection pressure for the change of glycogen metabolism into the synthesis of semi-crystalline polymers. Indeed, this would enable the storage of larger amounts of osmotically inert carbon with a lower 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 unusually large size of the granules without any detailed structural analysis.

Deschamps *et al.* (2008*a*) 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; 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.*, 2008*a*). 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.*, 2008*a*). The presence of amylose in the Clg1 starch prompted Deschamps *et al.* (2008*a*) 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.*, 2008*a*). 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.* (2008*a*) clearly proposed that the plastid ancestor was indeed a cyanobacterial ancestor of subgroup V.

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.* (2008*a*) 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 exemplified in *Entamoeba histolytica* (Loftus *et al.*, 2005). The relevance of this choice by the finding of a richer set and diversity of important enzymes in amoebas has previously been discussed. The starch metabolism network of *Crocosphaera watsonii* was chosen as a model subgroup V starch-accumulating diazotrophic cyanobacterium. Table 1 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.*, 2008*a*). In addition in phylogenetic trees, the common chloroplastidal and rhodophycean enzyme sequences display a common origin (Coppin *et al.*, 2005; Patron and Keeling, 2005; Deschamps *et al.*, 2008*a*). 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, it has not been possible 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 humans. This accumulation concerns many different tissues and organs, including the brain, thereby 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 because this unknown kinase may have been unsuitable to phosphorylate the more 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.

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 argue that the third scenario defines the only plausible situation. There are several complex reasons for this that are 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 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* 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 1000 genes and 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 observed during the evolution of plastids, although some of the experienced gene losses would have prevented independent life of the chromatophore.

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 three branching enzymes and two phosphorylases and possibly more if it is considered that isoamylases and soluble starch synthases may be of chlamydial rather than cyanobacterial origin (see below). 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 three 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, it could be argued that these enzymes are certainly not functionally equivalent to the cyanobacterial phosphorylase. Indeed, the latter seems able to attack solid cyanobacterial starch directly 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 in green algae 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). It must be stressed that, in the earliest diverging green algae, only the starch pathway seems to stand out by a high level of functional redundancy and subfunctionalizations. Other pathways have indeed not yet reached the level of complexity that can sometimes be found in terrestrial plants. There are only two alternatives to be considered at this point: either starch synthesis was exclusively cytosolic as we proposed or storage polysaccharides were synthesized both in the plastid and cytosol in contradiction to our hypothesis. In the first

case, the cyanobiont would, very early, have lost its ability to store starch, while, in the second, the latter would have been maintained.

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 whole starch metabolism network from the cytosol to the plastid. In following sections, the means by which this could be achieved will be detailed. Suffice to say now that this process does explain 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 a 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 any other photosynthesis genes whose products has remained in the cyanobiont does not require any gene duplication and enzyme subfunctionalization and these are indeed not observed as extensively in such networks.

Because the selective complexity of the starch metabolism network (at variance with land plant pathways, only starch metabolism displays a high level of gene redundancy in green algae) can only be explained through the complex rewiring mechanism that selectively takes place if the cyanobiont had lost the ability to store starch, we believe this pleads for an ancient cytosolic 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 more closely to resemble the ancestral Archaeplastida (the Glaucophyta) is among them.

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 both the UDP-glucose and ADP-glucose requiring starch synthases and all other enzymes of bacterial or host phylogeny. How could the 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 into 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 with that of spontaneous mutations in a given gene (Thorsness and Fox, 1990). Similar experimental results with similar 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 not filled with organelle sequences, one has to imagine that most of these events will be counter-selected and that losses of such nuclear sequences will be at least as frequent.

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 routinely to 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 harbour 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 funnelled 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.* (2008*a*) 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 Fig. 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 Fig. 5 is no more nor 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.

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 the 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 cross-talk or possibility of regulation between the two unrelated biochemical networks. Yet the carbon must be exported only when the cyanobiont is able to supply photosynthate 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 membrane then the net result for the homeostasis of the cyanobiont's metabolism is likely to be toxic and, 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

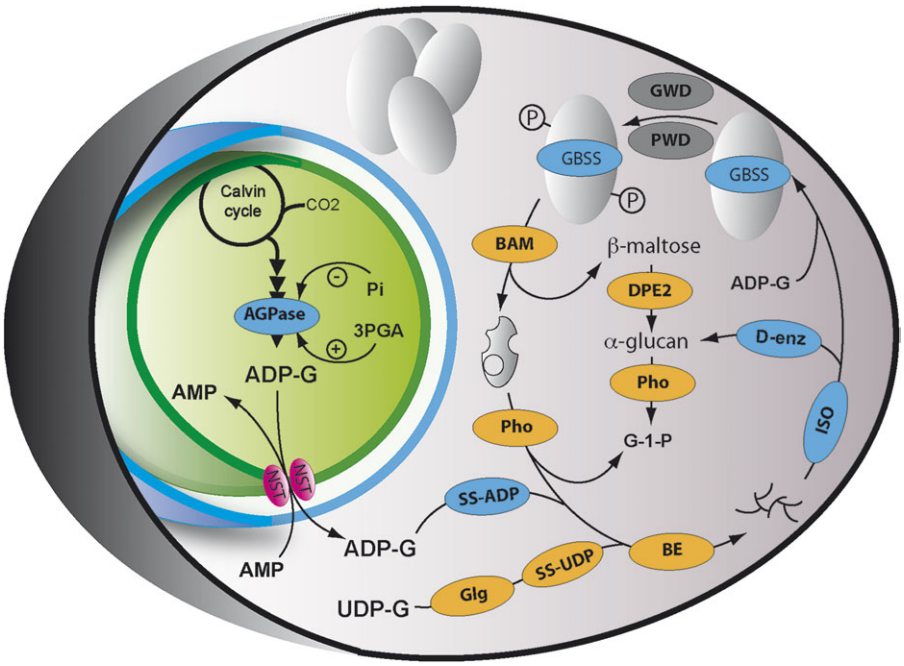

Fig. 5. Reconstruction of starch metabolism in the common ancestor of Archaeplastida. The cyanobiont is displayed in green with an emphasis on the inner (solid green line) selective membrane. The outer membrane is represented by a solid blue line and is thought to be much less selective and permeable to most metabolites including sugar-nucleotides. Enzymes are coloured with respect to their phylogenetic origin as bacterial (blue), host (orange), or unknown (grey) with the following symbols: AGPase, ADP-glucose pyrophosphorylase; SS-ADP, ADP-Glc requiring soluble starch synthase; SS-UDP, UDP-Glc requiring soluble starch synthase; Glg, glycogenin; Pho, phosphorylase; BE, branching enzyme; ISO, isoamylase; D-enz, disproportionating enzyme (α-1,4 glucanotransferase); DPE2, amylomaltase; BAM, β-amylase; GBSS, granule-bond starch synthase; GWD, glucan water dikinase; PWD, phosphoglucan water dikinase; NST, nucleotide sugar translocator. The carbon fixed by photosynthesis enters the Calvin cycle in the cyanobiont. The carbon committed to storage flows selectively through ADP-glucose pyrophosphorylase. The ADP-glucose accumulates in the cyanobiont to high levels because of the absence of storage polysaccharides within the latter. The ADP-Glc is exported in exchange for AMP and orthophosphate by an NST recruited from the host endomembrane system as described in Colleoni *et al.* (2010). Within the cytosol, the exported photosynthate is polymerized into starch thanks to the ADP-glucose-specific soluble starch synthase. The carbon is then mobilized through a completely host-driven polysaccharide degradation machinery following host needs. This machinery involves the GWD–PWD starch phosphorylation pathway followed by β-amylase digestion and maltose degradation. Degraded starch is symbolized by a smaller eroded and pierced granule.

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 Fig. 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 that part of bacterial metabolism which is committed to storage, that is to escape cyanobacterial metabolism and become at least temporarily unavailable in solid starch. Translocating ADP-glucose out of the cyanobiont and having the carbon stored in the cytosol will make very little difference. In addition, as ADP-glucose enters the cytosol it has virtually no impact on host metabolism since this glycosyl-nucleotide is not produced nor recognized by eukaryotes. Because of this absence of recognition, ADP-glucose's only fate will be to be used through a bacterial 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 the outer chains of storage polysaccharides and release ADP. ADP will re-enter the cyanobiont through the exchange reaction mediated through the ADP-glucose NST working on the cyanobiont's inner membrane. This will either be in the form of ADP or AMP since most known NSTs are documented to exchange the glycosylnucleotide with the corresponding nucleotide monophosphate. However, in the case of the exchange with AMP, the question of the missing orthophosphate's fate remains to be fully addressed. The cycle indeed needs to be balanced with respect to both purine nucleotides and orthophosphate to result 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 in glycogen or starch will efficiently buffer demand and supply for carbon. Indeed the pathway of storage polysaccharide breakdown is entirely under the 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 successfully to establish the endosymbiotic link. This would have initially led 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. 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 would have remained important since the host probably adopted first a mixotrophic cycle involving both phagocytosis and the use of photosynthate provided by its endosymbiont. Phagocytosis is expected to impact specifically and dramatically storage polysaccharide metabolism. Indeed large amounts of carbon need to be stored in a relatively short time in an otherwise oligotrophic environment. It is thought that the host has evolved complex regulatory mechanisms impacting the UDP-glucose specific glycogen synthase and triggering glycogen accumulation under such circumstances. The phagotroph model heterotrophic dinoflagellate *Crypthecodinium cohnii* was recently demonstrated, unlike other non-phagotrophic microorganisms such as yeasts or bacteria, to accumulate glycogen during the log phase as soon as the substrate becomes available and not at the transition between the log and stationary phases (Deschamps *et al.*, 2008*d*). This behaviour was interpreted as a response to nutrient bursts mimicking the consequences of prey ingestion. This useful physiological function and the concomitant regulation were 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).

Discovering the missing link of eukaryotic photosynthesis

When the flux displayed in Fig. 5 was first proposed, there was 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, and 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 monophosphate 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 Fig. 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) 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 K_m 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 outcompete 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 (see below).

The results reported by Colleoni *et al.* (2010) do not imply that the ancestors of the corresponding GDP-mannose translocators defined the hypothetical missing link for the establishment of eukaryotic photosynthesis. In fact, probably any NST3 purine nucleotide sugar transporter would have 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 Archaeplastida 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.

Early targeting of the ADP-glucose translocator to the cyanobiont inner membrane

One of the essential requirements to be met for the first carbon translocators was that they should be correctly targeted to the cyanobiont's inner membrane. Most researchers agree that the present sophisticated machinery of protein targeting to plastids took time to evolve. Yet a novel set of proteins probably needed to find its way to the endosymbiont's inner membrane to tap photosynthate into the periplasm. The latter would then leak to the cytosol, the outer membrane being much less selective.

Several options can be considered at this stage. However, one should also consider, in addition, that the the plant TPT (triose phosphate translocator) localizes to mitochondrial membranes when expressed in yeast without its transit peptide (Loddenkötter *et al.*, 1993). This could suggest that whatever mechanism or built in structural features the TPT uses to localize to mitochondrial membranes, these may also have been working for the localization of the TPT ancestors to the cyanobiont's inner membrane.

There is presently no consensus on the nature of early protein targeting to the cyanobiont or the cyanobiont inner membrane. Some authors support vesicular transport

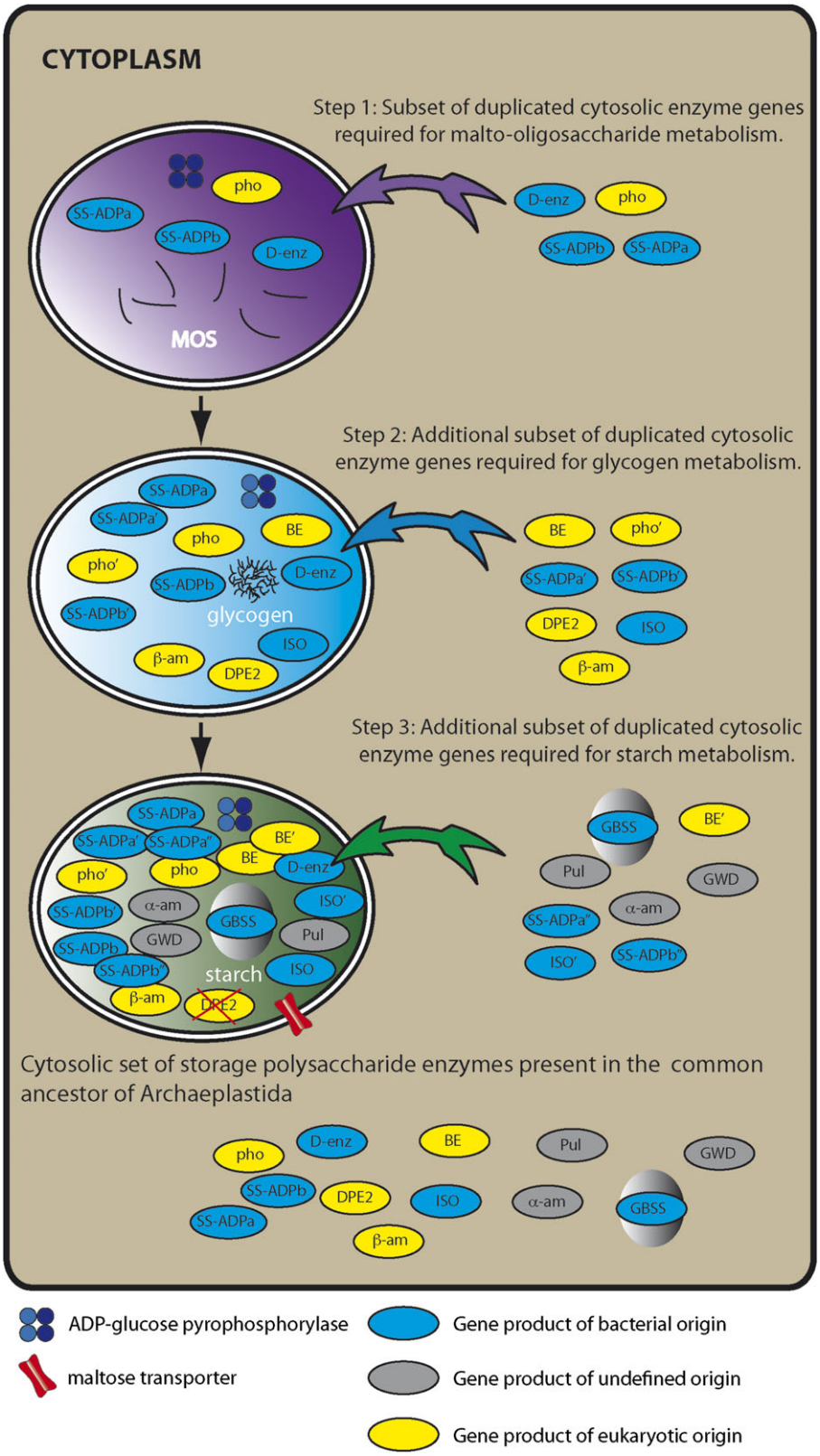

Fig. 6. Rewiring of starch metabolism in the emerging Chloroplastida. As detailed in Deschamps *et al.* (2008*c*), evolution of novel chlorophyll-rich LHCs in Chloroplastida possibly increased the photooxidative stresses experienced by the evolving chloroplasts. To obviate some of these stresses the presence of storage polysaccharides or oligosaccharides within the plastid became desirable. Because the rewiring of a 12 component network in a single step defines a very unlikely event we propose that this evolution involved two workable intermediate stages defined by the accumulation of a small pool of maltooligosaccharides and a larger pool of glycogen which co-evolved with the increasing complexity and requirements of the LHC antennae (symbolized by the purple phycobilisome

(Villarejo *et al.*, 2005) while others propose an early origin for a simplified version of the TIC-TOC import machinery (Bodyl *et al.*, 2009). Indeed, to target a protein to the cyanobiont inner membrane through a primitive TOC system would have only required an accidental duplication of a suitable NST gene that was accompanied by the substitution of the signal sequence by a pre-existing mitochondrial leader-like sequence. The primitive TOC system could have been defined by the sole cyanobacterial Omp85 protein whose orientation was reversed in the endosymbiont's outer membrane, as initially suggested by Bolter *et al.* (1998) and reviewed by Bodyl *et al.* (2009). This reversion would have been readily generated by the cytosolic expression of a copy of Omp85 that was transferred to the host nucleus (Bodyl *et al.*, 2009). Despite the fact that the recruited NST originally came from the host endomembrane system, we favour the primitive TOC hypothesis. This seems to be suggested by the fact that all extant pTPs are nowadays targeted by the TIC-TOC machinery. This primitive version of TOC would have been sufficient to direct a protein to the cyanobiont periplasm or the inner membrane. This suggests that both recruitment of the endomembrane NST and inversion of Omp85 were very early events, since both of these were simultaneously required to initiate endosymbiosis. It also implies that gene duplication of the recruited NST was a prerequisite for its use on the cyanobiont's inner membrane, since it had to be transported by an altogether different route than that taken by the original transporter targeted to the ER and Golgi.

Another possibility would be that the transporter was targeted to plastids through a selective modification of vesicular transport, a mechanism which is considered by some to have defined the first kind of transport machinery (Villarejo *et al.*, 2005). In this respect, the ER retention signal found in many pPT proteins (Martinez-Duncker *et al.*, 2003; Colleoni *et al.*, 2010) might have played a specific role to avoid the transporter being directed to the Golgi which defines the home of all purine nucleotide sugar transporters presently documented. However, in this case, the duplicated copies of the ancestral transporter would have switched later to transport through the TIC-TOC protein import machinery.

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

The simple cytosolic pathway depicted in Fig. 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 observed in secondary endosymbiosis lineages such as the apicomplexan parasites (Coppin *et al.*, 2005) and is suspected to be 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.*, 2008*b*). 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 counterselected. The means by which an entire pathway becomes targeted to plastids are 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

containing plastid, the blue-green intermediate with both LHCs and phycobilisomes and the green chloroplast). Each stage is thought to have lasted long enough for the rewired enzymes to have become optimized with respect to the type of products they were making. The movement from one stage to another (from oligosaccharides to glycogen or from glycogen to starch) can be initiated through a single (for instance BE to generate glycogen from the pre-existing MOS metabolism machinery) or at most two gene duplications (for instance starch synthase and phosphorylase to intiate MOS metabolism from ADP-glucose in the cyanobiont) with concomitant acquisition of transit peptides. The nature of the network desirable at each stage is deduced from that exemplified in bacteria for maltooligosaccharides and in amoebas for glycogen (although the elongation enzyme would be defined by the ADP-glucose specific starch sytnhases). The cytosolic network of the archaeplastidal ancestor is displayed at the bottom of the figure with enzymes of bacterial and host origin respectively in blue and yellow. This cytosolic network will be finally lost (with the exception of cytosolic phosporylase and DPE2) at the very last stage of the evolution process when the maltose transporter appeared (in red) and the plastidial DPE2 was lost. The rationale for enzyme subfunctionalization is detailed in the text.

Table 2. The identity and number of isoforms of the starch metabolism network rewired to plastids at each stage (MOS, glycogen, and plastidial starch) of the reconstruction are indicated in blue

To obtain the final theoretical numbers displayed in red, the number of plastidial enzymes (in blue) transferred at each stage must be summed together and with the number of enzymes of the ancestral cytosolic network (in orange) taking into account the selective losses that occurred mostly when cytosolic starch was finally lost. Gene losses are symbolized by the minus symbol followed by the number of genes lost that is coloured according to the cytosolic (orange) or plastidial (blue) localization of the product lost. Innovations are symbolized by the + symbol followed in bold by a number that is coloured according to its cytosolic (orange) or plastidial localization. The identity of the enzymes transferred at each stage corresponds to the catalytic activities that are known to be at work for MOS metabolism in bacteria, for glycogen metabolism in eukaryotes, and for starch metabolism. When several isoforms of the ancestral cytosolic network correspond to an activity required at a particular stage, it is assumed that the genes corresponding to all isoforms will be duplicated and their products targeted to the organelle. GWD, dpe2, and β-amylase are considered as enzymes of the eukaryote glycogen metabolism network either on the basis of their existence in extant non-opistokont eukaryotes or on the basis of the existence of a glycogen phosphorylation pathway in eukaryotes.

Activity	Ancient cytosolic Starch network	MOS network	Glycogen network	Plastidial starch network	Theoretical distribution	*O. taurii* network
ADP-glucose PPase	1	0	0	0	1	2
Starch synthase	2	2	2	2–2	6	5
GBSS I	1	0	0	1–1	1	1
Branching enzyme	1	0	1	1–1	2	2
Isoamylase	1	0	1	1–1	2	3
Phosphorylase	1	1	1	0	3	2
Glucanotransferase	1	1	0	0–1	1	1
Transglucosidase	1	0	1	−1	1	1
b-amylase	1	0	1	1–1	2	2
GWD	+1	0	1	1–1	2	2
MEX	0	0	0	+1	1	1

a few distinct steps where a defined balance of enzyme activities is, in addition, not critical may be rewired in this fashion. However, we 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 a linear fashion (Ball and Morell, 2003). Polysaccharide debranching also produces substrates for further elongation and branching and debranching (Ball and Morell, 2003). In addition, the balance of enzyme activities is 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.*, 2004*b*; Hennen-Bierwagen *et al.*, 2008).

Everything we know concerning starch metabolism tells us that the simultaneous acquisition of a transit peptide by all the required enzymes and the achievement of the proper activity balance in one step are 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.* (2008*c*) 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 progressively 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-crystalline 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.*, 2008*b*). 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. A bioinformatic analysis carried out in six green alga genomes demonstrated that, in the earliest diverging prasinophyte algae (the mamelliales), the full complexity of starch metabolism witnessed in higher plants is already present (Deschamps *et al.*, 2008*b*). Deschamps *et al.* (2008*c*) 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). The flux to chlorophyll is suspected to have increased in Chloroplastida which have substituted the phycobilisomes by chlorophyll-containing light-harvesting complexes. Hence the need for a plastidial source of ATP at night would have increased (Deschamps *et al.*, 2008*c*).

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.* (2008*b, c*) argue that the return to the chloroplast and starch metabolism complexity are mechanistically linked. The problems inherent to a rewiring of a whole suite of enzymes have been outlined above. However, such a return can be imagined if the whole process happened in sequential steps. These entail two fully functional intermediate stages involving first MOS production then glycogen synthesis (Deschamps *et al.*, 2008*c*). 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 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 to subject the rewired enzymes to mutations that would have optimized their activities with respect to the products they now 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 then 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. When a selection pressure reappeared to increase the plastidial storage carbohydrate pools further 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. It would probably have taken longer to accumulate mutations in the aforementioned MOS synthase in order 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 led 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 a reasoning applies to all enzymes of starch metabolism and the reader is referred to a recent review for further details (Deschamps *et al.*, 2008*c*).

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 six and two, since the initial cytosolic network contained, respectively, two and one isoforms 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). The establishment of starch in the plastid therefore 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. One must not forget that in addition to the process of duplication and enzyme subfunctionalization pertaining to the rewiring mechanism that we propose, other reasons may have prompted a minor proportion of enzyme subfunctionalizations. For instance, isoamylase was evidently duplicated and subfunctionalized in the cytosolic starch metabolism network independently of the rewiring mechanism outlined above. Similarly ADP-glucose pyrophosphorylase was also duplicated and subfunctionalized into large and small subunits despite its maintenance in the cyanobiont. In these two cases, however, strong functional constraints would have favoured such processes. For instance, cytosolic subfunctionalized isoamylases were required as debranching enzymes played essential functions during both synthesis and degradation of starch polysaccharides. Occasional duplications and subfunctionalizations are indeed found in many pathways. It is argued, however, that starch metabolism in Chloroplastida differ from all other pathways by the sheer magnitude of this phenomenon. Indeed this pathway stands out as the only one displaying this type of complexity among prasinophytes. In addition, there is general good agreement between the theoretical isoform numbers deduced in Table 2 from the reconstruction mechanism that is proposed here and those

observed in the Chloroplastida starch metabolism network. Such an unexpected agreement supports our initial proposal that the ancient starch metabolism pathway was indeed cytosolic.

Amylose synthesis and the presence of the pyrenoidal starch sheath

Many green and red algae and some of the secondary endosymbiosis lines contain pyrenoids in the plastid stroma. Pyrenoids are insoluble Rubisco aggregates that are seemingly involved in the CO_2 concentration mechanism (CCM) operating in these cells (Süss *et al.*, 1995; for a review see Giordano *et al.*, 2005). In most green algae and in hornworts, in porphyridiales red algae and in cryptophytes, the pyrenoid is surrounded by a starch sheath (named pyrenoidal starch) in addition to dispersed granules found either within the plastid of the green algae (often named 'stromal starch'), in the periplastidial space of the cryptophytes or in the cytosol of the porphyridiales red algae (Izumo *et al.*, 2007).

In the case of red algae and cryptophytes, the starch remains separated from the pyrenoid by the two rhodoplast membranes or the two inner membranes from the cryptophyte secondary plastids, respectively. The shape of such granules is typically distorted and results from the physical interaction between the starch sheath and the underlying pyrenoid structure. The growth of the starch sheath was shown to be induced in *Chlamydomonas* when switching from high to low CO_2 while stromal starch synthesis is favoured in the presence of high CO_2 and (or) acetate (Villarejo *et al.*, 1996). It has been demonstrated through the use of starch defective mutants in *Chlamydomonas* and *Chlorella* that starch is not required for normal induction of the CCM (Plumed *et al.*, 1996; Villarejo *et al.*, 1996). Hence the starch sheath most probably comes as a consequence of localized CO_2 fixation by pyrenoidal Rubisco yielding concentration gradients of important metabolites such as 3-PGA and hexoses rather than being an integral part of the CCM required for pyrenoid confinement and function.

The presence of amylose in starch could be an important feature required for the synthesis of an abundant pyrenoidal starch sheath. Indeed all photosynthetic eukaryotes that synthesize a pyrenoidal starch sheath contain amylose. The porphyridiales which define the only red algae whose floridean starch is synthesized in close association with the pyrenoid also defines the only group that contains amylose within the Rhodophyceae. Delrue *et al.* (1992) noted that *Chlamydomonas* mutants defective for GBSSI and, consequently, amylose synthesis displayed an altered pyrenoidal starch sheath. Because GBSS is responsible for building long chains within both amylose and amylopectin it is possible that these chains allows the starch granules to be more 'plastic' and adopt those shapes that are compatible with a close physical association to the pyrenoid. SSIII might play a similar and partly redundant function in Chloroplastida (Maddelein *et al.*, 1994). Interestingly both SSIII and GBSSI are induced when switching to low CO_2 in *Chlamydomonas* and *Chlorella* (Miura *et al.*, 2004; Oyama *et al.*, 2006). Hence it would be tempting to suggest that amylose and thus GBSSI was maintained in starch-storing algae for the purpose of pyrenoidal starch synthesis.

It must be stressed, however, that neither subgroup V cyanobacteria nor glaucophytes nor terrestrial plants other than hornworts are reported to contain pyrenoids; yet these organisms have maintained amylose and GBSSI. Nevertheless the maintenance of some 'plasticity' in starch granule morphogenesis which is exemplified by the pyrenoidal starch sheath might define one of the major selection pressures for the maintenance of amylose in Archaeplastida. This could be important in various other instances where substrate concentration gradients could lead to localized amylose synthesis.

Chlamydial genes in the starch pathway: evidence for ancient lateral gene transfers

One of the most surprising recent findings of phylogenetic inference consist in the unexpected relationship existing between the genes of Archaeplastida and Chlamydiae (reviewed in Horn, 2008). Chlamydiae define obligate intracellular gram-negative bacterial pathogens initially 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 enters 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 recently been identified, often showing greater metabolic capabilities (and genomes) than the human pathogens but with similar obligate intracellular life cycles typified by the same basic energy parasitism (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 39–55 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 Chlamydiae 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 analysing the phylogeny of the NTT transporters (Linka *et al.*, 2003). These transporters have been shown 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 indeed have been 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 support an LGT of an enzyme related to extant chloroplastidal SSIII-SSIV from Protochlamydia to plants (Moustafa *et al.*, 2008). However, in this case, only the ancestral Chlamydiae display convincing evidence for this transfer and the more classic animal pathogen-type of Chlamydiae proved to contain an unrelated enzyme.

One of the most parsimonious 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). During integration of the cyanobiont the recurrent presence of chlamydial parasites expanded the repertoire of useful LGTs, thereby facilitating establishment of endosymbiosis. For instance, a chlamydial parasite donated its ATP/ADP translocase enabling the cyanobiont to withstand in darkness the negative consequences of the loss of its storage polysaccharides. In addition, *Chlamydia* parasites may have been the source of the Archaeplastida isoamylases turning the pool of cytosolic glycogen into starch, thereby further increasing photosynthate export.

Present-day plants and algae are not known to be subjected to infection by *Chlamydia*. Therefore, this particular influence of Chlamydiae-related parasites was terminated at a very early stage of Archaeplastida evolution, possibly before the 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 parasites which are 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*.

The finding of isoamylase as a bona fide chlamydial gene has important consequences on our understanding of the evolution of starch metabolism in Archaeplastida. Indeed the finding of true starch and of GBSSI in unicellular diazotrophic cyanobacteria would tend to suggest that the cyanobiont might have donated the critical genes required for turning glycogen into starch. While this remains valid when one envisions the synthesis of amylose by GBSSI, the switch of hydrosoluble glycogen to starch in the archaeplastidal cytosol did not result from an LGT from a cyanobacterial debranching enzyme gene but rather by that of the corresponding chlamydial gene. This, in turn, could mean that semi-crystalline polymers may be synthesized by a different pathway in cyanobacteria. In any case this argues that semi-crystalline storage polysaccharides evolved independently in cyanobacteria and Archaeplastida. A critical evaluation of the phylogeny of debranching enzymes in apicomplexa parasites further suggests that the archaeplastidal isoamylase was not the source for the debranching enzyme gene in starch accumulating secondary endosymbiosis lineages (Coppin *et al.*, 2005; Deschamps *et al.*, 2008*c*). There again the switch from glycogen to starch may have occurred independently in the cytosol or periplast of these organisms.

General conclusion: starch metabolism evolution gives an unexpected window on those fundamental events that shaped the Archaeplastida

When we first pondered the evolution of the starch pathway in Archaeplastida we were faced with the problem of fitting a dual substrate storage polysaccharide metabolism relying on two types of glycosyl-nucleotides (ADP-glucose and UDP-glucose) in a single compartment defined by the common ancestor's cytosol. This could only be achieved and make any kind of physiological sense if ADP-glucose was exported from the cyanobiont. We thus stumbled on what actually could define the major metabolic link of plastid endosymbiosis as the only plausible solution to our compartmentalization problem. In that sense, examining the origin of starch metabolism in Archaeplastida gave us an unexpected window on the critical early steps that led to the establishment of the cyanobiont. The proposal that we make of an ancient cytosolic starch metabolism network also gives us the reasons underlying the otherwise unexplainable complexity of starch metabolism that characterize selectively the Chloroplastida compared with the Rhodophyceae and Glaucophyta. The history of the starch metabolism network in vascular plants is thus highly complex and fascinating. When going through this complex history it appears that synthesis of semi-crystalline starch-like polysaccharides evolved several times independently from glycogen metabolism: it appeared in unicellular diazotrophic cyanobacteria, it evolved in the cytosol of the Archaeplastida through the use of enzymes which are of chlamydial rather than cyanobacterial origin and it appeared again in the cytoplasm of the common ancestor of dinoflagellates and apicomplexa. Hence the fact that the cyanobiont ancestor was a starch accumulator may not have been a requirement for the evolution of semi-crystalline

polysaccharides per se in Archaeplastida. Nevertheless, there is one important component of starch which required the presence of such ancestors. Amylose is, in all cases, known to depend on the presence of GBSSI, an enzyme which, apart from Archaeplastida and a few secondary endosymbiosis derivatives, is only found to date in such unicellular diazotrophic starch-accumulating cyanobacteria. Because GBSSI has been proven to be significantly active only when associated with semi-crystalline amylopectin packaged into solid starch granules, the cyanobiont ancestors will have had to have been starch accumulators.

Another consequence from the complex history of starch metabolism is that the function of an enzyme seems to be just as much dependent on the historical record of the network to which this enzyme belongs as on its catalytic activity per se. A deep understanding of starch metabolism will require an understanding of the function of an enzyme through the entire evolutionary process: that is in cyanobacteria or chlamydia 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 as well as their secondary endosymbiosis derivatives.

Acknowledgements

This work was performed in the Unité de Glycobiologie Structurale et Fonctionnelle funded by the Region Nord Pas de Calais, the European Union, the French Ministry of Education, the CNRS and ANR grant 'starchevol'. In memory of Jean Montreuil recently deceased who fathered the field of Glycobiology in France.

References

Adl SM, Simpson AG, Farmer MA, *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*. *Journal of Bacteriology* **188,** 5266–5272.

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

Baldauf SL. 2003. The deep roots of eukaryotes. *Science* **300,** 1701–1703.

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, Visser R. 1998. Progress in understanding the biosynthesis of amylose. *Trends in Plant Science* **3,** 462–467.

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

Ballicora MA, Iglesias AA, 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.

Bodyl A, Mackiewicz P, Stiller JW. 2009. Early steps in plastid evolution: current ideas and controversies. *Bioessays* **31,** 1219–1232.

Bolter B, Soll J, Schulz A, Hinnah S, Wagner R. 1998. Origin of a chloroplast protein importer. *Proceedings of the National Academy of Sciences, USA* **95,** 15831–15836.

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–D238.

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, Biological Sciences* **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* **27,** 2691–2701.

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, *et al*. 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, *et al*. 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, USA* **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.

Delrue B, Fontaine T, Routier F, Decq A, Wieruszeski JM, van den Koornhuyse N, Maddelein ML, 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. *Journal of Bacteriology* **174,** 3612–3620.

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, USA* **103,** 11647–11652.

Deschamps P, Colleoni C, Nakamura Y, *et al*. 2008*a*. Metabolic symbiosis and the birth of the plant kingdom. *Molecular Biology and Evolution* **25,** 536–548.

Deschamps P, Guillebeault D, Devassine J, *et al*. 2008*d*. The heterotrophic dinoflagellate *Crypthecodinium cohnii* defines a model genetic system to investigate cytoplasmic starch synthesis. *Eukaryotic Cell* **7,** 247–257.

Deschamps P, Haferkamp I, Dauvillee D, *et al*. 2006. Nature of the periplastidial pathway of starch synthesis in the cryptophyte *Guillardia theta*. *Eukaryotic Cell* **5,** 954–963.

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, 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.

Edner C, Li J, Albrecht T, *et al*. 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 JA, Coyne RS, Wu M, *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 EJ. 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.

François J, Parrou JL. 2001. Reserve carbohydrates metabolism in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiology Reviews* **25,** 125–145.

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.

Giordano M, Beardall J, Raven J. 2005. CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annual Reviews of Plant Biology* **56,** 99–131.

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*. *Journal of General and Applied Microbiology* **50,** 289–295.

Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AG, Roger AJ. 2009. Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic 'supergroups'. *Proceedings of the National Academy of Sciences, USA* **106,** 3859–3864.

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

Horn M, Collingro A, Schmitz-Esser S, *et al*. 2004. Illuminating the evolutionary history of Chlamydiae. *Science* **304,** 728–730.

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

Izumo A, Fujiwara S, Oyama Y, Satoh A, Fujita N, Nakamura Y, Tsuzuki M. 2007. Physicochemical properties of starch in *Chlorella*

change depending on the CO_2 concentration during growth: comparison of structure and properties of pyrenoid and stroma starch. *Plant Science* **172,** 1138–1147.

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

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

Koonin EV. 2010. The origin and early evolution of eukaryotes in the light of phylogenomics. *Genome Biology* **11,** 209–221.

Kotting O, Santelia D, Edner C, *et al*. 2009. STARCH-EXCESS4 is a laforin-like phosphoglucan phosphatase required for starch degradation in *Arabidopsis thaliana*. The *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: Biological Sciences* **302,** 3–11.

Lin T-P, Caspar T, Somerville C, Preiss J. 1988. Isolation and characterisation of a starchless mutant of *Arabidopsis thaliana* (L.) Heynh. lacking ADP-glucose pyrophosphorylase activity. *Plant Physiology* **86,** 1131–1135.

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, USA* **90,** 2155–2159.

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

Maddelein ML, Libessart N, Bellanger F, Delrue B, D'Hulst C, van den Koornhuyse N, Fontaine T, Wieruszeski JM, Decq A, Ball S. 1994. Toward an understanding of the biogenesis of the starch granule. Determination of granule-bound and soluble starch synthase functions in amylopectin synthesis. *Journal of Biological Chemistry* **269,** 25150–25157.

Martin W. 2010. Evolutinary origins of metabolic compartmentalization in eukaryotes. *Philosophical Transactions of the Royal Society B: Biological Sciences* **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, *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, 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.

Miura K, Yamano T, Yoshioka S, *et al*. 2004. Expression profiling-based identification of CO_2-responsive genes regulated by CCM1 controlling a carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiology* **135,** 1595–1607.

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. *The 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, *et al*. 2005. Some cyanobacteria synthesize semi-amylopectin type α-polyglucans instead of glycogen. *Plant and 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.

Oyama Y, Izumo A, Fujiwara S, Shimonaga T, Nakamura Y, Tsuzuki M. 2006. Granule-bound starch synthase cDNA in *Chlorella kessleri* 11h: cloning and regulation of expression by CO_2 concentration. *Planta* **224,** 646–654.

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, *et al*. 2008. The pathway of starch synthesis in the model glaucophyte *Cyanophora paradoxa*. *Eukaryotic Cell* **7,** 247–257.

Plumed MP, Villarejo A, de los Rios A, Garcia-Reina G, Ramazanov Z. 1996. The CO_2-concentrating mechanism in a starchless mutant of the green unicellular alga *Chlorella pyrenoidosa*. *Planta* **200,** 28–31.

Pombert J-F, Keeling PJ. 2010. The mitochondrial genome of the entomoparasitic green alga *Helicosporidium*. *PLoS ONE* **5,** e8954.

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

Ral JP, Derelle E, Ferraz C, *et al*. 2004. Starch division and partitioning a mechanism for granule propagation and maintenance in the picophytoplanktonic green alga *Ostreococcus tauri*. *Plant Physiology* **136,** 3333–3340.

Raven JA, Beardall J, Flynn KJ, Maberly SC. 2009. Phagotrophy in the origins of photosynthesis in eukaryotes and as a complementary mode of nutrition in phototrophs: relation to Darwin's insectivorous plants. *Journal of Experimental Botany* **60,** 3975–3987.

Reinhold T, Alawady A, Grimm B, *et al*. 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, USA* **99,** 7166–7171.

Roach PJ. 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.

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. *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.

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, *et al*. 2008. Variations in storage α-polyglucans of the Porphyridiales (Bangiophycidae, Rhodophyta). *Plant and Cell Physiology* **49,** 103–116.

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.

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, USA* **100,** 8828–8833.

Süss K-H, Prokhorenko I, Adler K. 1995. *In situ* association of Calvin cycle enzymes, ribulose-1,5-bisphosphate carboxylase/oxygenase activase, ferredoxin-$NADP^+$ reductase, and nitrite reductase with thylakoid and pyrenoid membranes of *Chlamydomonas reinhardtii* chloroplasts as revealed by immunoelectron microscopy. *Plant Physiology* **107,** 1387–1397.

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*, *USA* **100,** 10659–10663.

Villarejo A, Buren S, Larsson S, *et al*. 2005. Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nature Cell Biology* **7,** 1224–1231.

Villarejo A, Martínez F, Plumed MP, Ramazanov Z. 1996. The induction of the CO_2 concentrating mechanism in a starch-less mutant of *Chlamydomonas reinhardtii*. *Physiologia Plantarum* **98,** 798–802.

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-aativated protein kinase, and the cyclin-dependent kinase Pho85p. *Molecular and Cellular Biology* **21,** 5742–5752.

Wattebled F, Dong Y, Dumez S, *et al*. 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.

Wilson WA, Roach PJ, Montero M, Baroja-Fernández E, Muñoz FJ, Eydallin G, Viale AM, Pozueta-Romero J. 2010. Regulation of glycogen metabolism in yeast and bacteria. *FEMS Microbiology Reviews* **34,** 952–985.

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 α-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. *Molecular Biology and Evolution* **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. *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. *The Plant Cell* **10,** 1699–1712.

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