

Evolution of Storage Polysaccharide Metabolism in Archaeplastida Opens an Unexpected Window on the Molecular Mechanisms That Drove Plastid Endosymbiosis

S.G. Ball

Abstract Plastid endosymbiosis was selected through the establishment of a biochemical link between the disconnected metabolic networks of the cyanobiont and its eukaryote host. This link is likely to have consisted of the efflux of photosynthetic carbon from the bacterial symbiont to the cytosol of the eukaryote. Storage molecules are suspected to have played a pivotal role in the establishment of such a flux. The latter provided an immediate opportunity to feed carbon upon a supply dictated by cyanobacterial metabolism while allowing the host to tap these resources upon demand according to its own regulatory circuits. The presence of the stores thus buffered the disconnected and unrelated sources and sink pathways of photosynthetic carbon metabolism during the early stages of plastid endosymbiosis. Comparisons of extant biochemical networks explaining storage polysaccharide metabolism in the three lineages that emerged after plastid endosymbiosis have enabled the reconstruction of the simplest hypothetical ancient network. The latter possibly consisted of the export of photosynthate from the cyanobiont in the form of the bacteria-specific metabolite ADP-glucose and the polymerization of the latter in the host cytosol through an ADP-glucose-specific glucan synthase. Neither the required ADP-glucose transporter nor the glucan synthase can be suspected to have been encoded by the cyanobacterial or host genomes prior to endosymbiosis. Nevertheless these critical components were required to trigger the event. The possible origin of these two key proteins is reviewed.

S.G. Ball (✉)

Unité de Glycobiologie Structurale et Fonctionnelle, Université des Sciences et Technologies de Lille, UMR8576 du CNRS, Lille, France

e-mail: Steven.Ball@univ-lille1.fr

Contents

Introduction	112
Storage: The Metabolic Answer to Connect Unrelated Biochemical Networks	113
The Comparative Biochemistry of Starch and Glycogen Metabolism	113
Starch Metabolism in Archaeplastida	116
Reconstruction of the Ancient Network of Starch Metabolism	121
Incompatibility of Phototrophy and Diazotrophy in Photosynthetic Eukaryotes	124
The ADP-Glucose Connection	127
ADP-Glucose Transport and Glucan Polymerization	129
References	131

Introduction

Oxygenic photosynthesis is a remarkably complex mechanism that evolved in cyanobacteria and durably changed the chemistry of the atmosphere through the massive release of oxygen from water. This very efficient mechanism of tapping energy from the sun to reduce carbon dioxide to organic components requires a vast number of proteins interacting in large complexes finely tuned to avoid the production of toxic waste. The complexity of the process made the progressive introgression to non-cyanobacterial taxa of the hundreds of genes required to achieve it an impossible task. Yet oxygenic photosynthesis was transmitted to eukaryotes allowing primary producers to benefit from the infinite possibilities offered either by eukaryotic multicellularity programs or mobility or by the ability of some of these organisms to colonize environments inaccessible to cyanobacteria. Later on, eukaryotic photosynthesizers were able to maximize primary production by the building of tall rigid structures, thus paving the way for the bloom of animal life. But how did eukaryotes succeed where non-cyanobacterial prokaryotes failed? How did they capture the hundreds of required cyanobacterial genes? The evolution of phagotrophy in early eukaryotes offered a unique opportunity for plant ancestors to capture cyanobacteria as prey (Raven et al. 2009). It is plausible that aborted phagotrophic events enabled stable symbiotic interactions between prey and predator. A particular eukaryotic phagotroph eventually established a specific cyanobacterial lineage as endosymbiont and went down the path of metabolic integration of the cyanobiont into a true cellular organelle: the plastid (for a general review, see Chan et al. 2011). This review focuses on those molecular events that established the initial carbon flux of photosynthesis at endosymbiosis.

Storage: The Metabolic Answer to Connect Unrelated Biochemical Networks

If we hypothesize that plastid endosymbiosis resulted from an aborted phagocytosis, then there is no reason to suppose the preexistence of any molecular connection between the ingested prey and its predator. If we further assume that plastid endosymbiosis was selected because it provided photosynthate to heterotrophic eukaryotes, then we can appreciate the magnitude of the problems facing the partners of plastid endosymbiosis. An optimized metabolic flux exporting carbon from the cyanobiont to its host had to be set up at the very beginning of the process. There was very little time for innovations and no time for adaptation to the problem of unsynchronized demand and supply of carbon. The host had no way of signaling the cyanobiont that it needed carbon nor could it be signaled that carbon could be made accessible by the cyanobiont. Yet the flux came directly under fierce natural selection and was responsible for the initial success of the organisms that had achieved endosymbiosis. How could this be done? It is unlikely that tapping just any metabolite out of the cyanobiont would be without consequences on biochemical networks that have undergone millions year of selection to work in an optimized fashion. Moreover it is equally unlikely that a sudden burst of metabolite in the cytosol would be neutral with respect to the highly coordinated cytosolic networks. An obvious solution to this problem would be to store the carbon in a form readily accessible to the host through catabolic networks defined by preexisting and thus optimized host catabolism responding solely to host needs. Because the size of the stored carbon pools was likely to vary widely as a function of unsynchronized carbon demand and supply the need emerged for a pool having little if any impact on host osmolarity. Storage polysaccharide metabolism therefore defines an obvious candidate for the establishment of the first connection between the partners of endosymbiosis and examining the evolution of this pathway in Archaeplastida and their cyanobacterial and eukaryote ancestors may provide useful information in this respect. In this review we will propose that plastid endosymbiosis has indeed been established through the forging of a metabolic link making uses of the similarities and differences between the preexisting networks of host and cyanobiont storage polysaccharide metabolism. To fully grasp the issues we will briefly first outline what these similarities and differences were.

The Comparative Biochemistry of Starch and Glycogen Metabolism

Both cyanobacterial and host partners can be safely assumed to have been alpha-glucan accumulators before endosymbiosis (Ball et al. 2011). This can be deduced from the finding of enzymes of starch metabolism in all Archaeplastida that are both of distinctive cyanobacterial phylogenetic origin and of eukaryotic host origin

(Coppin et al. 2005; Patron and Keeling 2005; Deschamps et al. 2008a). In addition all extant free-living cyanobacteria are either glycogen or starch accumulators (Nakamura et al. 2005) and many heterotrophic eukaryote lineages accumulate glycogen (Ball et al. 2011).

Storage α -glucans come in a variety of forms including maltooligosaccharides (MOS), glycogen, and starch (Boos and Shuman 1998; Buléon et al. 1998). Glycogen defines by far the most widespread form of storage polysaccharide found in all three domains of life (eukaryotes archaea and bacteria). Glycogen consists of small hydrosoluble particles with a maximal diameter of 40–50 nm made of glucose linked by α -1,4 glycosidic linkages with 8–12 % α -1,6 branches. Each branch creates a novel chain increasing progressively the density of the glucans (glucose chains) at the periphery of the particle and therefore limiting its size to less than 50 nm in diameter. The glycogen outer chains are readily accessible to enzymes of glycogen synthesis and degradation as if the glucose was in a soluble state. Hence glycogen defines a very dynamic form of storing glucose yet with little osmotic activity while remaining rapidly available to cellular metabolism (for a review of glycogen structure, see Buléon et al. 1998; Shearer and Graham 2002). Starch has the same composition and chemical linkages but a very different physicochemical state. It defines a solid insoluble semicrystalline structure inert osmotically but also unavailable to hydrosoluble enzymes. Polysaccharide crystallization is ensured by an ordered distribution of branches which remain concentrated in certain regions of the polysaccharide (for a detailed review of starch structure, see Buléon et al. 1998). This asymmetric distribution leads to a greater proximity of chains into clusters. This facilitates the alignment of the glucans within a cluster and of the latter with those of other clusters leading to hydrophobic structures that aggregate into insoluble material. These are packaged into a huge solid called the “starch granule.” The major branched fraction of starch is called amylopectin which defines the largest biological molecule known. Amylose, the minor fraction of starch, can be considered as an accessory smaller and linear (with very few branches) side product within starch (Buléon et al. 1998). Indeed, some organisms fail to accumulate this polysaccharide while building normal starch granules. The presence of amylose is due to the only enzyme which seems to be active within the semicrystalline polysaccharide matrix defined by the starch granule (all other enzymes being only active in the soluble state) (for a review of amylose biosynthesis, see Ball et al. 1998). The granule-bound starch synthase thus synthesizes amylose processively within the granule, thereby generating a linear polymer which is protected from the action of the soluble branching enzymes by the presence of the polysaccharide matrix. Starch is only found in Archaeplastida, cryptophytes, alveolates, and a subgroup of particular unicellular diazotrophic cyanobacteria (for a review, see Ball et al. 2011). All evidence points to the evolution of starch from the related glycogen metabolic pathways. The transition from glycogen to starch has happened several times independently: at least once in cyanobacteria, once in the common ancestor of Archaeplastida, and at least once in an alveolate ancestor.

At endosymbiosis two pathways of storage polysaccharide metabolism merged to generate the archaeplastidal starch metabolism network (Coppin et al. 2005;

Patron and Keeling 2005; Deschamps et al. 2008a). These consisted of the pathways of eukaryotic glycogen metabolism and that of cyanophycean starch metabolism. Prokaryotic and eukaryotic pathways of glycogen–starch metabolism both depend on polymerization of glucose through the so-called glycogen (starch) synthases. These glycosyl transferases transfer an activated glucose from either a purine (ADP-glucose) or pyrimidine (UDP-glucose) glucosyl-nucleotide to the nonreducing end of a growing α -1,4-linked chain (for a review of glycogen metabolism in eukaryotes, see Roach 2002; Wilson et al. 2010; for bacteria, see Preiss 1984; Wilson et al. 2010). These purine or pyrimidine nucleotide sugars are synthesized from either ATP or UTP and glucose-1-P through the corresponding nucleotide sugar pyrophosphorylases. Branches are introduced in both bacteria and eukaryotes through hydrolysis of a preexisting α -1,4 linkage within an α -1,4-linked chain and transfer of a segment of chain in α -1,6 position on a neighboring chain. In eukaryotes and bacteria glycogen (starch) phosphorylases are considered to be the major enzymes of glycogen breakdown. These enzymes use orthophosphate to break the outer α -1,4 linkage at the nonreducing end of the glucan chains, thereby releasing glucose-1-P. This energy-efficient mobilization recovers one of the two high-energy phosphate bonds that were used to synthesize the nucleotide sugar substrate from glucose-1-P and either UTP or ATP. Glycogen phosphorylase stops four residues away from the first branch it encounters and cannot break or bypass the α -1,6 linkage. Hence the complete mobilization of the glycogen particle requires further action of the so-called glycogen debranching enzymes. Bacteria use direct debranching enzymes which release the outer glycogen chains in the form of malto-oligosaccharides (MOS) which need to be further degraded by enzymes of MOS metabolism which include maltodextrin phosphorylases and α -1,4 glucanotransferases (disproportionating enzymes or bacterial amyloamylases) (for a review of bacterial MOS metabolism, see Boos and Shuman 1998). Eukaryotes use indirect debranching enzymes which display a complex mode of debranching yielding solely glucose. In addition to phosphorolysis, hydrolysis plays an important role in eukaryotic glycogen breakdown and possibly also in bacterial catabolism. Of particular relevance, eukaryotes with the noticeable exception of fungi and animals often contain an exohydrolase named β -amylase that produces β -maltose from the outer chains of glycogen and which like phosphorylase stops a few residues before the branch (Deschamps et al. 2008a; Ball et al. 2011). The β -maltose is then degraded through an α -1,4 glucanotransferase named dpe2 (or amyloamylase) which transfers glucose from maltose to the outer α -1,4 chains of glycogen or heteroglycans (for a review of dpe2 action in starch and glycogen metabolism, see Fettke et al. 2009). In addition to this eukaryotes may contain glucosidases or other hydrolases in selective compartments such as the lysosome or fungal vacuole where they access glycogen particles that have been redirected through autophagy-like mechanisms (Wang et al. 2001; see Wilson et al. 2010 and Roach 2002 for reviews).

When one compares the glycogen metabolism networks of bacteria and eukaryotes two major differences can thus be found. Most bacteria synthesize glycogen through the bacteria-specific metabolite ADP-glucose while eukaryotes

always use UDP-glucose. The second difference consists in the presence of direct debranching enzymes which, in bacteria only, tie glycogen to malto-oligosaccharide metabolism while eukaryotes seem devoid of cytosolic MOS-degrading enzymes other than the dpe2-type amyloamylases.

ADP-glucose is devoted to the synthesis of glycogen in bacteria while UDP-glucose is a substrate common to many distinct pathways in all living cells. Hence the synthesis of ADP-glucose through ADP-glucose pyrophosphorylase defines the first committed step of glycogen synthesis in bacteria and as such is finely tuned by allosteric regulations while it is glycogen synthase which is regulated in eukaryotes. Of particular relevance to our discussion is the cyanobacterial ADP-glucose pyrophosphorylase (AGPase) (for a review of ADP-glucose pyrophosphorylase structure and function, see Ballicora et al. 2003). In cyanobacteria the enzyme is tightly coupled to photosynthesis through not only its substrates (glucose-1-P and ATP) but also its allosteric activator (3-PGA) and inhibitor (orthophosphate). Hence cyanobacterial glycogen synthesis is finely tuned by the Calvin cycle status and the ATP to Pi ratio.

On the other hand, eukaryotic glycogen synthases and phosphorylases are in addition known to be the subject of activation or inhibition through kinase and phosphatase cascades. Indeed it was by studying the regulation of animal glycogen metabolism that protein kinases were discovered (Krebs 1983).

As we shall see, the above similarities and differences were used to establish the first biochemical connection between the cyanobiont and its host.

Starch Metabolism in Archaeplastida

Starch is usually thought of as a plastidial storage polysaccharide by plant biologists. In fact green plants and algae (Chloroplastida) are the only organisms accumulating storage polysaccharides in this organelle. Red algae and glaucophytes and all starch-storing secondary endosymbiosis derivatives accumulate starch in either the cytosol (alveolates) or the periplast (derived from the red alga cytosol) (reviewed in Ball et al. 2011). Hence plastidial starch metabolism can be considered as the exception rather than the rule. The enzyme network composition of Glaucophyta, Chloroplastida, and Rhodophyceae is summarized in Table 1, which also emphasizes the phylogenetic origin of these enzyme sequences. For comparison we have also listed the extant amoebal and cyanobacterial networks as examples reflecting the possible composition of the ancestral networks present in the future cyanobiont and eukaryotic host before endosymbiosis (see preceding paragraph for details). We have chosen a representative member of a particular group of cyanobacteria that we suspect to descend from the cyanobacterial clade that donated the plastid. This gene-rich diazotrophic yet unicellular group of cyanobacteria was the only clade not ruled out by the synapomorphies examined recently by Gupta (2009) within cyanobacteria. In addition it is the only group reported to contain the GBSSI gene (Deschamps et al. 2008a) which defines one of

Table 1 Storage polysaccharide enzyme networks

Activity	Cyanobacteria (<i>C. walsbyi</i>)	Eukaryotes (<i>E. histolytica</i>)	Chloroplastida (<i>C. tauri</i>)	Rhodophyceae (<i>C. merolae</i>)	Glaucoophyta (<i>C. paradoxa</i>)	Ancestor
ADP-glucose pyrophosphorylase	1	0	3	0	0	1
Starch synthase (ADPG) SSI-SSIV	2	0	3	0	1	1
Starch synthase (ADPG) SSI-SSII	0	0	2	0	0	
Starch (glycogen) synthase (UDPG)	0	1	0	1	2-3 ?	1-2
GBSSI	1	0	1	(0)1	2	1
Branching enzyme	3	1	2	1	3	3
Isoamylase/GlgX	1	0	3	2	3	1
Indirect debranching enzyme	0	1	0	0	1	1
Pullulanase	0	0	1	1	0	N/A
Phosphorylase	2	2	3	1	2	2
Glucanotransferase (dpe1)	1	0	1	0	0	N/A
Transglucosidase (dpe2)	0	2	1	2	3	1
β -amylase	0	4	2	1	1	1
Glucan water dikinase (GWD)	0	0	3	1	2	1
Phosphoglucan water dikinase	0	0	2	-	1	1
sex4-lafolin glucan phosphatase	0	1	1	1	1	1

The storage polysaccharide enzyme networks are displayed for *Entamoeba histolytica* (eukaryotic host glycogen metabolism reference), *Crocospira walsbyi* (representative unicellular cyanobacterium of clade B chosen as the cyanobiont reference), *Ostreococcus tauri* (Chloroplastida), *Cyanidioschyzon merolae* (Rhodophyceae), and *Cyanophora paradoxa* (Glaucoophyta). Little variation in number and type of isoforms occurs within each group and other species of Chloroplastida or Rhodophyceae would have been as well suited as those chosen here. The numbers correspond to the different number of isoforms found for each type of activity. In the case of Rhodophyceae we have listed GBSSI as (0)1 despite its absence in the *C. merolae* genome since this important enzyme of amylose synthesis has been reported in Porphyridiales red algae (Shimonaga et al. 2007). Phylogenetic origin of the different enzymes is represented by the background color: light blue for cyanobacteria, pink for eukaryotes, violet for Chlamydiales, and yellow for diverse proteobacteria. Note that pullulanase is polyphyletic in red and green algae. The reconstruction displayed as “ancestor” makes an assumption of simple vertical inheritance from the common ancestor which holds for all monophyletic enzymes and minimizes the number of isoforms to the minimal number needed to explain the present distribution of genes in the three archaeplastidial lineages. We have represented SSI-SSII with a darker blue background as the latter do not exist in extant cyanobacteria but were probably generated post-endosymbiosis by duplication and diversification from a cyanobacterial GBSSI ancestral gene. Because (continued)

Table 1 (continued)

GWD-PWD works in a background of a preexisting pathway of glycogen phosphorylation in eukaryotes, we consider these novel activities as host derived and represent them with a pink hatched background. The status represented in the ancestor column corresponds to the second stage depicted in Fig. 1b. At the onset of plastid endosymbiosis, it is thought that storage polysaccharides of the host cytosol were not defined by starch as suggested by Table 1 but by glycogen particles (Fig. 1a) as is presently the case for all eukaryotes groups unrelated to plastid endosymbiosis. This implies that at that early stage the domain fusions that generated the PWD-GWD group of enzymes and the mutations of the GlgX bacterial ancestor that transformed this enzyme into an isoamylase had not yet occurred. In turn, the GBSSI gene would have not yet been transferred, thereby postponing the later appearance of the SSI-SSII enzymes

the few cases where a cyanobacterial gene was used within the archaeplastidal starch metabolism network (see below). It also fulfills the requirement relatively to the presumed gene-rich and possibly diazotrophic nature of the ancestor (Deusch et al. 2008). All GBSSI containing cyanobacteria accumulate starch-like polymers and belong to subgroup V according to Honda et al. (1999) within clade B (Gupta 2009). GBSSI in both cyanobacteria and Archaeplastida is an enzyme solely active within the semicrystalline matrix of starch-like structures. In both cases it is responsible for the synthesis of amylose (see above). It displays little or no activity as a soluble enzyme. From these considerations we can deduce that the plastid donor was probably a starch accumulating cyanobacterium. Yet this property is entirely fortuitous as the transition to starch in the archaeplastidal cytosol seems to have occurred independently of this except for the presence of GBSSI.

The three distinct archaeplastidal biochemical starch metabolism networks illustrated in Table 1 essentially differ by two criteria: the first being the nature of the glycosyl nucleotide used for starch synthesis and the second consisting in the level of redundancy of the network (the number of enzyme forms for each step).

First, both Glaucophyta and Rhodophyceae accumulate starch from UDP-glucose in the cytosol (Nyvall et al. 2001; Plancke et al. 2008) while Chloroplastida synthesize plastidial starch from ADP-glucose (Lin et al. 1988; Zabawinski et al. 2001). Second, Rhodophyceae metabolize starch with a mere 11–12 genes while Glaucophyta use over 20 genes and Chloroplastida do so by using a minimum of 30 and often more than 40 genes. For the essential part this increase in complexity reflects an increase in redundancy of enzymes. For instance while only one branching enzyme, one soluble starch synthase, and one β -amylase are found in the rhodophycean network, three branching enzymes, four soluble starch synthases, and a minimum of three β -amylases are found in the green alga network. These redundancies in the green lineage stem from post-endosymbiosis gene duplications. The redundancies witnessed in the Glaucophyta are of a different nature. With the noticeable exception of direct debranching enzymes and dpe2 eukaryotic amylomaltases, most redundancies stem from pre-endosymbiosis duplications and reflect probably a greater diversity of eukaryotic glycogen metabolism enzymes present in the amoeba-like ancestor. In line with this observation, only Glaucophyta still contain the gene coding indirect debranching enzyme that must have been present in this ancestor before it engulfed the cyanobacterium.

Notwithstanding these differences the vast majority of common steps that are listed in Table 1 are controlled by enzymes which display a clear common phylogenetic origin (Coppin et al. 2005; Patron and Keeling 2005; Deschamps et al. 2008a) in agreement with Archaeplastida monophyly (Rodríguez-Ezpeleta et al. 2005). In addition, the Archaeplastida storage metabolism network has witnessed only one invention (a novel enzyme activity not found elsewhere): the glucan and phosphoglucan water dikinases (GWD and PWD) which are required to mobilize the crystalline structures of starch which would otherwise remain inaccessible to hydrosoluble enzymes (reviewed in Fettke et al. 2009). These enzymes phosphorylate the insoluble amylopectin crystals from the β -phosphate of ATP through a dikinase reaction. This renders the otherwise hydrophobic crystals more

accessible to attack by hydrosoluble enzymes of starch catabolism such as β -amylases. The “invention” of the GWD/PWD enzymes results from a fusion of a CBM20 (carbohydrate binding module 20) with a dikinase domain. It probably occurred in the host cytosol since the introduction of the phosphate also requires the hydrolysis of the latter through glucan phosphatases. CBM20-containing glucan phosphatases have been reported as enzymes preventing the accumulation of abnormal hyper-phosphorylated glycogen in animals (Tagliabracci et al. 2008). In particular, “laforin” was documented as such a phosphatase which when defective led to the pathological accumulation of lafora bodies (abnormal glycogen) in the brains of humans afflicted by “lafora disease.” A related defective “SEX4” glucan phosphatase of *Arabidopsis* was reported to prevent normal starch accumulation in this model plant (Kotting et al. 2009). Interestingly this defect could be complemented by the introduction of the animal laforin gene (Gentry et al. 2007). Hence, invention of the GWD/PWD enzymes was greatly facilitated by the preexisting metabolism of phosphorylated glycogen in eukaryotes. However, no bacteria have ever been reported to accumulate phosphorylated glycogen and this pathway is presumed to be inexistent in these organisms and in particular in cyanobacteria. Hence, the routes of starch catabolism in cyanobacteria are entirely different relying possibly on distinct enzymes and mechanisms. If the cyanobiont was initially a starch accumulator it remains highly unlikely that the loss of a catabolic enzyme from the cyanobiont genome could be complemented by the targeting of a corresponding host protein (i.e., of eukaryotic phylogeny) to the evolving symbiont.

The archaeplastidal starch metabolism network differs from that of eukaryotic glycogen metabolism (exemplified by *Entamoeba histolytica* in Table 1) by two critical steps. The first is defined by the aforementioned GWDs and PWDs, and the second concerns the presence of direct debranching enzymes (named isoamylase) which are never found in storage polysaccharide metabolism of heterotrophic eukaryotes. In green plants and algae the direct debranching enzymes are responsible for generating the asymmetrical and ordered distribution of branches that generates the cluster structure of amylopectin which in turn is required for aggregation into solid semicrystalline structures. These two features are sufficient to explain the difference between glycogen and starch metabolism in eukaryotes.

Looking at the starch synthases present in the three archaeplastida lineages it appears that the major soluble starch synthases responsible for amylopectin synthesis display a phylogenetic origin in agreement with its substrate specificity. Hence the UDP-glucose-specific enzymes are clearly related to enzymes of glycogen synthesis of CAZy family GT5 that use UDP-glucose in several eukaryotic lineages (with the noticeable exception of fungi and animals which use a GT3 enzyme) while SSIII–IV of Chloroplastida are related to bacterial GT5 enzymes that use ADP-glucose. However, the source of the chloroplastidal SSIII–IV enzyme is likely to be chlamydial rather than cyanobacterial (Moustafa et al. 2008; Ball et al. 2013). The source for the SSI–SSII enzymes likely was a duplication of the GBSSI gene that occurred selectively in Chloroplastida while the source of GBSSI is distinctively cyanobacterial (Deschamps et al. 2008a; Ball et al. 2013).

If we further look at the phylogenetic origin of the enzymes (Table 1) it appears that glaucophytes contain a complete network of eukaryotic glycogen metabolism. Rhodophyceae only lack indirect debranching enzyme while Chloroplastida lack indirect debranching enzyme and the GT5 UDP-glucose-specific glycogen(starch) synthase. However and very importantly, the only contribution of cyanobacteria to the archaeplastidal network consists of GBSSI in all three lineages and of GBSSI and ADP-glucose pyrophosphorylases in Chloroplastida. Enzymes that were previously thought of as cyanobacterial, like SSIII, SSIV, DPE1 (D-enzyme), pullulanase, and isoamylase turned out to display very clear distinct bacterial origins. Most importantly, SSIII-IV and isoamylase originated most probably from chlamydial intracellular pathogens (Moustafa et al. 2008).

Reconstruction of the Ancient Network of Starch Metabolism

The current predominant view states that Archaeplastida are monophyletic and that plastid endosymbiosis defines a common ancestor for the whole group (Rodríguez-Ezpeleta et al. 2005). We can then assume that most genes (with the exception of pullulanase which is not monophyletic and *dpe1* which was selectively transferred horizontally from an unknown bacterium to the Chloroplastida) will have been transmitted vertically from the common ancestor to the three Archaeplastida lineages. Therefore, one can reconstruct the minimal network of enzymes which had to be present in the common ancestor to explain the present distribution of genes encoding them in Glaucophyta, Rhodophyceae, and Chloroplastida. In order to do this we further minimized the number of isoforms to a single enzyme when we have good reasons to suspect that gene duplications and subfunctionalization have occurred post-endosymbiosis. The ancestral pathway reconstructed in Table 1 relies on both ADP-glucose and UDP-glucose. This ancestral pathway contains a complete set of cytosolic eukaryotic glycogen metabolism enzymes but lacks all but two cyanobacterial enzymes (ADP-glucose pyrophosphorylase and GBSSI). Because all three lineages display the same pattern of loss (with the exception of ADP-glucose pyrophosphorylase which was kept by Chloroplastida only) it is more parsimonious to suppose that the loss occurred once in the common ancestor at a very early stage. We believe this implies that the cyanobiont had lost the opportunity to metabolize storage polysaccharides and that the latter were only present in the ancestor's cytosol. Indeed, if the cyanobiont gene losses occurred at a very early stage the complementation of a missing enzyme by the supply of a corresponding host enzyme would have been problematic. First the major plastidial protein targeting machinery (TOC–TIC) may not have been yet routinely efficient and second, as mentioned earlier, prior to plastid endosymbiosis starch-storing cyanobacteria had evolved mechanisms of polysaccharide mobilization entirely different from those found in Archaeplastida or heterotrophic glycogen

accumulating eukaryotes. This would make complementation of gene loss by supply of corresponding eukaryotic enzymes unlikely. The suggestion that the ancestor had lost the ability to store glucose in the cyanobiont is further strengthened by three additional observations.

First, Henrissat et al. (2002) have noted that parasites and symbionts in general tended to lose storage polysaccharide metabolism as a function of their obligatory intracellular lifestyle. This seems to be the case for the only other photosynthetic cyanobacterial endosymbiont reported: that carried by *Paulinella chromatophora* where the chromatophore genome seems to have already lost the genes of storage polysaccharide metabolism (Nowack et al. 2008).

Second, if storage polysaccharide metabolism was lost at a very early stage this would imply that the ancestor of Chloroplastida synthesized starch in the cytosol and the pathway was redirected to plastids when the green algae evolved. We have previously reviewed in detail the possible reasons and the problems dealing with the redirection of starch metabolism to plastids (Deschamps et al. 2008b, c). Suffice it to say here that this probably defined two intermediate stages (MOS and glycogen accumulation) that likely generated a requirement for duplications of enzyme forms followed by enzyme subfunctionalizations (Deschamps et al. 2008c). Indeed, the Chloroplastida have selectively experienced such duplication and subfunctionalization rounds leading to their characteristic highly redundant pathway (see above). This further suggests that the ancestral network was, as proposed initially, exclusively cytosolic.

Third, Glaucophyta and Rhodophyceae still synthesize starch exclusively in the cytosol today with no evidence for the presence of plastidial storage polysaccharides. Since Glaucophyta in general are assumed to have conserved a greater number of ancestral features it would further support that cytosolic starch deposition was ancestral in Archaeplastida and that plastidial starch synthesis is derived.

Taken together, the loss of the vast majority of enzymes of cyanobacterial starch metabolism and the three aforementioned observations make a compelling case for an ancient localization of storage polysaccharides exclusively in the host cytosol shortly after endosymbiosis. This would imply that the ADP-glucose-specific starch synthase would have been active in the cytosol. This in turn would require the presence of ADP-glucose in this compartment. However, ADP-glucose is not synthesized by eukaryotes. Hence, one would be tempted to place the cyanobacterial ADP-glucose pyrophosphorylase in the host cytosol shortly after endosymbiosis, thereby generating the ancestral cytosolic dual substrate pathway. The problem with such a hypothesis is its lack of physiological relevance. Dual substrate pathways are not common in biochemistry and the advantage of producing ADP-glucose in addition to UDP-glucose in the host cytosol is anything but obvious. As noted previously, ADP-glucose pyrophosphorylase in cyanobacteria and plants is an enzyme finely tuned by photosynthesis and the Calvin cycle. Because these processes have never moved out of the plastid, it seems reasonable to assume that ADP-glucose pyrophosphorylase will have never left this compartment during the whole evolution process. If the enzyme is left within the future

plastid stroma then the system to become functional requires the presence of an ADP-glucose transporter on the cyanobiont inner membrane to feed the cytosolic ADP-glucose requiring glycogen (starch) synthase. With such a transporter the physiological relevance of the reconstructed dual substrate pathway becomes enlightening. We have outlined above that storage can be predicted to define an interesting buffer between the unsynchronized supply and demand for carbon during plastid endosymbiosis. The proposed reconstruction of storage polysaccharide metabolism in the common ancestor of Archaeplastida lends considerable support to this prediction. Upon close examination of the carbon flux generated in this proposed reconstruction, it seems impossible to imagine a better suited first connection between the unrelated partners of plastid endosymbiosis. The carbon that flows through ADP-glucose pyrophosphorylase within the cyanobiont would have normally been committed to storage since this nucleotide sugar is devoted to glycogen (starch) synthesis in bacteria. Hence this carbon committed to leave cyanobacterial metabolism by becoming temporarily unavailable in the form of solid cyanophycean starch will similarly escape the latter by physically moving out of the cyanobiont through the ADP-glucose translocator. The cyanobacterial pathways have been optimized to generate and control this escape for millions of years and there are no penalties to be expected for such an export of carbon. This to our knowledge would not be the case for any other possible carbon substrate. Upon arrival in the host cytosol ADP-glucose is unlikely to affect host metabolism which does not recognize it. ADP-glucose will thus exclusively feed cytosolic glycogen synthesis and increase the available storage carbohydrate pools. The only very modest penalty will be the uncontrolled increase in the osmotic impact of glycogen which can be considered as negligible. Access to the cytosolic glycogen pools will be through the eukaryotic glycogen catabolism regulatory networks that have been tailored by millions of years of selection to respond optimally to the various needs of this ancient protist in a changing environment. The system was thus optimal at the very moment phagocytosis aborted and the connection was established. Hence, reconstruction of starch metabolism in the archaeplastida ancestor opens an unexpected window on the nature of the biochemical connection that drove plastid endosymbiosis. It should be emphasized here that reconstruction of starch metabolism in the common ancestor does not exactly reflect the situation present at the time endosymbiosis was established but rather shortly thereafter but before the three Archaeplastida lineages diverged. Indeed, in our reconstruction proposal starch is present in the cytosol and the GWD–PWD pathway of starch catabolism also. This situation is very close to that found in extant glaucophytes. It is however unlikely that the transition from glycogen to starch occurred immediately as this required the simultaneous “invention” of the dikinase-CBM20 gene fusion to generate the GWD–PWD required to catabolize the glucan crystals and the recruitment of a modified bacterial debranching enzyme to generate the “crystals.” Indeed the bacterial source of this enzyme did not display the required “isoamylase” type of substrate specificity to begin with. However the latter is thought to be needed to synthesize crystalline amylopectin. The bacterial ancestors displayed a much narrower substrate specificity consisting of hydrolysis of external glycogen chains

of three to four glucose residues long (Dauvillée et al. 2005). The required modifications will have required gene duplications and the accumulation of mutations changing the enzyme specificity on the duplicated locus. Similarly the enzymes that work downstream from the phosphorylated crystals had to accumulate mutations in their genes that optimized their action by comparison to their previous analogous role in glycogen breakdown. This all took time and suggests that the transition came about later. It nevertheless happened fast enough to allow for the recruitment of the cyanobacterial GBSS gene by EGT which was otherwise likely to have been very quickly lost since the cyanobiont presumably became starchless very early on. The early glycogen and late starch accumulation stages are displayed in Fig. 1a, b. The reconstructed pathway from Table 1 reflects the late starch accumulation stage (Fig. 1b).

Incompatibility of Phototrophy and Diazotrophy in Photosynthetic Eukaryotes

As stated above, cyanobacterial metabolism has been tailored by natural selection to allow for escape of excess carbon in the form of ADP-glucose. Yet, we must still admit that it has also been selected to be able to tap carbon in the polysaccharide stores when required and this will certainly happen mostly in darkness. Hence the tolerance for carbon escape does not necessarily mean that cyanobacteria can do with no carbon stores. Mutants of storage polysaccharide synthesis completely lacking glycogen and starch, respectively, have been obtained in yeast (Thon et al. 1992), *E. coli* (Damotte et al. 1968), *Chlamydomonas reinhardtii* (Zabawinski et al. 2001), *Arabidopsis thaliana* (Lin et al. 1988), and more recently in both *Synechocystis* sp. PCC 6803 (Miao et al. 2003) and *Synechococcus elongatus* (Suzuki et al. 2010). In yeast, *E. coli*, and *Chlamydomonas* there is virtually no impact on growth of these microorganisms under laboratory conditions. In *Arabidopsis thaliana*, starchless mutants grow normally under continuous light. However, under day and night cycles growth of the mutant plants becomes stunted. In cyanobacteria, mutants lacking glycogen have been produced that carry a defect for the single cyanobacterial ADP-glucose pyrophosphorylase subunit (Miao et al. 2003; Suzuki et al. 2010). Growth of these mutants was monitored under continuous illumination. Under these conditions growth proved to be normal and the requirement for glycogen pools minimal. However, photosynthesis and respiration were impacted, with a significant reduction in photosynthesis especially under high light and a 50 % reduction in respiration activity in darkness. The reduction in photosynthesis was thought to be due to a decrease in the regeneration of oxidized NADP^+ allowed through glycogen synthesis. The decreased respiration was attributed to the fact that glycogen breakdown accounted for a significant portion of the accessible carbohydrate substrate pools.

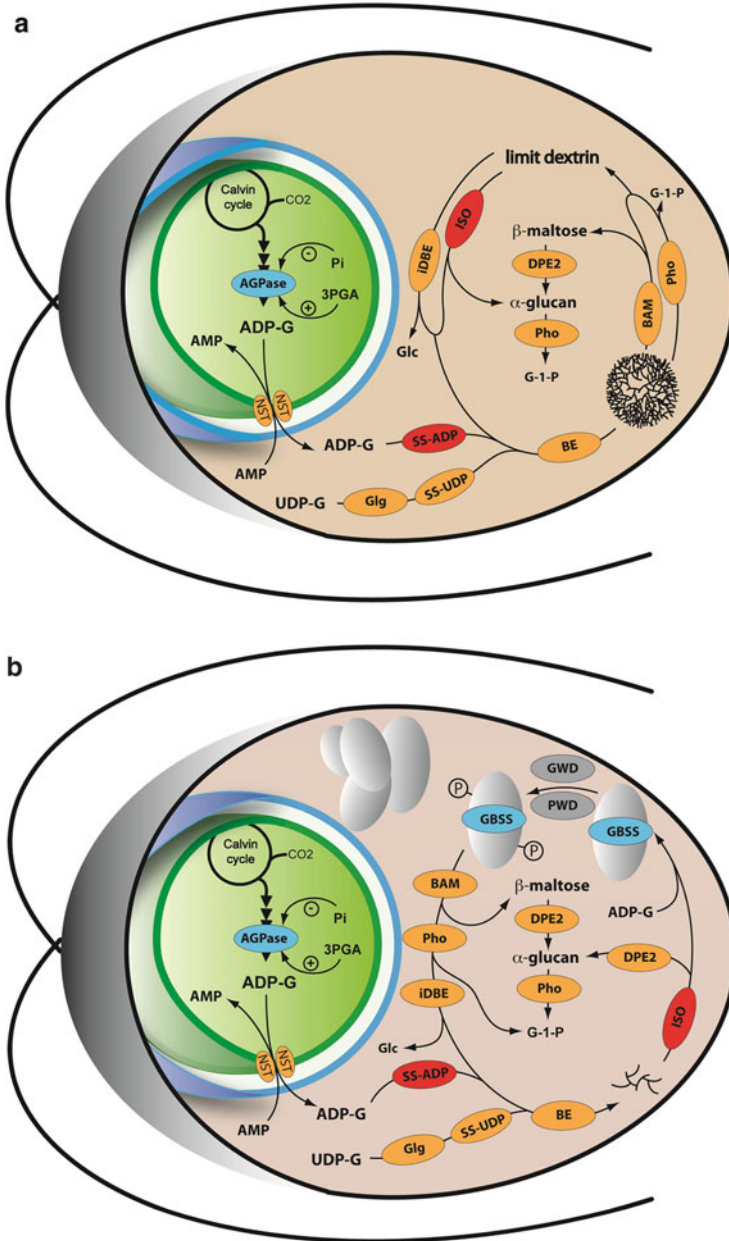


Fig. 1 (a) The storage polysaccharide network at the onset of endosymbiosis. In this reconstruction the transition to starch in the common ancestor cytosol has not yet occurred. This implies that the direct debranching enzyme (labeled *iso*) of bacterial phylogeny did not yet duplicate to generate the isoamylase required for amylopectin crystallization. As a consequence there was no need yet for the evolution of the GWD–PWD dikinases. The bacterial direct debranching enzyme still displays its ancestral bacterial function which is to debranch the product of glycogen

Quite interestingly, the *Synechococcus* mutants were shown to be also more sensitive to salt and photooxidative stresses. Indeed it was hypothesized by Deschamps et al. (2008c) that ATP in darkness may define a critical limitation in the absence of stored carbon. Under these conditions magnesium chelatase would not be able to assemble, thereby leading to the accumulation of photoactive intermediates of chlorophyll synthesis which upon the return of light would induce photooxidative stresses (Reinhold et al. 2007). It was in addition hypothesized by Deschamps et al. (2008c) that the selective increase in the chlororophyll synthesis that could have accompanied the evolution of chlorophyll *b*-containing antennae may have defined the selection pressure that has propelled the return of storage polysaccharides to the chloroplasts of the evolving green lineage.

Fig. 1 (continued) degradation (labeled *limit dextrin*) by glycogen phosphorylase (labeled *pho*). The eukaryotic indirect debranching enzyme has a similar function. However unlike the eukaryotic indirect debranching enzyme (idBE), the bacterial enzyme releases the maltotetraose outer chains (labeled *α-glucan*) in the cytosol which may have been subjected to degradation by a combination of the *dpe2* amylomaltase and the glycogen phosphorylase. The cytosolic dual substrate pathway of glycogen accumulation relies on both UDP-glucose (UDP-G) generated through host biochemical networks in its cytosol according to host needs and ADP-glucose generated by the cyanobacterial ADP-glucose pyrophosphorylase (AGPase) which is activated by 3-PGA and inhibited by orthophosphate according to the cyanobiont's networks and physiology. To be incorporated into cytosolic glycogen this substrate has to be exported by a nucleotide sugar translocator (labeled *NST*) of host origin which exchanges ADP-glucose with AMP. The ADP-glucose substrate in the cytosol will have to be incorporated through an ADP-glucose-specific glucan synthase (labeled *SS-ADP*). On the other hand, the host UDP-glucose pools will be directed to glycogen according to the highly regulated eukaryotic UDP-glucose-specific glucan synthase (labeled *SS-UDP*). This enzyme unlike the bacterial glucan synthase requires a primer to elongate a glucan. This primer is defined by glycogenin, an autoglucosylating protein (labeled *Glg*). The glucans elongated through both glucan synthases will then be branched into glycogen by branching enzyme (labeled *BE*). The glycogen outer chains will be degraded through either β-amylase (labeled *BAM*) or glycogen phosphorylase (labeled *Pho*) to generate maltose and glucose-1-P, respectively. The maltose will be metabolized by the *dpe2* amylomaltase. Enzymes of host phylogenetic origin are colored in *beige*. Those of cyanobacterial origin in *blue* and those of chlamydial origin in *red*. **(b)** Cytosolic storage polysaccharide metabolism has been reconstructed as detailed in the text. This early stage corresponds to the common ancestor starch metabolism after the transition from glycogen to starch has occurred. This transition required the duplication and evolution of the bacterial direct debranching enzyme into a functional isoamylase (*iso*). This enzyme processes the branches generated randomly on the hydrophilic branched polysaccharides generated by branching enzymes. The debranched glucans (labeled *α-glucan*) will be metabolized through a combination of amylomaltases and phosphorylases. Simultaneously a gene fusion between a CBM20 (carbohydrate binding module) possibly from laforin (see text) and a dikinase domain enabled the phosphorylation and loosening of the otherwise undegradable amylopectin crystals (displayed by the *circled P* attached to the white starch granules). This fusion generated the archaeplastidal GWD–PWD inventions (labeled in *gray*) which were required to initiate starch catabolism through the β-amylase and phosphorylases (see above). The presence of polysaccharides aggregated into semicrystalline starch granules enabled the binding and function of the cyanobacterial GBSS (displayed bound to starch) responsible for amylose synthesis within the polysaccharide matrix

In face of these results and speculations, we can predict how the loss of storage polysaccharide would have impacted the cyanobiont at endosymbiosis. We believe that unlike the cyanobacterial mutants, there would have been no impact at all on photosynthesis since the cyanobiont is still able to consume reducing equivalents through the synthesis and export of ADP-glucose in a fashion entirely similar to glycogen or starch synthesis of wild-type algae. In addition the cyanobiont being sheltered within an eukaryotic cytosol, we do not believe that it would have still required a particular resistance to osmotic stress. Yet we expect that respiration in darkness would have been dramatically reduced. Hence, there may have been a strong requirement for an ability to import ATP in darkness which was exacerbated by the need to assemble magnesium chelatase to obviate photooxydative stresses (Reinhold et al. 2007; Deschamps et al. 2008c). This may define the reason why all three Archaeplastida lineages have recruited ATP/ADP transport proteins (NTT, nucleotide transporter) on the inner membrane of their plastids which drive the unidirectional import of ATP in exchange for ADP. The gene encoding these transporters was acquired by lateral gene transfer from a Chlamydiale source (Linka et al. 2003). This transporter may have defined a critical early requirement for successful endosymbiosis. Although import of cytosolic ATP may have been sufficient to obviate photooxydative stresses we believe this import was unlikely to have allowed the maintenance of diazotrophy in the cyanobiont. Indeed, the cyanobacterial ancestor may be affiliated to extant unicellular diazotrophic cyanobacteria of clade B. A number of studies have suggested that these cyanobacteria reach the anoxia status required for nitrogen fixation in darkness thanks to the respiration of their large starch pools (Colón-López et al. 1997; Schneegurt et al. 1994). Not only was the vast amount of ATP and reducing power needed to feed nitrogenase but above all the consumption of the local O₂ by respiration would by itself be required. When the cyanobiont lost its storage polysaccharide pools very early on (possibly even at the onset of endosymbiosis) it lost the ability to fix nitrogen at the same time and thus very quickly also lost the *nif* genes. Indeed, while the NTT transporter could in theory supply the ATP to obviate photooxydative stresses it could not compensate for the decrease in the respiration activity absolutely required to reach anoxia. We believe this explains why in nature no member of the Archaeplastida has retained the ability to fix nitrogen that was originally displayed by the cyanobiont's ancestor. This also suggests that after endosymbiosis the cyanobiont had to be provided with some form of reduced nitrogen by the host.

The ADP-Glucose Connection

As mentioned above, the incorporation into glycogen of glucose from ADP-glucose in the common ancestor's cytosol required the transport of this glycosyl-nucleotide out of the cyanobiont. It is unlikely that the cyanobiont encoded such a transporter. Indeed the physiological significance of a protein exporting ADP-glucose into the

extracellular medium would be hard to imagine for free-living bacteria. Likewise the host is not likely to harbor such a transporter since ADP-glucose is neither synthesized nor used by eukaryotes. Clues to the elusive origin of the putative ancient ADP-glucose translocator came when Weber et al. (2006) (see also Facchinelli and Weber 2013) examined the phylogeny of the major extant plastidial carbon translocators. These belong to a family of proteins known as the pPT (phosphosugar phosphate translocator) proteins which exchange sugar phosphates for orthophosphate. The best studied transporter of this family is the TPT or triose phosphate translocator which is responsible for the export of carbon from the chloroplast to the plant leaf cell cytosol. Weber et al. (2006) demonstrated that the whole family of transporters found in green and red algae and secondary endosymbiosis derivatives was monophyletic. It was further assumed to have originated through duplication and evolution of a gene for a host endomembrane transporter. The authors postulated that this ancient transporter probably was involved in establishing the symbiotic flux.

Upon looking closer to the eukaryotic origin of these transporters these appeared to consist of members of a family of nucleotide sugar translocators (for a review, see Handford et al. 2006) known as the NST3 family (defined in Martinez-Duncker et al. 2003). Among the nucleotide sugar translocator families, NST3 defines a family that transports not only pyrimidine sugar nucleotides but also many purine sugar nucleotides (Martinez-Duncker et al. 2003). Interestingly, NST3 contains many GDP-mannose translocators, the latter defining a structural analog of ADP-Glc. Colleoni et al. (2010) thus tested the abilities of GDP-mannose translocators from yeast and *Arabidopsis* to transport ADP-Glc in yeast membrane-derived liposomes. They were able to show that the *Arabidopsis* GDP-mannose translocator was able to transport ADP-Glc as efficiently as GDP-mannose but displayed a lower affinity for the non-physiological substrate. Nevertheless, the K_m for ADP-glucose remained at a 1–5 mM concentration range which is in agreement with a possible role of such a transporter in establishing the initial symbiotic link. Indeed, it is expected that mutants which are blocked in the utilization of ADP-glucose will see the size of their ADP-glucose pools rise above 1 mM as was demonstrated in cereal endosperm mutants (Shannon et al. 1996). The cyanobiont, having lost the ability to polymerize glucans but not to synthesize the nucleotide sugar substrate, was in precisely that situation.

One of the obvious problems faced by the host-endomembrane-derived putative ancestral ADP-glucose translocator was how to reach the cyanobiont's inner membrane at a time when no plastidial protein targeting machinery was likely to have existed. An interesting observation was published by Loddenkötter et al. (1993) concerning the expression in yeast of the TPT deprived of its transit peptide sequence. The protein was found to be located on the yeast mitochondrial membranes. Although contamination of ER membranes could not be definitively ruled out, the authors reported that in vitro also the protein was associated with yeast or *Neurospora* mitochondria in an energy- and receptor-independent fashion, strongly suggesting that this protein displayed an innate ability to reach the organelle membranes in the absence of a functional transport system. If such a property

was displayed by the ancestral transporter it would have greatly facilitated its recruitment at the onset of plastid endosymbiosis.

We thus postulate that such a transporter accidentally reached the cyanobiont's inner membrane. A duplicated gene encoding this transporter subsequently enabled it to be expressed and regulated independently of the endomembrane sugar nucleotide translocators. Later, evolution further explored the numerous possibilities of substrate exchange offered by this family of transporters whose expression had been optimized with respect to photosynthate export. This yielded, thanks to other duplicated copies, the pPT family that allowed for a more integrated solution to the export of carbon from plastids. Both the ADP-glc translocator and the pPT coexisted until the Archaeplastida lineages diverged. This happened when the emerging Archaeplastida lost the ability to synthesize glucans from ADP-glc in the cytosol. In Rhodophyceae, it happened when the ADP-glucose-specific starch synthase was lost while in Chloroplastida it happened at the final stages of rewiring of the storage polysaccharide metabolism network to plastids, i.e., when starch disappeared from the cytosol.

ADP-Glucose Transport and Glucan Polymerization

The second condition that had to be met at the onset of plastid endosymbiosis was the presence in the host cytosol of a glucan synthase able to use ADP-glc. The eukaryotic glycogen synthases use UDP-glc with little or no activity for purine nucleotide sugars as substrates. Immediate establishment of the symbiotic flux of carbon was however required to allow for natural selection of plastid endosymbiosis. This did not give the required time for the accumulation and selection of mutations in the host glucan transferase gene. Clearly an efficient ADP-glucose utilizing glycogen/starch synthase had to be present in the host cytosol at the onset of the event. Such enzymes are never observed in eukaryotes and have only been described in the bacterial or archaean domains (with the exception of course of green algae and plants). To get further insights into this problem we have examined the phylogeny of extant archaeplastidal starch synthases that use ADP-glc (Ball et al. 2013). Two monophyletic groups are found in green algae and plants (Chloroplastida): the GBSSI–SSI–SSII group and the SSIII–IV group (Deschamps et al. 2008a; Ball et al. 2013). The GBSSI–SSI–SSII clade can be reasonably rooted by considering that the enzyme source is defined by the cyanobacterial GBSSI gene. With such a root in mind, the published phylogenetic trees support a transfer of the GBSSI gene to the Archaeplastida ancestor before the three lineages diverged. When GBSSI was bound to starch in the ancestor's cytosol it was exposed to the presence of both UDP-glc and ADP-glc that drove the ancient dual substrate pathways of storage polysaccharide synthesis. The gene therefore accumulated mutations that turned this low-affinity cyanobacterial enzyme that originally only used ADP-glc into a bifunctional synthase accepting both glycosyl nucleotides as

substrates to achieve amylose synthesis whenever either the ADP-glc or the UDP-glc pools rose above the required levels.

Under this scenario with cyanobacteria at the root of the clade, the GBSSI gene duplicated and accumulated mutations that turned the duplicated gene product into a soluble (unbound) activity. This probably happened selectively in the Chloroplastida lineage as the pathway was redirected to plastids. Hence SSI–SSII were not available at the time of endosymbiosis and no bacterial glucan synthase shows significant proximity to these enzymes despite the presence of hundreds of available bacterial whole genome sequences (Ball et al. 2013). Under this hypothesis, which is consistent with the phylogeny and a cyanobacterial source of GBSSI, we can conclude that only wild-type GBSSI could have been available in the host cytosol. However this can only be imagined, provided an LGT had just happened shortly before or at endosymbiosis because of the phagotrophic habit of the ancestral protist (“you are what you eat”). This otherwise nonproductive event could thus have been selected to establish the symbiotic flux. However, this hypothesis does not stand in face of the biochemical properties of GBSSI. GBSSI is an enzyme that displays very little activity when expressed as soluble protein both in vivo (Dauvillée et al. 1999) and in vitro as a recombinant enzyme unbound to starch (Edwards et al. 1999). Yet at the time of endosymbiosis, the eukaryote ancestor synthesized glycogen and not starch which evolved shortly thereafter. A wild-type GBSSI protein would not have been able in such a context to polymerize glucan from ADP-glucose onto glycogen.

SSIII–SSIV presently define the only extant archaeplastidal transferases whose ancestor could have played a role in supplying the symbiotic link. One of the most surprising findings of the recently established *Cyanophora paradoxa* genome sequence (Price et al. 2012) was the description of an SSIII–SSIV-like enzyme sequence which is presumably involved in cytosolic starch synthesis. We believe this enzyme still uses ADP-glc only in present-day *Cyanophora*, but this yet needs to be demonstrated. This could thus suggest that *Cyanophora paradoxa* might very well define a living fossil of the putative ancestral dual substrate pathway, although it is apparently lacking both ADP-glucose pyrophosphorylase and the ADP-glucose translocator. In line with this suggestion, the GBSSI of glaucophytes displays similar affinities for both nucleotide sugars and the *C. paradoxa* phosphorylase is surprisingly exquisitely sensitive to mixed inhibition by ADP-glucose (Plancke et al. 2008). The source of the ADP-glucose in glaucophytes needs to be ascertained, but it is already known that the reversible sucrose synthase reaction using ADP in place of UDP is not involved since these organisms lack sucrose metabolism altogether (Price et al. 2012). What would be the rationale for glaucophytes to have kept this enzyme in its cytosol while they have lost the ability to produce the ADP-glucose substrate in plastids? Clues can be found in the exceptional properties displayed by the SSIII–SSIV starch (glycogen) synthases. These enzymes are involved in the priming of starch granules and thus control their numbers and sizes (Roldán et al. 2007; Szydlowski et al. 2009). In addition, unlike other starch synthase mutant combinations a double SSIII–SSIV defective mutant abolishes starch synthesis. This essential in vivo function correlates with the in vitro

ability displayed at least by SSIII to prime polysaccharide synthesis (Szydlowski et al. 2009). The GT3 UDP-glc requiring glycogen synthase of fungi and animals requires the presence of glycogenin, an autoglucosylating protein used as a primer (Cheng et al. 1995). The *Cyanophora paradoxa* genome contains no convincing glycogenin candidate sequence. Hence, the GT5 UDP-glucose requiring glycogen (starch) synthase of glaucophytes may have become dependent on the SSIII–SSIV-like enzyme for polysaccharide synthesis priming.

The SSIII–SSIV clade defines a monophyletic group consisting of the plant enzymes as well as a number of related enzymes from cyanobacteria, proteobacteria, and Chlamydiales (Ball et al. 2013). The phylogeny of this group is complex because of the existence of several LGT events splitting the Chlamydiales into two groups. Despite this complexity, all possible scenarios reject the cyanobacteria as the source of the archaeplastidal enzyme (Ball et al. 2013). The presence of the Chlamydiales at the base of the clade and the fact that the pathogens define the only group of organisms containing this glucan synthase as sole enzyme used for glycogen metabolism suggest that the enzyme may have evolved and acquired its exceptional biochemical properties in the pathogens. These genes were then passed on to Archaeplastida, proteobacteria, and cyanobacteria.

Clearly, the ancestor of extant SSIII–SSIV does qualify as a serious candidate to provide the enzyme used to establish the symbiotic link between the cyanobiont and its host. Why would such an enzyme have been present in the host cytosol at the advent of plastid endosymbiosis? We believe this question may be answered when a clear understanding of glycogen metabolism function will be reached in the group of organisms that are the most likely source for the LGT to Archaeplastida: the Chlamydiales intracellular pathogens (Moustafa et al. 2008).

References

- Ball S, van de Wal M, Visser R (1998) Progress in understanding the biosynthesis of amylose. *Trends Plant Sci* 3:462–467
- Ball SG, Colleoni C, Cenci U, Raj JN, Tirtiaux C (2011) The evolution of the glycogen and starch pathway in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. *J Exp Bot* 62:1775–1801
- Ball SG, Subtil A, Bhattacharya D, Moustafa A, Weber APM, Gehre L, Colleoni C, Arias MC, Cenci U, Dauvillée D (2013) Metabolic effectors secreted by bacterial pathogens; essential facilitators of plastid endosymbiosis? *Plant Cell* 25(1):7–21
- Ballicora MA, Iglesias AA, Preiss J (2003) ADP-glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogensynthesis. *Microbiol Mol Biol Rev* 67:213–225
- Boos W, Shuman H (1998) Maltose/maltodextrin system of *Escherichia coli*: transport, metabolism, and regulation. *Microbiol Mol Biol Rev* 62:204–229
- Buléon A, Colonna P, Planchot V, Ball S (1998) Starch granules: structure and biosynthesis. *Int J Biol Macromol* 23:85–112
- Chan XC, Gross J, Yoon HS, Bhattacharya D (2011) Plastid origin and evolution: new models provide insights into old problems. *Plant Physiol* 155:1552–1560

- Cheng C, Mu J, Farkas I, Huang D, Goebel MG, Roach PJ (1995) Requirement of the self-glucosylating initiator proteins Glg1p and Glg2p for glycogen accumulation in *Saccharomyces cerevisiae*. *Mol Cell Biol* 15:6632–6640
- Colleoni C, Linka M, Deschamps P, Handford MG, Dupree P, Weber APM, Ball SG (2010) Phylogenetic and biochemical evidence supports the recruitment of an ADP-glucose translocator for the export of photosynthate during plastid endosymbiosis. *Mol Biol Evol* 27:2691–2701
- Colón-López MS, Sherman DM, Sherman LA (1997) Transcriptional and translational regulation of nitrogenase in light-dark- and continuous-light-grown cultures of the unicellular cyanobacterium *Cyanothece* sp. strain ATCC 51142. *J Bacteriol* 179:4319–4327
- Coppin A, Varre JS, Lienard L, Dauvillée D, Guerardel Y, Soyer-Gobillard MO, Buléon 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. *J Mol Evol* 60:257–267
- Damotte M, Cattaneo J, Sigal N, Puig J (1968) Mutants of *Escherichia coli* K 12 altered in their ability to store glycogen. *Biochem Biophys Res Commun* 32:916–920
- Dauvillée D, Colleoni C, Shaw E, Mouille G, D’Hulst C, Morell M, Samuel MS, Bouchet B, Gallant DJ, Sinskey A, Ball S (1999) Novel starch-like polysaccharides are synthesized by a soluble form of granule-bound starch synthase in glycogen accumulating mutants of *Chlamydomonas reinhardtii*. *Plant Physiol* 119:321–330
- Dauvillée D, Kinder 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. *J Bacteriol* 187:1465–1473
- Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux JL, Buléon A, Haebel S, Ritte G, Steup M, Falcon LI, Moreira D, Loeffelhardt W, Nirmal Raj J, Plancke C, D’Hulst C, Dauvillée D, Ball S (2008a) Metabolic symbiosis and the birth of the plant kingdom. *Mol Biol Evol* 25:536–548
- Deschamps P, Moreau H, Worden AZ, Dauvillée D, Ball SG (2008b) Early gene duplication within chloroplasts and its correspondence with relocation of starch metabolism to chloroplasts. *Genetics* 178:2373–2387
- Deschamps P, Haferkamp I, D’Hulst C, Neuhaus E, Ball S (2008c) The relocation of starch metabolism to chloroplasts: when, why and how. *Trends Plant Sci* 13:1802–1816
- Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol Biol Evol* 25:748–761
- Edwards A, Borthakur A, Bornemann S, Venail J, Denyer K, Waite D, Fulton D, Smith A, Martin C (1999) Specificity of starch synthase isoforms from potato. *Eur J Biochem* 266:724–736
- Facchinelli F, Weber APM (2013) Insertion of metabolite transporters into the endosymbiont membrane(s) as a prerequisite for primary endosymbiosis. In: Löffelhardt W (ed) *Endosymbiosis*. Springer, Wien New York, pp 53–80
- Fettke J, Hejazi M, Smirnova J, Höchel E, Stage M, Steup M (2009) Eukaryotic starch degradation: integration of plastidial and cytosolic pathways. *J Exp Bot* 60:2907–2922
- Gentry MS, Downen RH, Worby CA, Mattoo S, Ecker JR, Dixon JE (2007) The phosphatase laforin crosses evolutionary boundaries and links carbohydrate metabolism to neuronal disease. *J Cell Biol* 178:477–488
- Gupta RS (2009) Protein signatures (molecular synapomorphies) that are distinctive characteristics of the major cyanobacterial clades. *Int J Syst Evol Microbiol* 59:2510–2526
- Handford M, Rodriguez-Furlán C, Orellana A (2006) Nucleotide-sugar transporters: structure, function and roles *in vivo*. *Braz J Med Biol Res* 39:1149–1158
- Henrissat B, Deleury E, Coutinho PM (2002) Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? *Trends Genet* 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. *J Mol Evol* 48:723–739

- Kotting O, Santelia D, Edner C, Eicke S, Marthaler T, Gentry MS, Comparot-Moss S, Chen J, Smith AM, Steup M, Ritte G, Zeeman SC (2009) STARCH-EXCESS4 is a laforin-like phosphoglucan phosphatase required for starch degradation in *Arabidopsis thaliana*. *Plant Cell* 21:334–346
- Krebs EG (1983) Historical perspectives on protein phosphorylation and a classification system for protein kinases. *Philos Trans R Soc B Biol Sci* 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 Physiol* 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 relationships of non-mitochondrial nucleotide transport proteins in bacteria and eucaryotes. *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. *Proc Natl Acad Sci USA* 90:2155–2159
- Martinez-Duncker I, Mollicone R, Codogno P, Oriol R (2003) The nucleotide-sugar transporter family: a phylogenetic approach. *Biochimie* 85:245–260
- Miao X, Wu Q, Wu G, Zhao N (2003) Changes in photosynthesis and pigmentation in an *agp* deletion mutant of the cyanobacterium *Synechocystis* sp. *Biotechnol Lett* 25:391–396
- 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
- Nakamura Y, Takahashi J, Sakurai A, Inaba Y, Suzuki E, Nihei S, Fujiwara S, Tsuzuki M, Miyashita H, Ikemoto H, Kawachi M, Sekiguchi H, Kurano N (2005) Some cyanobacteria synthesize semi-amylopectin type alpha-polyglucans instead of glycogen. *Plant Cell Physiol* 46:539–545
- Nowack ECM, Melkonian M, Glöckner G (2008) Chromatophore genome sequence of *Paulinella* sheds light on acquisition of photosynthesis by eukaryotes. *Curr Biol* 18:410–418
- Nyvall P, Pelloux J, Davies HV, Pedersen M, Viola R (2001) Purification and characterization of a novel starch synthase selective for uridine 5'-diphosphate glucose from the red alga *Gracilaria tenuistipitata*. *Planta* 209:143–152
- Patron NJ, Keeling PK (2005) Common evolutionary origin of starch biosynthetic enzymes in green and red algae. *J Phycol* 41:1131–1141
- Plancke C, Colleoni C, Deschamps P, Dauvillée D, Nakamura Y, Haebel S, Ritte G, Steup M, Buléon A, Putaux JL, Dupeyre D, D'Hulst C, Ral JP, Löffelhardt W, Ball SG (2008) The pathway of starch synthesis in the model glaucophyte *Cyanophora paradoxa*. *Eukaryot Cell* 7:247–257
- Preiss J (1984) Bacterial glycogen synthesis and its regulation. *Annu Rev Microbiol* 38:419–458
- Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, Schwacke R, Gross J, Blouin NA, Lane C, Reyes-Prieto A, Durnford DG, Neilson JAD, Lang BF, Burger G, Steiner JM, Löffelhardt W, Meuser JE, Posewitz MC, Ball S, Arias MC, Henrissat B, Coutinho PM, Rensing SA, Symeonidi A, Doddapaneni H, Green BR, Rajah VD, Boore J, Bhattacharya D (2012) *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science* 335:843–847
- 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. *J Exp Bot* 60:3975–3987
- Reinhold T, Alawady A, Grimm B, Beran KC, Jahns P, Conrath U, Bauer J, Reiser J, Melzer M, Jeblick W, Neuhaus HE (2007) Limitation of nocturnal import of ATP into Arabidopsis chloroplasts leads to photooxidative damage. *Plant J* 50:293–304
- Roach PJ (2002) Glycogen and its metabolism. *Curr Mol Med* 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. *Curr Biol* 15:1325–1330
- Roldán I, Wattedled F, Lucas MM, Delvallé D, Planchot V, Jiménez S, Pérez R, Ball S, D'Hulst C, Mérida A (2007) The phenotype of soluble starch synthase IV defective mutants of *Arabidopsis thaliana* suggests a novel function of elongation enzymes in the control of starch granule formation. *Plant J* 49:492–504
- 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. *J Bacteriol* 176:1586–1597
- Shannon JC, Pien FM, Liu KC (1996) Nucleotides and nucleotide sugars in developing maize endosperms: synthesis of ADP-glucose in brittle-1. *Plant Physiol* 110:835–843
- Shearer J, Graham TE (2002) New perspectives on the storage and organization of muscle glycogen. *Can J Appl Physiol* 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*. *Mar Biotechnol* 9:192–202
- Suzuki E, Ohkawa H, Moriya K, Matsubara T, Nagaïke Y, Fujiwara S, Tsuzuki M, Nakamura Y (2010) Carbohydrate metabolism in mutants of the cyanobacterium *Synechococcus elongatus* PCC 7942 defective in glycogen synthesis. *Appl Environ Microbiol* 76:3153–3159
- Szydłowski N, Ragel P, Raynaud S, Lucas MM, Roldán I, Montero M, Muñoz FJ, Ovecka M, Bahaji A, Planchot V, Pozueta-Romero J, D'Hulst C, Merida A (2009) Starch granule initiation in *Arabidopsis* requires the presence of either class IV or class III starch synthases. *Plant Cell* 21:2443–2457
- 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. *J Biol Chem* 283:33816–33825
- Thon VJ, Vigneron-Lesens C, Marianne-Pepin T, Montreuil J, Decq A, Rachez C, Ball SG, Cannon JF (1992) Coordinate regulation of glycogen metabolism in the yeast *Saccharomyces cerevisiae*: induction of glycogen branching enzyme. *J Biol Chem* 267:15224–15228
- Wang Z, Wilson WA, Fujino MA, Roach PJ (2001) Antagonistic controls of autophagy and glycogen accumulation by Snf1p, the yeast homolog of AMP-activated protein kinase, and the cyclin-dependent kinase Pho85p. *Mol Cell Biol* 21:5742–5752
- Weber APM, Linka M, Bhattacharya D (2006) Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryot 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 Microbiol Rev* 34:952–985
- Zabawinski C, Van Den Koornhuysen 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 a heterotetrameric ADP-glucose pyrophosphorylase. *J Bacteriol* 183:1069–1077