Chapter 4
The Transition from Glycogen to Starch Metabolism in Cyanobacteria and Eukaryotes

Steven Ball, Christophe Colleoni, and Maria Cecilia Arias

Abstract α-1,4-linked glucan chains branched through α-1,6 glucosidic lineages define the most frequently found storage polysaccharides in living cells. These glucans come in two very distinct forms known as glycogen and starch. The small water-soluble glycogen particles distribute widely in Archaea, Bacteria, and heterotrophic eukaryotes, while semicrystalline solid starch seems to be restricted to photosynthetic eukaryotes. This review focusses on the so-called glycosyl-nucleotide-dependent pathway of starch and glycogen synthesis. Through comparative biochemistry of storage polysaccharide metabolism in distinct clades, we will review the evidence sustaining that starch has evolved from preexisting glycogen metabolism several times during the evolution of photosynthetic eukaryotes and cyanobacteria. This review will also describe the possible function of storage polysaccharide metabolism in establishing metabolic symbiosis during plastid endosymbiosis. We will detail the evidence sustaining that storage polysaccharide metabolism was used by three distinct organisms to establish a tripartite symbiosis that facilitated metabolic integration of free-living cyanobacteria into evolving organelles.

Keywords Starch • Amylopectin • Glucan • Amylose • Glycogen • Photosynthesis • Endosymbiosis • Chlamydia

4.1 The Comparative Biochemistry of Starch and Glycogen Metabolism in Bacteria

4.1.1 The Bacterial Eukaryote and Archeal Domains

Three distinct domains are currently recognized among living cells. A consensus is slowly being reached that defines both the Archaea and the Bacteria domains as the most ancient organizations of life on Earth. Eukaryotes are presently

S. Ball (✉) • C. Colleoni • M.C. Arias
Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576 CNRS – Université Lille 1, 59655 Villeneuve d'Ascq Cedex, France
e-mail: steven.ball@univ-lille1.fr

Y. Nakamura (ed.), *Starch*, DOI 10.1007/978-4-431-55495-0_4

thought to branch from within the Archaea as a sister to a group to the so-called TACK superphylum consisting of the Thaumarchaeota, the Aigarchaeota, the Crenarchaeota, and the Korarchaeota (Williams et al. 2012). The appearance of the first living cells and biochemical pathways will not be considered in this review. We will focus on the function, evolution, and distribution of the glycosyl-nucleotide-based pathway of glycogen synthesis which is the most widely distributed pathway in this respect but are well aware that a number of alternatives have been described, notably in bacteria, where synthesis of glycogen from maltose, trehalose,, or sucrose can operate without glycosyl-nucleotide synthesis (Chandra et al. 2011). Eukaryogenesis, the process by which the first eukaryotes have emerged from within the Archaea, is a highly controversial issue and multiple conflicting hypotheses have been proposed (Martijn and Ettema 2013). With such unknowns to relate the eukaryotic pathways of glycogen metabolism to their bacterial and archaeal relatives; deduce the implementation of this pathway during eukaryogenesis, is a speculative and perilous exercise, that we shall not attempt. On the other hand, such an approach applied to plastid endosymbiosis is today entirely feasible, as there is a large consensus on the nature of the cells involved (at minima a standard biflagellated heterotrophic phagotroph and a free-living cyanobacterium). In this review, we will compare the nature and function of the classical glycosyl-nucleotide-based pathway of storage polysaccharide metabolism in several lineages and deduce, whenever possible, the evolutionary history of this biochemical network. Figure 4.1 represents the current status of eukaryote phylogeny, with a proposed root of the tree between the Excavata and the Amoebozoa. As seen from this bird's-eye view of eukaryote diversity, glycogen is uniformly present on the so-called unikont branch (see legend of Fig. 4.1 for a definition) of the eukaryotic tree, while either α-(glycogen or starch) or β-glucan (paramylon or soluble glucans) storage polysaccharides are evidenced in the bikont branch. While Archaeplastida are uniformly α-glucan accumulators, a mix of either α- or β-glucan accumulators are evidenced in the Excavata and the polyphyletic Hacrobia group, as well as the SAR superphylum. Starch has been clearly documented in all three Archaeplastida lineages. It is also found in many alveolates and in all cryptophytes. Cryptophytes define a subgroup of the so-called Hacrobia uniting them to the haptophytes, a group presently believed to be polyphyletic (hence, they do not define a real clade anymore). Alveolates are members of the SAR (which stands for Stramenopiles, alveolates, and Rhizaria) which contain two other major groups: the Stramenopiles and Rhizaria. Both of these are β-glucan accumulators and accumulate neither glycogen nor starch (with the noticeable exception of the Blastocystis glycogen-accumulating unicellular stramenopile gut parasite (Yoshikawa et al. 2003)). It must be stressed that β-1,3-linked storage β-glucans are widely distributed in the bikont branch of the eukaryotes (Fig. 4.1, see legend for definition of bikont) and that very few reports address the biochemistry of their synthesis or mobilization (Bäumer et al. 2001; Goldemberg and Marechal 1963; Tomos and Northcote 1978; Vogel and Barber 1968). Most soluble β-glucan forms, such as laminarin, chrysolaminarin, and mycolaminarin, contain, in addition, β-1,6 branches, while the very highly crystalline paramylon found in *Euglena* (an excavate unicellular green alga) and

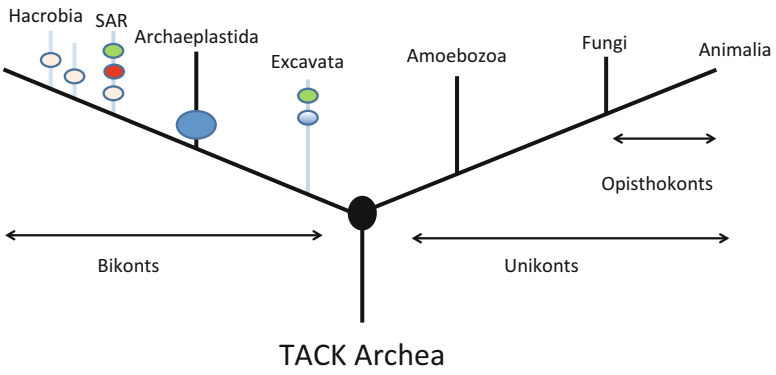

Fig. 4.1 Schematic representation of the eukaryotic domain. A simplified view of the phylogeny of eukaryotes is presented. Traditionally the eukaryotes are divided into the unikonts and bikonts which were previously thought to contain only uniflagellated cells or biflagellated cells. Although this has been brought into question, this division in two branches is kept for the sake of convenience and clarity. The tree is displayed as rooted within the Archaea (the TACK subdivision, see text). The most frequently proposed root is displayed by a *large-size black dot*, but other roots remain possible. Plastid endosymbiosis are displayed within the bikonts by oval-colored shapes. The *large-size blue-colored oval* represents the unique primary plastid endosymbiosis. The *small-size blue-colored oval* represents the Paulinella chromatophore primary endosymbiosis. All other small-size ovals display secondary endosymbioses. *Pink* represents secondary endosymbiosis of a red alga followed by fusion of the outer plastid membrane with the ER. *Red* represents secondary endosymbiosis of a red alga not followed by a fusion with the ER. *Green* represents secondary endosymbiosis of a green alga not followed by a fusion with the ER

possibly also in chlorarachniophytes (rhizarian unicellular green algae) display very high levels of crystallinity (Kiss et al. 1988). Interestingly, chlorarachniophytes are derived as *Euglena* from a β-glucan-accumulating heterotrophic phagotroph that internalized a green alga through secondary endosymbiosis. Strikingly, storage β-glucans in eukaryotes can thus be found, as is the case for α-glucans, in either water-soluble or solid crystalline form. The very few studies performed on paramylon metabolism suggest that such storage polysaccharides are synthesized from UDP-Glc by a glycosyltransferase (Bäumer et al. 2001). However, there is not one single report of a mutant, or antisense RNA or RNAi with decreased biochemical activity, associated with an alteration in β-glucan metabolism to confirm this reasonable hypothesis. To date, the sequenced genomes of all storage β-glucan-accumulating organisms have been proven to lack glycogen or starch metabolism.

As seen in Fig. 4.1, the bikont branch of the eukaryotes has experienced a number of distinct endosymbiotic events leading to acquisition of photosynthesis, eventually followed by photosynthesis loss in several clades. Bikonts have experienced one single primary endosymbiosis event and numerous secondary endosymbioses. Primary endosymbiosis of the plastid can be defined as the capture of a free-living cyanobacterium by a heterotrophic eukaryote phagotroph. There is one single major primary endosymbiosis that enabled eukaryotes to gain the ability to perform oxygenic photosynthesis. This event dated between 0.9 and 1.5 billion

years defines the founding event of the Archaeplastida clades (glaucophytes and red and green algae) (Shih and Matzke 2013; Yoon et al. 2004). All plastids in eukaryotic algae and plants are derived directly or indirectly from this event. Another much more recent primary endosymbiosis of marginal ecological impact is nevertheless known. This event involves a rhizarian cercozoan phagotroph and an α-cyanobacterium. However, to date, *Paulinella chromatophora* seems to define the sole offspring of this fairly recent, 60-million-year-old event, and no other eukaryotic alga seems to be derived from it (Bhattacharya et al. 2007; Marin et al. 2005). The Archaeplastida primary plastids have, in turn, been captured by other heterotrophic eukaryotic phagotrophs. Several events consisting of the capture of a red or a green alga have indeed been documented. These events will ultimately lead to four membrane plastids where the outermost membrane defines the phagocytosis vacuole while the second outermost membrane defines the plasma membrane of the internalized eukaryotic alga. The two innermost membranes represent the two membranes of the primary red or green plastids. Some secondary endosymbiosis-derived algae such as cryptophytes or the rhizarian chorarachniophytes still harbor nucleomorphs, which are the remnants of the red or green alga nucleus, between the 2nd and 3rd membrane, in a compartment known as the periplastidial space, that corresponds to the internalized red or green alga's former cytosol. This very complex endosymbiotic history of bikonts is further complicated by numerous instances of primary or secondary loss of photosynthesis that may be accompanied by loss of photosynthesis genes, reduction or loss of plastid DNA, and eventually plastid loss.

When approaching the evolution of storage polysaccharides metabolism in eukaryotes, it is of paramount importance to consider this endosymbiotic history for several reasons. First, this story may impact the structure of the network through acquisition of genes of endosymbiotic origin. Second, acquisition of photosynthesis may have depended chiefly on establishment of an optimal biochemical connection, at the very onset of the event. This optimal connection had to be established between two unrelated and disconnected networks. In other words, a free-living photoautotroph, internalized by phagocytosis, will not spontaneously reduce or increase supply of carbon to suit the host's demands and needs. One obvious way to overcome this would be to feed photosynthetic carbon into a host cytosolic molecular buffer that overcomes the disconnection between source and sink of carbon. Such a buffer is provided by storage compounds in the host cytosol, whose catabolism responds to the highly regulated host demands. Hence, storage lipids and polysaccharides are very high on the list of compounds that could play such a role and particular attention must be paid to evolution of their metabolism in this respect.

In this review we will show an ever-increasing amount of evidence sustaining that α-glucans have played such a molecular buffer function during primary endosymbiosis of the plastid. As far as secondary endosymbioses go, we presently have no evidence for or against such a function in establishment of the eukaryotic symbionts. A possible role of storage lipids in the early phase of the distinct symbioses remains in this respect a possibility worthy of further exploration. It must be stressed that the secondary plastids harbor distinct membrane arrangements. In most cases involving

a red alga symbiont, the first membrane which corresponds to the phagocytic vacuole membrane has fused with the ER. This in effect would place the 2nd plastid membrane called the periplastidial membrane adjacent and in contact with the ER lumen of the host. In many eukaryotes, the ER lumen contains the full suite of enzymes required for TAG biosynthesis (Beller et al. 2010). Hence, installment of a symbiotic flux, relying on export of symbiont lipids to the periplastidial membrane, would lead to TAGs budding off from the ER into the host cytosol. Such lipid bodies would offer a very convenient buffer to initiate secondary endosymbiosis and storage polysaccharides would not be per se required to trigger symbiosis. In the unique secondary endosymbiosis of a red alga that possibly generated the alveolates and in both cases of green alga secondary endosymbiosis (Euglena and chlorarachniophytes), the phagocytosis vacuole membrane did not fuse with the ER. Remarkably in all three instances, the storage polysaccharide (starch or paramylon) appears to have switched from a soluble to a solid crystalline form (although this is not entirely clear with respect to chlorarachniophytes), thereby suggesting that a sizable increase of the storage polysaccharide pool was selected in all three cases. We believe that this may suggest an essential function of storage polysaccharide metabolism in the metabolic integration of these secondary plastids. A lot more work and convincing evidence is nevertheless required before such a proposal is seriously considered.

4.1.2 Glycogen Synthesis in Bacteria

The paradigm of bacterial glycogen metabolism is defined by the ADP-Glc specific pathway, which has been abundantly studied in *E. coli* for several decades (Preiss 1984). This pathway relies on the production of ADP-Glc by ADP-Glc pyrophosphorylase, the product of the *GlgC* locus in *E. coli* which uses glucose-1-P and ATP as substrates to generate ADP-Glc and pyrophosphate. The single documented fate of this glycosyl nucleotide is to feed glycogen or starch synthesis in bacteria and Chloroplastida (green algae and plants). An additional function of ADP-Glc has been documented in cyanobacteria, where it was shown to be required for the synthesis of glucosylglycerol, an osmoprotective compound (see below). ADP-Glc is next used by a GT5 (CAZy classification glycosyltransferase family 5) glycogen synthase (GS) to generate an additional glucose bound at the nonreducing end of a growing α-1,4-glucan chain with concomitant release of ADP. Interestingly, Ugalde et al. have demonstrated that the GlgA GS of *Agrobacterium tumefaciens* is able to prime glycogen synthesis through autoglucosylation (Ugalde et al. 2003). This is important because, as we shall see below, the GT3 glycogen synthases from fungi and animals are unable to prime polysaccharide elongation and rely on the presence of glycogenin and autoglucosylating protein for this function. We presently do not know if the properties described for the *Agrobacterium* glycogen synthase apply to all other enzymes of this type in bacteria. Anyhow, no other mode of glycogen synthesis priming has been convincingly described in the bacterial

domain. The elongated glucans are then branched through branching enzyme (BE), which belongs to the GH13 group of glycosyl hydrolases (although other CAZy families have been documented in some bacteria including cyanobacteria) which is referred as the product of the *E. coli* glgB gene. BEs typically hydrolyse an α-1,4 chain and transfer a segment of chain in α-1,6 position. Catabolism is through the action of glycogen phosphorylase, the product of the glgP gene, which catalyzes orthophosphate-mediated phosphorolysis, yielding glucose-1-P from the nonreducing end of the glycogen particle. Glycogen phosphorylase typically stops digestion 4 glucose residues from any given α-1,6 branch or from any reducing end. Glycogen with short external chains of 4 glucose residues is called glycogen phosphorylase limit dextrin. These short external stubs are then released from the particle, through the action of the GlgX gene product, which directly hydrolyses the α-1,6 branch. For this reason this enzyme was named direct debranching enzyme. It is worth stressing that at least in *E. coli*, the GlgX enzyme is quite exacting and will not hydrolyze at significant rates glucans longer than 4 glucose residues (Dauvillée et al. 2005). This prevents the futile cycling of branches in the presence of BE, which transfers chains longer than 6 glucose residues. The maltotetraose released is then processed through enzymes of maltooligosaccharide metabolism, which typically include a combination of α-1,4-glucanotransferase and a specific form of phosphorylase called maltodextrin phosphorylase. The α-1,4-glucanotransferase, product of the *MALQ* (amylomaltase) gene, will hydrolyse out the reducing-end glucose from maltose and longer maltooligosaccharides and transfer the remaining oligosaccharide to other maltooligosaccharides, maltose itself being a very poor donor in this reaction (Palmer et al. 1976). The MalP gene product (maltodextrin phosphorylase) will recess back to maltotetraose all long glucans produced by amylomaltase. Some bacteria, like *E. coli*, contain two genes encoding distinct forms of phosphorylases acting selectively either on glycogen (glycogen phosphorylase) or on MOS (maltodextrin phosphorylase). Other bacteria contain a single bifunctional glucan phosphorylase. Other maltooligosaccharide-processing enzymes, such as α-glucosidase (product of the malZ gene product), act like MalP to avoid synthesis of longer MOS that would ultimately feed glycogen synthesis. In fact, in the absence of a functional MalP gene product, *E. coli* will very effectively synthesize glycogen even in the absence of functional glycogen synthase (Park et al. 2011). This suggests that an ADP-Glc-independent pathway of glycogen synthesis may be effective and functional in bacteria species feeding from MOS-rich media.

It must also be stressed that 21 % of bacterial clades investigated by Chandra et al. (2011) are able to use maltose-1-P as substrate for glycogen synthesis, through the use of α-1,4-glucan:maltose-1-phosphate maltosyltransferase, the product of the *GLGE* gene (note that GlgE is not present in *E. coli*). Over half of these clades are able to convert trehalose into maltose-1-P through the use of the product of *TRES* which converts trehalose into maltose and of Pep2 (maltokinase) which will phosphorylate maltose in the presence of ATP. This alternate pathway of glycogen synthesis is absent from *E. coli* and still relies on GlgB and the other genes described above of the classical pathway.

Finally while ADP-Glc remains, by far, the predominant glycosyl nucleotide used for glycogen synthesis in bacteria, there is one documented case of synthesis mediated through UDP-Glc in *Prevotella* (Lou et al. 1997). In this case, it must be stressed that LGTs of genes of glycogen metabolism from eukaryotic clades have been documented in related taxa (Arias et al. 2012).

The regulation of glycogen metabolism in bacteria is complex and exceedingly diverse. We will refer the reader to recent efforts addressing this topic and will make no attempts to review this (Wilson et al. 2010).

From our present knowledge of bacterial glycogen metabolism, we can confidently state that ADP-Glc pyrophosphorylase and ADP-Glc-specific GT5 GS are specific to bacterial glycogen synthesis, while debranching through a direct debranching enzyme and the ensuing coupling of the released chains to MOS metabolism is equally specific to bacteria (and in this case also to same degree to Archaea). The restricted eukaryotic distribution of such enzymes to Archaeplastida and some of their secondary endosymbiosis relatives, in face of the universal presence of the eukaryotic pathway (see below), leaves little doubt that these photosynthetic eukaryotic clades have gained the corresponding genes through lateral gene transfers from bacteria.

The classical bacterial glycogen metabolism network, as seen in *E. coli* and many other proteobacteria, is most often exceedingly simple and consists of 6–7 genes (*GLGC*, *GLGA*, *GLGB*, *GLGP*, *GLGX*, *MALQ*), depending on the specificity of the glucan phosphorylase with respect to MOS, organized in one or two operons, often in company of the *PGM* gene (phosphoglucomutase responsible for conversion of glucose-6-P to glucose-1-P) or eventually of *GLGE* (see above). In addition to this core pathway, some bacteria, including cyanobacteria, may contain additional isoforms for several steps of the network as well as additional enzymes from different CAZy families catalyzing analogous steps. This is especially true among the large-size cyanobacterial genomes which are of particular interest in this review.

4.1.3 Starch and Glycogen Biosynthesis in Cyanobacteria

In cyanobacteria and plants, oxygenic photosynthesis produces oxygen, NAD(P)H, and protons, which subsequently drive the conversion of ADP to ATP. Both NADP(H) and ATP will fuel the Calvin cycle to reduce carbon dioxide (CO_2). This primary carbon fixing reaction or carboxylase reaction is catalyzed by the ribulose-1,5 bisphosphate carboxylase/oxygenase (RuBisCO), which transfers CO_2 onto ribulose-1,5 bisphosphate producing two molecules of 3-phosphoglycerate (3-PGA). It should be stressed that this reaction is not favored in an aquatic environment since most of the CO_2 is found as bicarbonate (HCO_3^-) and carbonate (CO_3^{2-}) forms (Falkowski and Raven 2007). To circumvent this problem, cyanobacteria have evolved an extremely efficient CO_2-concentrating mechanism (i.e., active carbon transporters, carbonic anhydrases, carboxysome), which promotes the carboxylase reaction by enhancing CO_2 concentration in the vicinity of RuBisCO (for a review of this topic, see Badger and Price 2003).

Assimilated carbon is mainly stored as glycogen particles or starch granules in cyanobacteria and plants, respectively. Both homopolymers of α-D-glucose share the same α-1,4 and α-1,6 chemical linkages. Glycogen is predominately observed in most of cyanobacteria as tiny hydrosoluble particles between thylakoids membranes. Starch appears as nonaqueous semicrystalline granules in plants and several eukaryotic algae and is usually made of two α-polysaccharides: amylopectin and amylose. Until recently, starch granules were only described in Archaeplastida (plants, red alga, and glaucophytes) and other phylogenetically related organisms (Ball et al. 2011). However, Sherman's group performed the first description and partial characterization of abnormal carbohydrate granules in *Cyanothece* ATCC51142 (Schneegurt et al. 1994) but did not recognize this material as starch like. Later, a survey of storage polysaccharides in different species of cyanobacteria reported the presence of solid granules in other cyanobacterial species (Nakamura et al. 2005). Detailed characterization now indicates that such carbohydrate granules, including those of *Cyanothece* ATCC51142, are composed of a high molecular weight polysaccharide similar to amylopectin (Suzuki et al. 2013). More recently, *Cyanobacterium* sp. CLg1, a new strain isolated in the tropical North Atlantic Ocean phylogenetically related to *Crocosphaera watsonii*, accumulates starch granules made of both amylopectin and amylose (Deschamps et al. 2008a; Falcon et al. 2002). Interestingly, as in plants, the presence of amylose can be correlated to the identification of a polypeptide showing a high similarity in amino acid sequence to GBSS (granule-bound starch synthase) of plants. The latter is required for the synthesis of amylose in starch granules of various photosynthetic organisms (Delrue et al. 1992; Nelson and Rines 1962). So far, the GBSS gene is identified only in two cyanobacterial species, *Cyanobacterium* sp. CLg1 and *Crocosphaera watsonii*. Nevertheless, because GT5 ADP-Glc-specific glucan synthases are in essence prokaryotic and have never been found in eukaryotes and because archaeplastidal GBSS is related to bacterial sequences, it seems reasonable for us to propose that this type of sequence has evolved first among amylopectin-storing cyanobacteria.

Our understanding of the storage metabolism pathway in prokaryotes relies mostly on the studies in enterobacteria such as *Escherichia coli* (see preceeding paragraph for review). As pointed out earlier, the cyanobacterial storage polysaccharide network can be distinguished from that evidenced in most bacteria by an often surprisingly higher number of isoforms present at each stage of the pathway.

4.1.3.1 Biosynthesis

4.1.3.1.1 The Synthesis of ADP-Glc

In *Cyanobacteria* as in bacteria in general, ADP-Glc pyrophosphorylase (GlgC or Agp) regulates the first committed step in biosynthesis of α-glucan storage polysaccharide and for some species the synthesis of osmoprotectants. Like enterobacteria and plants, this enzyme mediates the synthesis of ADP-Glc, the sole nucleotide sugar used as substrate by the glucosyl transferase (glycogen(starch) synthase) that

belongs to the GT5 family (Cazy classication) to synthesize the α-1,4 linkages. AGPase is tightly regulated by allosteric effector defined by orthophosphate (Pi, inhibitor) and 3-phosphoglycerate (activator), which reflects the activity of the Calvin cycle. In contrast to plants, most of cyanobacteria strains possess one single glgC gene suggesting that AGPase is a homotetrameric enzyme. Nevertheless, a survey of genome sequences reveals the existence of two genes coding for glgC in *Cyanothece* ATCC51142, *Cyanothece* CCY01010, and *Acaryochloris marina* MBIC11017 (Fig. 4.2). To date, it is not clear if AGPase is active as a homotetrameric or heteromeric enzyme in these cyanobacteria. Phylogenetic analysis of the starch metabolism network of plants highlights the cyanobacterial origin of AGPase, which has evolved to maintain this cross talk with the Calvin cycle. Interestingly, the ancestral AGPase gene has been duplicated and subfunctionalized very early in the green lineage (the Chloroplastida), the small subunit considered as the catalytic part while large subunits fine tune AGPase activity. This could be correlated to starch relocalization in the chloroplast (Deschamps et al. 2008d). Because of its similarity to plant AGPase, the AGPase activity of Anabaena PCC7120 was intensively characterized by mutagenizing important amino acids responsible for allosteric effector regulation (Charng et al. 1992; Frueauf et al. 2002; Sheng and Preiss 1997).

In plants and green algae and to some extent in enterobacteria, the synthesis of ADP-Glc is exclusively committed to storage polysaccharide synthesis. Thus, mutants impaired in the synthesis of ADP-Glc grow normally in optimal growth conditions. In cyanobacteria, defects in glycogen biosynthesis lead to several physiological effects such as a decrease of photosynthesis activity, loss of viability, high salt sensitivity, and other defects (Grundel et al. 2012; Suzuki et al. 2010; Xu et al. 2013). The salt sensitivity phenotype results from a decrease of osmoprotective compounds: trehalose and glucosylglycerol (GG), which are synthesized by breaking down glycogen or by using ADP-Glc, respectively. Among the three different trehalose biosynthetic pathways reported in the literature, two are described in cyanobacteria (Wolf et al. 2003).

First, the production of trehalose from a glycogen/starch type of polysaccharide can occur through a three-step pathway: a glycogen debranching enzyme (TreX), which produces suitable maltooligosaccharides for a maltooligosyl trehalose synthase (Mts or TreY). The latter transforms the last α-1,4 linkage at the reducing end into an α-1,1 glucosidic bond. Maltooligosyl trehalose hydrolase (Mth or TreZ) will then hydrolyze the terminal trehalosyl unit of maltooligosaccharide (Higo et al. 2006). The reaction will produce trehalose and shorter maltooligosaccharides.

Second production of threhalose can occur from a starch-/glycogen-independent pathway in two steps. First trehalose phosphate synthase (TPS) will transfer the glucose residue of ADP-Glc or UDP-Glc to glucose-6-phosphate. Then trehalose-phosphate is dephosphorylated by trehalose-phosphate phosphatase (TPP). Interestingly, like TPS, cyanobacterial sucrose-phosphate synthase (SPS) is not specific with respect to the nucleotide-sugar substrate. In crude extracts of Scytonema, the specific activity of a TPS/SPS mixture is increased fivefold with ADP-Glc over that measured with UDP-Glc (Page-Sharp et al. 1999).

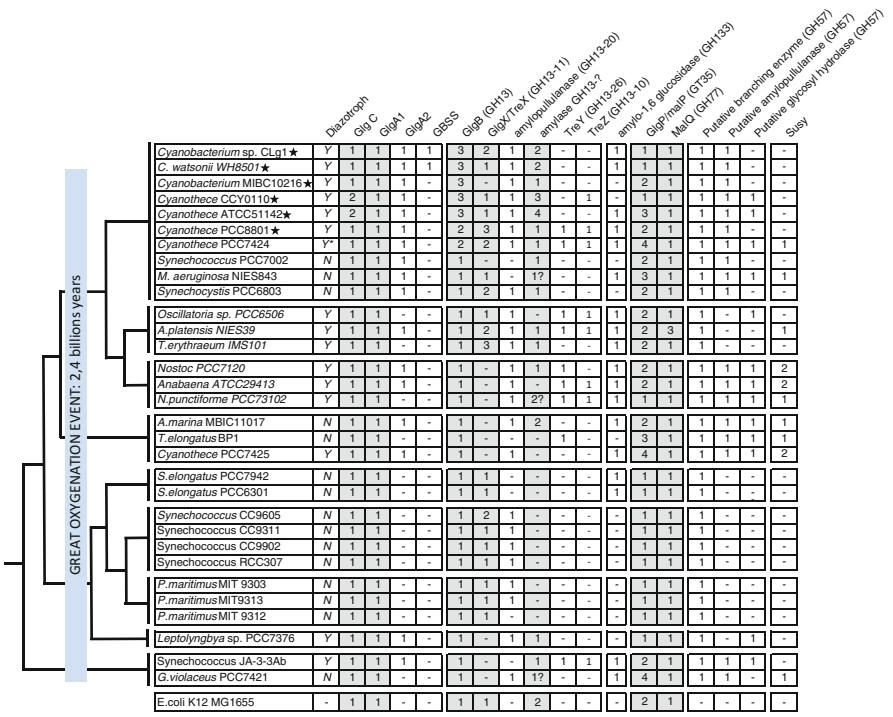

Fig. 4.2 *Black lines* depict the phylogenetic evolutionary relationship between cyanobacteria species in respect with the Great Oxygenation Event (GOE) which occurred around at 2.4 billion years ago. Representatives of each cyanobacteria clade were selected based on the availability of genome sequences. The number of enzyme isoforms found for each class of glycogen/starch and trehalose metabolism was determined using blast searches on NCBI, CyanoBase, and CAZy. Cyanobacteria strains labeled with a star were characterized to synthesize starch granules. *Y* (yes) and *N* (no) symbols reflect the ability to fix nitrogen in aerobic growth condition except for *cyanothece* PCC7424 which requires anoxic growth condition (Y*). *Escherichia coli* was used as reference for the type of isoforms found in cyanobacteria (*gray column*). Key: *BE* putative branching enzyme, *GlgC* ADP-glucose pyrophosphorylase, *GlgX/TreX* debranching enzyme, *iDBE* indirect debranching enzyme, *GlgB* branching enzyme, *GlgP/ MalP* glycogen phosphorylase/maltodextrin phosphorylase, *GlgA1/ GlgA2* glycogen synthase, *GBSS* granule-bound starch synthase, *MalQ* α-1,4-glucanotransferase, *TreY* trehalose synthase, *TreZ* trehalose hydrolase

Glucosylglycerol (GG) is the osmoprotectant compound synthesized only from ADP-Glc. GG-phosphate synthase catalyzed the first reaction using ADP-Glc and glycerol-3-phosphate. Then, a specific enzyme, GG-phosphate phosphatase will dephosphorylate GG-phosphate to produce GG. The characterization of *agp* (ADP-Glc pyrophosphorylase) mutants in *Synechocystis* PCC6803 shows clearly that in addition to the defect for glycogen biosynthesis, those mutants do not produce GG. The latter is compensated by large amount of sucrose after osmotic shock (Miao et al. 2003).

Interestingly, the presence of low amount of GG in the glgC mutant of *Synechococcus* PCC7002 suggests the existence of alternative pathway for the synthesis of ADP-Glc (Guerra et al. 2013). The first alternative source of ADP-Glc could come from sucrose synthase (Susy: EC2.4.1.13) activity. Susy activity is widespread in the plant kingdom and also found in filamentous heterocyst-forming cyanobacteria and in unicellular cyanobacteria. This enzyme catalyzes a reversible reaction in the synthesis of UDP-Glc and fructose by cleaving sucrose in the presence of UDP and vice versa. In plants, it is assumed that the major role of Susy is to supply UDP-Glc for cell wall synthesis. However, biochemical characterization of Susy of *Anabaena* PCC7119 highlighted major differences in the enzymatic properties (Porchia et al. 1999). Hence, cyanobacterial Susy activity displays a better affinity for fructose in the presence of ADP-Glc than UDP-Glc. In addition, Susy activity is sensitive to inhibition by ATP in both directions when UDP-Glc/UDP is used as the nucleotide-sugar/nucleoside substrate, while no activity was observed in the presence of adenine nucleotide. Recently, studies on the glycogen contents in mutants lacking Susy activity have shown a fourfold reduction in glycogen accompanied by an increase in ADP-Glc pyrophosphorylase activity in the knockout mutant of Susy (Curatti et al. 2008). In the presence of a nitrogen source, glycogen can be detected only in the mutant overexpressing Susy and not in the wild-type cells. Altogether, these data may suggest that Susy could contribute to the pool of ADP-Glc and consequently impact the pool of glycogen in *Anabaena*. However, a final but yet to be made convincing demonstration would consist of proving that GlgC mutants of *Anabaena* are not completely defective for glycogen accumulation.

A second theoretically possible source for ADP-Glc could consist of UDP-Glc pyrophosphorylase (GalU), which produces UDP-Glc from UTP and glucose-1-phosphate. A growing body of evidence suggests that the UGPase from *E.coli* or yeast may have some nonspecific activity toward ATP, generating ADP-Glc as an alternate product (Moran-Zorzano et al. 2007; Zea and Pohl 2004). However, the in vivo relevance of such speculations remains to be demonstrated.

4.1.3.1.2 Soluble Glycogen/Starch Synthase

Glycosyl transferases activities catalyze the α-1,4 linkages by transferring the glucose residue of ADP-Glc or UDP-Glc onto the nonreducing end of α-1,4-glucan chain. Two families of glycosyltransferase activities (GT3 and GT5) share this function among eukaryotes and prokaryotes (Coutinho and Henrissat 1999). The GT3 family is exclusively observed in the glycogen metabolism pathway of opisthokonts and of some amoebozoa and excavates, while the ADP-Glc utilizing GT5 enzymes are distributed in both prokaryotes and plants. It must be reminded that many other eukaryotic enzymes synthesize glycogen through a GT5 UDP-Glc-specific enzyme.

Another remarkable feature of glycogen synthase GT3 is their dependence on a self-autoglycosylation protein (glycogenin) to initiate glycogen synthesis (Torija et al. 2005). So far, there is no evidence of a comparable priming mechanism

in prokaryotes. However, as mentioned above, study of the *Agrobacterium* GS suggests that both priming and elongation activity are supported by this activity (Ugalde et al. 2003). Interestingly, starch synthase activities were subfunctionalized in plants for priming and elongation reactions. Thus, starch synthase III and IV are thought to be involved in polysaccharide synthesis priming, while starch synthase II and I are dedicated to the elongation of amylopectin chains (Szydlowski et al. 2009).

Except for the group of *Prochlorococcus-Synechococcus*, most of the cyanobacteria display two genes, coding for GlgA1 and GlgA2 isoforms (Fig. 4.2). Interestingly, the existence of both isoforms in the early cyanobacteria lineages might suggest that those isoforms were already present in the last cyanobacterial common ancestor. Although they were maintained through billions of years of evolution, it seems very likely that the *Prochlorococcus-Synechococcus* species, which emerged later in the evolution of cyanobacteria, have lost the GlgA2 gene during the genome reduction process. A detailed phylogenomic analysis of glycogen synthase reveals that glgA2 isoforms belong to a distinct clade composed of SSIII/SSIV starch synthase of plants and a dozen of α-proteobacteria. Interestingly, GlgA2 gene is restricted to a somewhat smaller group of bacteria (Chlamydiales, α-proteobacteria), whereas GlgA1 types of glycogen synthase are widespread among prokaryotes.

The respective functions of GlgA1 and GlgA2 activities have been recently investigated in *Synechocystis* PCC6803. Both single mutants of GlgA1 and GlgA2 produce unchanged wild-type amounts of glycogen (Grundel et al. 2012; Yoo et al. 2014). At first glance, these data suggest that there is an overlapping function of GlgA1 and GlgA2 activities. Both glycogen synthase activities are able to support both the priming and elongation reaction. However, structural characterization of the glycogen polysaccharides produced by single mutants is suggestive of different elongation properties (Yoo et al. 2014). Indeed, null *glgA2* mutants accumulate an altered glycogen structure, which contains fewer long glucans in comparison to the null *glgA1* mutant and to the wild-type reference strain. Based on the structural characterization of glycogen, the authors conclude that GlgA2 behaves as a processive enzyme while GlgA1 displays a distributive activity. Such conclusions have however not been confirmed directly by biochemical characterization of the purified enzymes.

Discovery of a GBSS gene in the genome of two cyanobacteria, *Crocosphaera watsonii* and *Cyanobacterium* sp. CLg1, defines a remarkable finding clarifying the origin of the enzyme of amylose synthesis. Phylogenetic analysis suggests a cyanobacterial origin for the GBSS found in Archaeplastida (see below). This gene gave rise by duplications to the soluble starch synthase I and II isoforms during the evolution of the Chloroplastida. GBSS has been demonstrated to define the sole enzyme required for amylose synthesis and its presence correlates perfectly with that of this starch fraction. In contrast to other starch synthase activities, this enzyme displays a unique elongation property, which consists to synthesize long glucan chains exclusively in the presence of semicrystalline polysaccharide (Maddelein et al. 1994).

4.1.3.1.3 Branching Enzyme Activity

Branching enzymes (GlgB or BE) – α-1,4-glucan: α-1,4-glucan 4-α glucosyl transferases – catalyze the formation of α-1,6 linkages in the storage polysaccharide. BE can be classified in three groups according to their enzymatic properties (e.g., preference for chain acceptors/donors) (Sawada et al. 2014). A survey of gene content in *Cyanobacteria* emphasizes the presence of two types of branching enzyme families: glycosyl hydrolase family 13 (GH13), which belongs to a large group of α-amylase family and glycosyl hydrolase family 57 (GH57) (Colleoni and Suzuki 2012). The latter is characterized by $(\beta/\alpha)_7$ barrel fold and five conserved domains important in the organization of the catalytic site (Santos et al. 2011). This family consists of diverse enzymatic activities (e.g., amylopullulanase, amylase, α-1,4-glucanotransferase, branching enzymes, and uncharacterized activities) and is restricted to Archaea and a few Eubacteria (Murakami et al. 2006; Zona et al. 2004). It is striking to observe that one putative branching enzyme GH57 gene is conserved in all cyanobacterial species, while the number of genes coding the GH13 branching enzyme varies from one to three (Fig. 4.2). Unlike the GH13 branching enzymes, the function of GH57 activities in starch or glycogen metabolism remains to be investigated. Nevertheless, a mutant of *Synechocystis* PCC6803 disrupted for its unique GH13 branching enzyme still accumulates half the wild-type amount of storage polysaccharide, while comparable mutants of yeast or *E. coli* witness a dramatic decrease or wipeout of glycogen synthesis. It seems likely that the GH57 branching enzyme is responsible for synthesis of the branches observed in the residual glycogen (Yoo et al. 2002).

The number of GH13 branching enzymes is variable among cyanobacteria. Phylogenetic analysis reveals three major classes among cyanobacteria: BE1, BE2, and BE3 (Colleoni and Suzuki 2012). The BE1 class is distributed in all cyanobacteria species, while the BE2 and BE3 classes are restricted to the order *Chroococcales*, which includes starch-synthesizing cyanobacteria. In plants, it is widely accepted that branching enzyme isoforms shape the starch granule architecture. Hence, the existence of distinct GlgB isoforms could in theory explain the structural diversity of storage polysaccharides among cyanobacteria. This reasonable assumption is apparently contradicted by the characterization of many starch-defective mutants in *Cyanobacterium* sp. Clg1 which all turned out to be defective for a GlgX-like starch debranching enzyme (Cenci et al. 2013). Had a particular BE been essential for starch rather than glycogen synthesis in this organism, mutants defective for this activity should have indeed appeared at either comparable or possibly lower but nevertheless detectable frequencies.

Characterization of branching enzyme activity is pretty scarce in the literature. A comparison of primary structure of *Synechococcus* PCC7942 reveals that the middle portion of the protein shares 62 % similarity with the *E.coli* branching enzyme, while the N-terminus displays little homology (Kiel et al. 1990). However, the N-terminus end has been reported to condition the length of the transferred glucan chain (Devillers et al. 2003). The lack of homology at the N-terminus could thus explain distinct chain-transfer patterns among branching enzymes. Recently,

branching enzyme activities of plants, yeast, human, and prokaryotes including GlgB of *Synechococcus* PCC7942 (BE1 class) have been subjected to detailed characterization. Based on transferred glucan chain preferences, data have shown that branching enzyme activities can be classified independently of their primary structure into three groups (see Sawada et al. 2014 for more details). From these studies it appears that both GlgB *E.coli* and *Synechococcus* PCC7942 belong to the group of OsBE1 (rice BEI), which introduce preferentially a new branching point on an acceptor glucan accordingly to the position of a preexisting branching point.

4.1.3.1.4 Convergent Evolution of Glucan Trimming Mechanism in Starch-Accumulating Cyanobacteria

As mentioned previously, glycogen particles define the most important type of storage polysaccharide among cyanobacteria. However, a small group of cyanobacteria belonging to the *Chroococcales* synthesize nonaqueous carbohydrate granules identified as starch granules (Fig. 4.2). A mutagenesis campaign was carried out on starch-accumulating *Cyanobacterium* sp. CLg1. Based on iodine staining of cell patches, over a hundred mutants were identified and subsequently categorized according to the ratio of soluble to insoluble polysaccharide. Among them, a dozen mutants harbor an increase in the water-soluble glycogen-like fraction and a disappearance of starch granules. This phenotype is correlated with a defect in direct debranching enzyme (DBE) activity belonging to glycosyl hydrolase family 13 (Cenci et al. 2013). Interestingly, biochemical characterization reveals that the mutants were impaired in the same type of ISA-GlgX activity as those involved in the crystallization process of starch in plants or in glycogen catabolism in *E. coli*.

The amylopectin aggregation in plants and probably in many others species relies on an isoamylase-type debranching enzyme activity (James et al. 1995; Kubo et al. 1999; Mouille et al. 1996; Wattebled et al. 2005). Like the CLg1 mutant, a defect in isoamylase activity in green algae and plants results in a substitution of starch granules by glycogen biosynthesis. Today, it is accepted that isoamylase removes selectively misplaced short glucans, which prevent amylopectin crystallization (Ball et al. 1996). Although starch-accumulating cyanobacteria predate the emergence of Archaeplastida, detailed phylogenetic studies of isoamylase enzymes failed to show the expected cyanobacterial origin of isoamylase. We can thus conclude that the cyanobacterial and plant enzymes have undergone convergent evolution to generate DBEs with similar properties. This remarkable finding suggests that there may be only one solution to generate starch from a preexisting glycogen metabolism network. We must nevertheless emphasize that some starch-accumulating cyanobacteria and the cryptophytes as well (see below) were reported to lack candidate DBE (Colleoni and Suzuki 2012; Coppin et al. 2004; Curtis et al. 2012). This raises the intriguing question of whether unidentified glucan hydrolase has evolved to ensure this function or whether combination of branching enzyme pattern could be sufficient to allow the aggregation of amylopectin (Streb et al. 2008).

4.1.3.1.5 Did the Great Oxygenation Event Trigger the Appearance of Crystalline Storage Polysaccharide in Unicellular Diazotrophic Cyanobacteria?

Cyanobacteria are one of the oldest phyla and most fascinating prokaryotes on Earth. Biomarkers (e.g., 2-methylhopanoid) and, more recently, geochemical analyses have traced back their oxygenic photosynthesis activity at least three billion years when the atmosphere was similar to anoxic environment (Crowe et al. 2013; Summons et al. 1999). It is widely accepted that the Great Oxygenation Event (GOE) occurred thanks to the release of oxygen from the splitting of water through cyanobacterial oxygenic photosynthesis 2.4 billion years ago (Kopp et al. 2005). The transition from a reductive to an oxidative environment triggered the diversification of cyanobacterial lineages and the appearance of new traits (e.g., morphology, cell size) (Blank and Sanchez-Baracaldo 2010; Latysheva et al. 2012; Sanchez-Baracaldo et al. 2014). A remarkable adaptation was achieved in nitrogen-fixing cyanobacteria. The reduction of dinitrogen to ammonium is catalyzed by an extremely oxygen-sensitive molybdenum-dependent ATP-hydrolyzing protein complex called nitrogenase (Burris 1991). Recent, phylogenetic reconstructions indicate that the last cyanobacterial common ancestor, which predates the GOE, was probably a freshwater nitrogen-fixing unicellular cyanobacterium (Larsson et al. 2011; Sanchez-Baracaldo et al. 2014). As oxygen rose, diazotrophic cyanobacteria have evolved to protect their nitrogenase activity (Bergman et al. 1997). In diazotrophic filamentous cyanobacteria, nitrogenase is localized in thick-walled specialized heterocyst cells lacking the PSII oxygen-generating complex. The large amount of energy required to fuel nitrogenase (16 ATP/N2) is supplied by neighboring cells displaying normal photosynthetic activity. Physical separation of two mutually exclusive biological processes, i.e., photosynthesis and nitrogen fixation, is possible only through the evolution of multicellularity and the ensuing cell specialization. For this reason, it was believed that nitrogen fixation could not occur in unicellular cyanobacteria. However, in 1970, two unicellular cyanobacteria strains, Gloeothece sp. and Cyanothece sp., were the first unicellular cyanobacteria described as aerobic nitrogen fixers (Singh 1973; Wyatt and Silvey 1969). In contrast to filamentous cyanobacteria that are able to fix nitrogen during the day, unicellular diazotrophic cyanobacteria perform nitrogen fixation exclusively at night. This group of unicellular cyanobacteria solved the problem by developing a temporal separation throughout circadian-clock regulation of the diurnal cycle. In the light, an accumulation of inclusion bodies is observed that disappears at night (Schneegurt et al. 1994). More recently, a survey and structural characterization of storage polysaccharide reveal that the inclusion bodies reported previously are composed of glucose residues forming semicrystalline polysaccharides similar to starch granules in plants (Deschamps et al. 2008a; Nakamura et al. 2005; Suzuki et al. 2013). This raises the question of whether the GOE was a driving force for the transition from glycogen to semicrystalline starch granules. It was noted that nitrogen fixation is engaged when the level of oxygen is low enough to allow nitrogenase

activity. To reach anoxia, unicellular aerobic nitrogen-fixing cyanobacteria exhibit high rates of dark respiration thereby consuming and decreasing O_2 levels locally in addition to ensuring the supply of the additional energy cost of N_2 fixation (Compaore and Stal 2010). Thus, unicellular diazotrophic cyanobacteria might have evolved a more efficient storage polysaccharide structure that maximizes the storage polysaccharide pool thereby allowing an efficient rate of nitrogen fixation in dark aerobic conditions. This hypothesis is reminiscent of the physiological adaptation of green algae to the anoxic production of hydrogen by hydrogenase. In that case it was indeed demonstrated that the predominant class of mutants recovered by screening for defective hydrogen production were mutants that had reverted from starch to glycogen accumulation because of a defect in isoamylase debranching enzyme activity! In addition, this assumption has been recently supported by the analysis of six Cyanothece species under various growth and incubation conditions (Bandyopadhyay et al. 2013); despite the fact that the authors did not notice that they have chosen starch- and glycogen-accumulating cyanobacteria in their experiment. It is striking for us to observe that those reported to accumulate starch as storage polysaccharide exhibit the highest rate of nitrogen fixation when compared to the glycogen-accumulating Cyanothece species in aerobic growth condition. Because we cannot of course exclude that genetic background may explain this variability in nitrogen fixation (e.g., Cyanothece PCC7525 requires anoxic condition to fix nitrogen), further analyses should be undertaken to correlate carbon limitation and specific rates of nitrogen fixation.

4.1.3.1.6 Physiological Importance of Storage Polysaccharide in Cyanobacterial Survival During Nitrogen Starvation

Glycogen biosynthesis is tightly regulated in response to different environmental cues, for instance, nitrogen starvation enhances glycogen accumulation in various bacteria species.

In non-diazotrophic cyanobacteria, limitation of the nitrogen source results in the storage of polysaccharide and a mobilization of the photosynthetic apparatus. This phenomenon named “chlorotic response” reflects a depletion of photosynthetic pigments (i.e., phycobiliprotein), which are used to supply transiently the cells in nitrogen. This nitrogen source allows protein synthesis before switching to a dormant state. Interestingly, a defect either in the ADP-Glc pyrophosphorylase or in both glycogen synthase activities results in a non-bleaching phenotype. Indeed, after a couple of days in depleted nitrogen medium, marine or freshwater cyanobacteria do not break down photosynthetic pigments and lose their viability (Grundel et al. 2012, 2013). Although the responsible signaling cascade is unknown, it is clear that the cyanobacterial survival relies on the presence of massive amounts of storage polysaccharide.

4.1.3.2 Catabolism

Photosynthetic organisms are subjected to alternating light and dark cycles. As mentioned previously, storage polysaccharide is synthesized during the day between thylakoid membranes. At night, cyanobacteria maintain their levels of ATP and NADP(H) by consuming storage polysaccharide. Several classes of catabolizing enzymes hydrolyze α-glucans into glucose and glucose-1-phosphate residues. Like in enterobacteria, glycogen phosphorylase (GlgP) and debranching enzyme (GlgX) work in synergy to cleave off the α-1,6 linkages. Glycogen phosphorylase releases directly a glucose-1-phosphate (G-1-P) from the nonreducing end of glucan chains and stops four glucose residues before the branch point. Short branched glucan are specifically cleaved off from phosphorylase limit dextrin by glycogen debranching enzyme GlgX. The number of isoforms involved in storage polysaccharide catabolism varies greatly according to the species. In cyanobacteria or *E.coli*, a defect in GlgX activity results in the accumulation of phosphorylase limit dextrin, which cannot be further metabolized (Colleoni and Suzuki 2012; Dauvillée et al. 2005). Recently, a screening for high temperature-sensitive strain of *Synechocystis* PCC6803 highlights a divergent function of both glycogen phosphorylases isoforms (Fu and Xu 2006). Surprisingly, GlgP-sll1356 is essential for growth at high temperature, whereas GlgP-sll1367 isoform is required for glycogen mobilization. Interestingly, despite the increase of GlgP-sll1356 activity in the null *glgP-sll1367* mutant, glycogen is inefficiently catabolized at night. Further investigations are required to understand the function of the GlgP-sll1356 isoform. However, we can hypothesize that sll1356 gene might encode a maltodextrin phosphorylase (MalP). This enzyme, involved in maltodextrin catabolism, predominantly acts on linear glucan oligosaccharides and is quite inefficient in its ability to digest the glycogen outer chains. Hence, like *glgP* mutants of *E.coli*, the MalP activity cannot substitute with the function of GlgP in the glycogen catabolism pathway (Alonso-Casajus et al. 2006).

Surprisingly, several cyanobacterial genomes, if complete, do not seem to encode any classical GlgX-type debranching enzymes. However, such genomes appear to contain amylopullulanase (GH13-20 and/or GH57) and amylo-1,6-glucosidase (GH133) genes (Fig. 4.2) that could supply such functions. Little is known about their role in α-glucan catabolism. Recently, amylopullulanase GH13 has been characterized from the filamentous cyanobacteria *Nostoc punctiforme* (Choi et al. 2009). In contrast to GlgX/TreX activities, this enzyme displays hydrolysis activity toward both α-1,6 and α-1,4 linkages when incubated with soluble branched polysaccharides (e.g., starch, amylopectin). The glucan chains made of 8 residues of glucose are specifically released and the resulting maltooligosaccharides are then hydrolyzed at the reducing end yielding shorter maltooligosaccharides ($G_{(n)} = G_{(n-1)} + G$). Because *N. punctiforme* does not contain any GlgX/TreX GH13 activity but harbors both maltooligosyl trehalose synthase (TreY) and maltooligosyl trehalose hydrolase (TreZ) (Fig. 4.2), it is quite possible that the GH13 amylopullulanase substitutes for the missing TreX activity, which is required to produce the suitable maltooligosaccharide for TreY. Amylo-1,6-glucosidase (GH133) shows similarity to the amino

acid sequence of the C-terminal domain of indirect debranching enzyme activity of animals and fungi. This activity could thus cleave out single α-1,6 glucose residues branched on glucan chains. So far, no biochemical characterization has ever been carried out on these cyanobacterial amylo-1,6-glucosidases. However, the characterization of the null GlgX mutant of *Synechococcus* PCC7942 brought some insights on the putative function of the amylo-1,6-glucosidase activity (Suzuki et al. 2007). Indeed, two genes encoding for candidate debranching enzyme activities, GlgX and amylo-1,6-glucosidase (GH133), have been identified in the genome of *Synechococcus* PCC7942 (Fig. 4.2). Although residual glycogen is enriched in short branched maltooligosaccharides, probably trimmed by the missing GlgX activity, the *GlgX* mutant of *Synechococcus* did not display a glycogen-excess phenotype comparable to that reported for the *E. coli* mutant (Dauvillée et al. 2005). Hence, it is entirely possible that the amylo-1,6-glucosidase may partially overlap with the GlgX activity in *Synechococcus*.

4.2 The Comparative Biochemistry of Starch and Glycogen Metabolism in Eukaryotes

A major difference between eukaryotes and bacteria is generally thought to be defined by a larger contribution in evolutionary histories of vertical inheritance and a lesser impact of horizontal gene transfer between unrelated clades. This is usually explained by the specific nature of gene sharing implied by the eukaryotic sexual cycle. The presence of a significant vertical inheritance component makes reconstitution of ancient pathways more feasible and less blurred by seemingly random gene sharing as often seen among bacteria. This in turn is required if evolutionary histories of biochemical networks are to be explained and enlightened by phylogenetic inferences. Hence, vertical inheritance usually leads to a higher level of congruence between diversification of eukaryotes and enzyme phylogenies. This assumption while still valid today must however be tempered by the presence of extensive HGTs observed as consequences of the very diverse and abundant endosymbiotic events. One may downplay the negative consequences of this type of HGT called EGT (endosymbiotic gene transfer) on the ease with which gene histories can be reconstituted. Indeed, the partners of the event are usually clearly identified and the events are also often well dated and placed within the eukaryotic tree of life. EGTs are defined as duplication of an endosymbiont gene that happens during metabolic integration of the endosymbiont followed by transfer to the host nucleus of the duplicated gene. This in turn is followed by expression and accurate localization of the gene product thereby allowing the corresponding loss of the endosymbiont gene copy. EGTs are evoked to explain the massive presence of cyanobacterial genes within the Archaeplastida (plant) nuclear genomes. At first glance this seems to only enrich gene histories without affecting our ability to reconstitute ancient pathways and recapitulate biochemical

network histories thereby enlightening gene function. Unfortunately life is indeed complicated and metabolic integration of endosymbionts does not always involve only the endosymbiont and host genomes. In fact in a majority of cases, the nuclear gene copy that replaces the endosymbiont genome's copy is not of endosymbiotic origin (and thus not a true EGT). This of course is not true for most (but not all) functions required for the maintenance and replication of the symbiont genome and for highly integrated process such as those governed by electron transport chains on evolving organelle membranes. However, for other less integrated but nevertheless numerous biological processes, the general rule concerning the origin of the gene that will replace the endosymbiont copy seems to be "whatever works!" Apart from chance (whoever gets in first!) a major factor governing the choice of a suitable source for endosymbiotic gene replacement by the nucleus could be the facility with which the product may be synthesized in the cytosol and translocated to the compartment where it will be active. What the endosymbiont does always transmit are the blueprints, the building instructions, for the novel biochemical pathways it encodes, but what the eukaryotic nucleus will choose as molecular tools to build this network is anyone's guess. In turn, this will create a lot of diversity and novelty in the constitution of the evolving organelle proteome. Fortunately for us the biochemical pathway of storage polysaccharide metabolism for most clades in eukaryotes has evolved in the eukaryotic cytosol and not in the endosymbionts from the preexisting cytosolic host glycogen metabolism pathway. As such, most of the starch metabolism enzymes are congruent with diversification of eukaryotes. This can be explained by the fact that the major pool of storage polysaccharides in all eukaryotes has always been, at least initially, exclusively cytosolic (Deschamps et al. 2008c). In the green algae and land plants, the plastidial localization of storage polysaccharides evolved after metabolic integration of the protoplastid had been partly achieved. The cryptophytes define the only case where the storage polysaccharide seems to have been maintained in the symbiont throughout the metabolic integration process of the evolving organelle. We will indeed see that in this case and in this case only, this has led to an astonishing patchwork of enzymes of distinct phylogenetic origins. Finally despite the ancient cytosolic origin of the storage polysaccharide metabolism pathways, we will see that this pathway has been indeed used to establish the first metabolic connection between endosymbiont and host at least in the case of the Archaeplastida.

4.2.1 Glycogen Metabolism in Opisthokonts

Opisthokonts define a very well-supported group uniting two major eukaryotic kingdoms (Fungi and Animalia) to a few of their unicellular ancestors; their name is derived from the presence of a unique flagellum propulsing flagellated cells from the posterior end of the cell (Cavalier-Smith et al. 2014). Glycogen metabolism in animals has been studied for nearly a century. These studies have yielded seminal contributions to our understanding of cell biology and biochemistry

in general. A review of the mass of studies that has allowed the understanding of glycogen metabolism regulation is out of the scope of this effort. Suffice it to say here that these studies have revealed the existence of glycogen itself, of cAMP, of protein kinases and phosphatases, of signal transduction, of glycosyl nucleotides themselves, and of their transferases, thereby leading to several well-deserved Nobel prizes (for review see Brautigan 2013, Wilson et al. 2010). The paradigm for opisthokont glycogen metabolism holds for both fungi and animals. It consists of glucan elongation from UDP-Glc through UDP-Glc-specific glycogen synthases, belonging to a CAZy family distinct from bacteria: the so called GT3 GS. These enzymes are notoriously unable to prime glucan elongation and require the presence of glycogenin. Glycogenin defines an autoglucosylating protein known to glucosylate selective tyrosine residues from UDP-Glc and then extend them to over 10 glucose residues (Albrecht et al. 2004). It is therefore believed that the GT3 GS uses this glucosylated primer to further extend glucans. In the absence of this primer, deregulated GT3 yeast GS is sometimes able to sustain synthesis of some glycogen in a stochastic fashion, leading to recurrent gain and losses of glycogen in the deregulated cells (Torija et al. 2005). As with bacterial glycogen synthesis (see above), the elongated glucan is branched through a GH13 glycosyl hydrolase named branching enzyme (BE). The mature glycogen particle is subjected to degradation through GT35 glycogen phosphorylase, whose basic structure and properties are analogous to the bacterial enzyme (see above). As with bacteria, the outer chains of the glycogen particle are recessed and stop 4 glucose residues from the next α-1,6 branch, generating thus glycogen phosphorylase limit dextrin. However, eukaryotes process this limit dextrin quite differently from bacteria. While bacteria directly release the short maltotetraose stubs through direct debranching enzymes, eukaryotes never produce such maltooligosaccharides. Instead, they use a bifunctional enzyme carrying two distinct catalytic domains involved first in the transfer of maltotriose from the branch to a neighboring chain within the same glycogen particle, thereby unmasking the remaining glucose at the branch, which in turn will be detached by the second α-1,6-glucosidase domain of the bifunctional debranching enzyme, which is therefore called indirect debranching enzyme (Nakayama et al. 2001). Through this mechanism, only glucose and extended glycogen chains, which define further substrates for glycogen phosphorylase, are produced. Hence, unlike bacteria or even archaea, no MOS are released by debranching, and therefore, there is no coupling to MOS catabolism in the eukaryotic cytosol. The presence of indirect DBE (iDBE) is distinctively eukaryotic, and no fusions of the two critical domains have ever been observed in prokaryotes. In addition to degradation by glycogen phosphorylase and iDBE, a hydrolytic pathway is known to operate either in lysosomes (animals) or in the fungal acid vacuole by glucoamylase or acid maltase (glucosidase). The mechanisms targeting and partitioning glycogen between the cytosol and the vacuole is an active domain of current research (for review see Wilson et al. 2010). Finally, a glucan phosphatase named laforin is known to be targeted in animals, but not in fungi, to abnormal glycogen particles in order probably to dephosphorylate them (Tagliabracci et al. 2007). The exact mode of action of laforin is presently controversial, but an

interesting possibility would be that a hypothetical glycogen kinase phosphorylates abnormal glycogen, to signal its targeting to the lysosome for degradation (Cenci et al. 2014). In the normal process of lysosome targeting and ensuing particle hydrolysis, a dephosphorylation step through this phosphatase may be mandatory. It is known that, in humans, a block at this step will have tragic consequences through the accumulation of hyperphosphorylated solid glucan bodies in the brain known as lafora bodies (Lafora and Glueck 1911). The patients suffering from the non-curable lafora disease die at a young age from epilepsy.

4.2.2 Glycogen Metabolism in Amoebozoa

Amoebozoa define a monophyletic lineage of unicellular or multicellular organisms placed among the unikonts but which in most cases define wall-less amoebal phagotrophs during their vegetative phase (for review see Glöckner and Noegel 2013). Two distinct genome sequences have been established within the monocellular amoebas: the Archamoebae (the human pathogen *Entamoeba histolytica*) and the Acanthamoebidae (*Acanthamoeba castellanii*). In addition to this, one multicellular model social amoeba genome (*Dictyostelium discoideum*) is also available among the Dictyosteliida (Glöckner and Noegel 2013). Some efforts in the 80s and 90s have been invested in dissecting glycogen metabolism in *Dictyostelium*. This has led to an in-depth biochemical characterization of the partially purified glycogen synthase activity and of two distinct glycogen phosphorylase isoforms (Rogers et al. 1992; Rutherford et al. 1992; Williamson et al. 1996). In the case of glycogen phosphorylase, the two genes encoding two distinct enzymes were selectively disrupted, and the consequences of these disruptions during the different phases of the complex Dictyostelium life cycle, were studied (Rogers et al. 1994). Disruption of Gp1 led to a 20-fold increase in the size of the glycogen pool at stationary phase in vegetative amoeba cultures. Gp2 was thought to be specifically involved in the conversion of glycogen to cellulose that occurs during sporogenesis. The gene disruption failed to demonstrate this, because of cross compensation through transcriptional and posttranslational regulation of Gp1. Interestingly, a similar function in the supply of glucose for cellulose synthesis was inferred for glycogen phosphorylase in *Acanthamoeba castellanii*. siRNA designed against Acanthamoeba's glycogen phosphorylase indeed resulted in a decrease of cyst wall assembly in this species (Lorenzo-Morales et al. 2008). A mutant defective for glycogen synthase has been selected in *Dictyostelium discoideum* (Tresse et al. 2008). In this species two glycogen synthase genes, belonging to the GT3 and GT5 families, have been reported. The candidate GT5 enzyme is fused to a truncated α-amylase domain, while the GT3 enzyme displays significant homology to the opisthokont enzyme. Disruption of the GT3 enzyme gene led to a threefold decrease of the glycogen levels, thereby suggesting that the GT5 GS ensures the synthesis of the remainder third of the glycogen pool. The mutants displayed delayed sporulation and were altered in spore stalk morphology because of abnormal autophagic cell

death of stalk cells. Glycogen synthase had been previously partly purified from wild-type cells and assumed to be of GT3 type, since at that time only one GT3 type of enzyme was thought to be present in the *D. discoideum* genome. Interestingly the laboratory of Charles Rutherford had described the presence of glucan primer-dependent and primer-independent forms, as well as that of a G6P-dependent and G6P-independent forms. The possible presence of the GT5 enzyme suggests that some of these properties may be carried specifically by either the GT5 or the GT3 forms, thereby requiring further experimentation. The gene content of the three distinct amoeba genomes is summarized in Fig. 4.3. The amoebas include all genes of the opisthokont glycogen metabolism network with the noticeable exceptions of bona fide opisthokont-like glycogenins and of laforin. The simultaneous presence of both GT3 and GT5 enzymes in *D. discoideum* and the presence of the sole GT5 enzyme in *E. histolytica* may suggest that a GT5 autoglucosylating enzyme could provide a primer for the GT3 enzyme, but this hypothesis also needs further experimentation. In addition to the opisthokont enzymes, the amoebas all contain β-amylase and the DPE2-like amylomaltase. Hence, amoebozoa contain a richer suite of glycogen metabolism enzymes than both fungi and animals. The analysis of these genomes strongly suggests that β-amylase and DPE2 that were previously thought to be plant-specific enzymes, required for starch degradation, are in fact very ancient enzymes of the eukaryotic glycogen metabolism network. β-amylase or true DPE2-like α-1,4-glucanotransferases are never found in bacteria, although some distantly related amylomaltases have been found in a few bacterial clades. We do not know however if β-amylase and DPE2 are cytosolic or lysosomal nor do we know how defects in such activities would impact glycogen content and metabolism. Because we can reasonably suspect that the opisthokonts have lost these two genes very early on, in their common ancestor, we believe that amoebozoa provides us with the paradigm of the ancient eukaryotic cytosolic glycogen metabolism network. It is most probably from such a cytosolic host network that starch metabolism evolved in Archaeplastida shortly after plastid endosymbiosis.

4.2.3 Starch Metabolism and Structure in Rhodophyceae

The red algae or Rhodophyceae (previously rhodophytes) are a large group of photosynthetic eukaryotes belonging to the super phylum Archaeplastida (Rodriguez-Ezpeleta et al. 2005). Over the past few years, phylogenetic analyses and genome sequencing have provided significant insights into the evolution of Rhodophyceae (Le Gall and Saunders 2007; Ragan et al. 1994; Yoon et al. 2006). Today it is widely accepted that Rhodophyceae define a monophyletic clade, currently composed of seven major classes: Bangiales, Florideophyceae, Porphyridiales-(1), Porphyridiales-(2), Porphyridiales-(3), Compsopogonales, and Cyanidiales. The latter consists of three genera Cyanidioschyzon, Cyanidium, and Galdieria and has received a lot of attention during the last decade. These unicellular red algae, with relatively small genomes (12 Mpb), live and thrive in particularly extreme

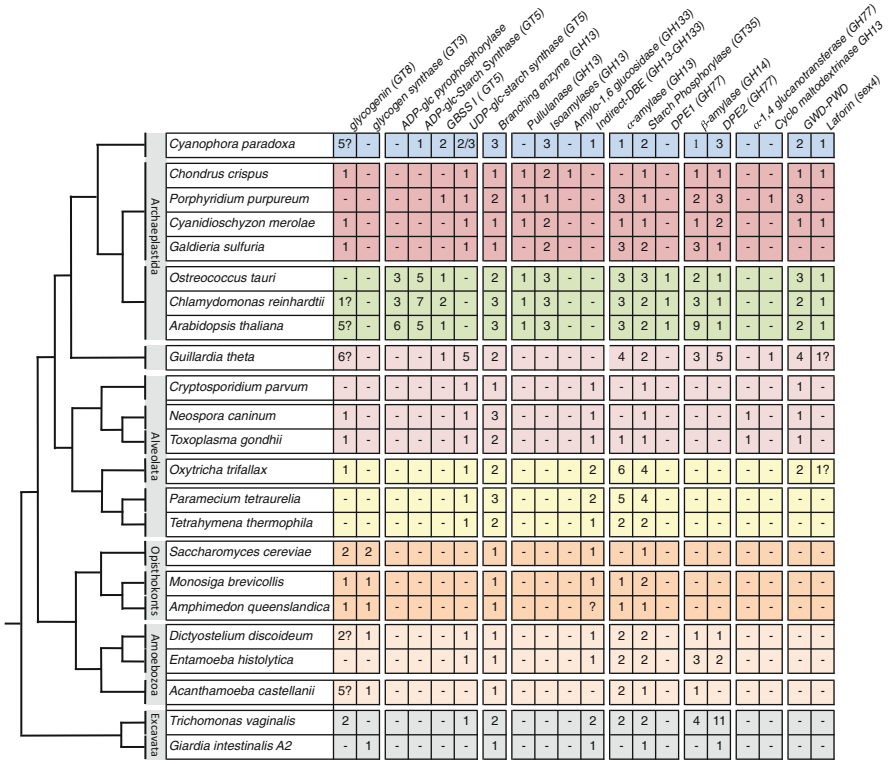

Fig. 4.3 Storage polysaccharide metabolism enzyme network in the eukaryotic domain. The number of enzyme isoforms is listed in each eukaryotic clade represented and listed according to the phylogeny of eukaryotes. Glaucophytes are represented in *blue*, Rhodophyceae in *red*, Chloroplastida in *green*, Cryptophytes in *dark purple*, apicomplexa in *light purple*, ciliates in *yellow*, opisthokonts in *pink*, amoebozoa in flesh, and excavates (metamonads) in *gray*. The number of enzyme isoforms found for each enzyme class of glycogen/starch metabolism was determined using blast searches on NCBI and CAZy

environmental conditions (pH 0.05–3; above 56 °C) unlike most eukaryotes. This remarkable adaptation required both genome reduction and extensive horizontal gene transfers (HGTs) from bacteria and archaea (Schonknecht et al. 2013). Based on phylogenetic studies, Cyanidiales are considered to be one of the most ancestral algae, which possibly diverged around 1.3 billion years ago (assuming a 1.5 billion-year-old date for plastid endosymbiosis) at the root of the Rhodophyceae (Yoon et al. 2002, 2004).

Like its green sister lineage (the Chloroplastida (previously known as "chlorophytes" or "viridiplantae" including green algae and plants), the photosynthetic carbon dioxide assimilation occurs in the plastid (i.e., rhodoplast) through the Calvin-Benson cycle. Subcellular localization of carbohydrate metabolizing enzymes of the oxidative pentose phosphate and glycolytic pathways appears to be essentially

identical with plants (Moriyama et al. 2014). Nevertheless, in contrast to the green lineage, assimilated carbon is efficiently exported to the cytosol by the triose-phosphate translocator (TPT) to promote cytosolic storage polysaccharide biosynthesis (Linka et al. 2008). Except for *Galdieria* sp. and *Cyanidium* sp. which have probably reverted from starch to glycogen synthesis (Hirabaru et al. 2010; Stadnichuk et al. 2007), most of the photosynthetic carbon is deposited in the cytosol of red alga as amylose-free starch granules called "floridean" starch (Borowitza 1978; Meeuse et al. 1960; Meeuse and Kreger 1954). Nevertheless, the absence of amylose fraction does not appear to define a universal feature of floridean starch. Extended survey of floridean starch composition highlights the presence of a significant amylose fraction in the genus Porphyridiales (McCracken and Cain 1981; Shimonaga et al. 2007). Floridean starch defines a semicrystalline polysaccharide whose crystallinity ranges from 16 to 36 % depending on red alga species (Shimonaga et al. 2008). The amylopectin of red algae contains significantly less long glucan chains (DP > 37) compared with that of plants. This CL distribution results in a semi-amylopectin structure similar to starch-accumulating cyanobacteria (Nakamura et al. 2005).

The access to four sequenced genomes within red algae, *Cyanidioschyzon merolae* (Matsuzaki et al. 2004), *Galdieria sulphuraria* (Schonknecht et al. 2013), *Porphyridium purpureum* (Bhattacharya et al. 2013), and a multicellular red alga *Chondrus crispus* (Florideophyceae) (Collen et al. 2013), have brought a complete overview of candidate genes required for the biosynthesis and breakdown of storage polysaccharides and have confirmed the UDP-Glc-based pathway in Rhodophyceae (for review Viola et al. 2001).

Compared to plants and green algae where 30–40 genes seem required, rhodophycean starch metabolism appears astonishingly simple, as less than 12 genes are required. For example, both the unicellular thermoacidophile *Cyanidioschizon merolae* and the multicellular mesophile *Chondrus crispus* contain only one starch synthase and one branching enzyme phylogenetically related to the BEI-BEII superfamily of plants. This seems sufficient to catalyze the formation of both α-1,4 and α-1,6 linkages (Sawada et al. 2014). Hence, Rhodophyceae are able to accumulate a complex semicrystalline floridean starch using a minimum set of three activities including the isoamylase-type activity responsible for the crystallinity of amylopectin.

In addition, the analysis of genome sequences clearly reveals the absence of genes encoding ADP-Glc pyrophosphorylase (AGPase) and ADP-Glc utilizing starch synthase in all red algae species. These genomic data conflict with previous biochemical characterizations, which have suggested the presence of both activities (Fredrick 1968; Nagashima et al. 1971; Sesma and Iglesias 1998; Sheath et al. 1979). This discrepancy may reflect the nonspecificity of UDP-Glc pyrophosphorylase and UDP-Glc-utilizing starch synthases of red algae toward ATP and ADP-Glc, respectively, as reported for others organisms (Deschamps et al. 2006; Moran-Zorzano et al. 2007; Zea and Pohl 2004).

The regulation of floridean starch biosynthesis is not well known in red algae. In contrast to plants, it seems unlikely that the amount of floridean starch is directly

controlled by the flux of UDP-Glc. Indeed, UDP-Glc serves as substrate for many enzymes, like trehalose-phosphate synthase, galactose-1-phosphate uridylyltransferase, or UDP-galactose-4-epimerase (Lluisma and Ragan 1998). This nucleotide sugar is involved in multiple biosynthetic pathways such as cell wall, isofloridoside, and floridoside synthesis (Pade et al. 2014). Interestingly, floridoside shares functions analogous to those of sucrose in plants: it is an osmoprotectant, potentially used to translocate carbon between cells, and represents a major soluble source of storage carbon (Ekman et al. 1991). Hence, the floridean starch pathway relies, just like that of eukaryotic glycogen metabolism, on the sole use of UDP-Glc. However, there is presently no evidence for or against the existence of posttranslational floridean starch synthase modifications (i.e., phosphorylation) analogous to those described for the opisthokont GT3 glycogen synthase.

Phylogenetic trees show that UDP-Glc starch synthases of red algae belong to the same GT5 family as those found in the Glaucophyta sister archaeplastidal lineage and also in other glycogen or starch-accumulating bikonts and amoebozoa derived or not from secondary endosymbiosis (Coppin et al. 2004). The latter is completely unrelated to the GT3 UDP-Glc glycogen synthase found in animals and fungi (Ball et al. 2011; Price et al. 2012). Only granule-bound starch synthase, the enzyme responsible for amylose synthesis within starch granules, shows a common cyanobacterial origin for all Archaeplastida (Cenci et al. 2014). Interestingly, the presence of a GT5 UDP-Glc starch synthase in red algae is correlated with the presence of a gene encoding a candidate glycogenin. However, such correlations are not universal features of other protist GT5 or GT3 glycogen (starch) synthases. It has been shown that the GT5 UDP-Glc starch synthase of *Gracilaria* sp. displays a significant glycosyl transferase activity toward amylose, whereas little or no activity was observed toward glucan chains made of 2–7 residues of glucose (Nyvall et al. 1999). Thus, it is reasonable to hypothesize that the glycogenin of *Gracilaria* sp. might prime the synthesis of longer glucan chains (>7) required for the starch synthase of *Gracilaria* sp. Nevertheless, the genome of *Porphyridium* sp. lacks glycogenin genes. So far, no biochemical characterization has been carried out to show whether this starch synthase has evolved to initiate directly the synthesis of glucan chains as reported for some glycogen synthases of prokaryotes (Ugalde et al. 2003). An alternative hypothesis would be that one of the multiple α-1,4-glucanotransferase activities (DPE2-like) may be responsible for the priming reaction.

Based on genome content, it seems that floridean starch catabolism is similar to the starch degradation pathway in green plants and algae. With the exception of the few glycogen-accumulating red algae such as *G.sulphuraria*, glucan water dikinase (GWD) or phosphoglucan water dikinase (PWD) and laforin genes are found in red alga genomes (Fig. 4.3). Laforin of red algae and glaucophytes have their glucan-binding and dikinase domains organized in a fashion similar to the animals and unlike that which is found in Chloroplastida (SEX4: phosphoglucan phosphatase). In plants, both GWD and PWD are required to initiate the first committed step in starch catabolism through the phosphorylation of semicrystalline starch granules at the C6 and C3 positions, while laforin is involved in cleaving

out the phosphate from phosphorylated glucans (Kotting et al. 2005, 2009; Yu et al. 2001). Interestingly, in both *C. merolae* and *Porphyridium* sp., the GWD activity appears sufficient to loosen the tight crystal packing of amylopectin and to allow an efficient degradation of amylopectin through actions of phosphorylases and both β-and α-amylase activities. Maltose molecules released by the β-amylase activity are probably further metabolized in the cytosol thanks to the action of DPE2. The latter defines an α-1,4-glucanotransferase which catalyzes the transfer of a glucose residue from maltose to an acceptor glucan. This disproportionation reaction allows then further actions of hydrolytic enzymes. Interestingly, some red algal species such as *Gracilaria sp*. possess an α-glucosidase activity belonging to the GH31 family localized in the rhodoplasts (Yu et al. 1999; Yu and Pedersen 1993). This enzyme defines an α-1,4-glucan lyase (EC 4.2.2.13) producing 1,5-anhydrofructose from the nonreducing ends of either long α-1,4-linked glucans or short maltooligosaccharides such as α- and β-maltose (Nyvall et al. 2000). The function of this 1,5-anhydrofructose is not clear in *Gracilaria*. It is known that 1,5-anhydrofructose is a highly reducing carbohydrate (Yu and Pedersen 1993). For this reason, Yu and collaborators have proposed that it might protect rhodoplasts against reactive oxygen species (ROS). However, the cytosolic localization of starch biosynthesis raises the question of the source of maltooligosaccharides in the rhodoplasts. To address this question, further investigations should be undertaken to prove whether maltooligosaccharides are synthesized in the rhodoplasts as proposed by Viola et al. (2001) or whether they are directly imported from the cytosol through an uncharacterized transporter.

4.2.4 Starch Metabolism and Structure in Glaucophyta

Glaucophytes represent one of the three clades that emerged from plastid endosymbiosis. Although the order of divergence of the Archaeplastida lineages remains a topic of recurrent debate, an ancient early divergence of the glaucophytes is generally favored. This correlates with many traits that are believed to be ancestral in these algae, as, for example, the presence of a muroplast with many cyanobacterial morphological features (including a peptidoglycan wall and carboxysome-like structures) (Facchinelli et al. 2013; Fathinejad et al. 2008; Pittenauer et al. 1993). *Cyanophora paradoxa*, the model glaucophyte species, defines one of the 14 biflagellated single-cell algae living in acidic freshwater ponds. Together with *Glaucocystis nostochinearum*, *Cyanophora paradoxa* is presently the only experimentally tractable organism. No sexual reproduction is known to occur in glaucophytes and only one genome sequence has been quite recently reported for *Cyanophora paradoxa* (Price et al. 2012). A haploid genome sequence of 70 Mbp is presently hypothesized for this vegetative possibly diploid species with around 28,000 putative protein-coding genes. This is two- to threefold more genes and a fivefold larger genome than those reported for unicellular or multicellular red algae. Hence, Glaucophyta may not have experienced the extent

of gene losses and genome simplification that occurred in the early diverging Rhodophyceae. In this context, one can view the exceedingly simple pathway of floridean starch synthesis in red algae as a highly derived character and the network evidenced in *Cyanophora paradoxa* as more ancestral. Remarkably, the *Cyanophora* genome contains the whole suite of genes that have been proven to be active for starch synthesis in Archaeplastida, with the only exception of ADP-Glc pyrophosphorylase. *Cyanophora* contains all the eukaryotic glycogen metabolism genes, including iDBE, as evidenced in *Dictyostelium discoideum*. The presence of bona fide indirect DBE came as complete surprise as multiple GLGX-ISA-like direct debranching enzymes are also evidenced in this genome. Some of the host-derived glycogen metabolism functions have remained simple such as α- and β-amylase and laforin, while others are present in multiple copies such as GT5 UDP-Glc starch synthases (2–3 genes), branching enzymes (3 genes), phosphorylases (2 genes), and DPE2-like amylomaltases (3 isoforms). Among the redundant enzymes we believe that multiple forms of both phosphorylases and BEs may have existed prior to plastid endodsymbiosis, as they do not present cases of more recent gene duplication followed by subfunctionalization. Hence, the multiple BE isoforms are, by no means, related to the BEI, BEII, BEIIa, and BEIIb subfunctionalizations that occurred specifically in the Chloroplastida lineage. In support of an early diversification of several enzymes in eukaryotes, we have indeed seen that glycogen phosphorylases often exist in distinct isoforms in amoebozoan genomes. On the other hand, some of the *Cyanophora* host glycogen metabolism network enzyme genes may have duplicated and subfunctionalized after plastid endosymbiosis. DPE2 amylomaltase-like genes usually present in 1 copy in glycogen metabolism networks are indeed present in multiple copies. This correlates with the presence of multiple copies of the GlgX-ISA-like DBEs. Because the eukaryotic glycogen metabolism network only contains DPE2 as an oligosaccharide-processing enzyme which is coupled to the presence of the maltose-producing β-amylase, we can hypothesize that duplication of DPE2 genes followed by subfunctionalization into enzymes with a less restrictive substrate specificity would have allowed metabolism of maltotetraose and longer glucans possibly produced by the multiple GlgX-ISA-like enzyme forms. This could also be the case when multiple DPE2s are evidenced in Rhodophyceae. These speculations are in agreement with the biochemical observation in *Cyanophora paradoxa* (Fettke et al. 2009) not only of a DPE2-like activity with high affinity to glycogen as in Chloroplastida but also of an enzyme form displaying lower affinity with respect to the glycogen acceptor. The *Cyanophora* genome also contains spectacular evidence of LGTs of chlamydial genes (see below). In phylogenetic trees there is even one of the chlamydial GlgX-ISA-like proteins that appears nested within the group of Chlamydiales, while two other copies are found together with the rhodophycean enzymes. The most spectacular finding in the analysis of the recently sequenced genome of *C. paradoxa* is no doubt the presence of a GT5 enzyme belonging to the SSIII-SSIV subgroup of ADP-Glc-utilizing archaeplastidal starch synthases (Price et al. 2012). In 2008, we speculated that the common ancestor of Archaeplastida probably contained an ADP-Glc-specific glucan synthase of the SSIII-SSIV family

(to account of the presence of such enzymes in the chloroplasts of green algae and plants) that enabled the polymerization of ADP-Glc into the cytosolic host storage polysaccharide stores (Deschamps et al. 2008b). Finding such an enzyme candidate sequence in the *Cyanophora* genome that moreover defines a common LGT from chlamydial intracellular pathogens further convinced us that Glaucophyta indeed defined the coelacanths of plastid endosymbiosis. Before the *Cyanophora paradoxa* genome had been sequenced, some biochemical investigations have been performed by Plancke et al. (2008) and by Fettke et al. (2009) with respect to the nature of the starch metabolism pathway and to its structure. Starch was shown to consist of high amylose (over 30 %) granules with an amylopectin fraction enriched in very small glucans displaying an otherwise bimodal chain-length distribution (Plancke et al. 2008). Multiple small-size cavities have been detected on the granules. These cavities too small to accommodate standard pyrenoid structures suggest that starch granules may be associated with specific cellular structures in the cytoplasm. The nature of such structures is unclear and deserves future attention. UDP-Glc-specific starch synthases had been clearly evidenced and studied as well as multiple forms of starch phosphorylases. One of the starch phosphorylase isoforms was found as a starch-bound 92 kDa protein. A bifunctional GBSS using both UDP-Glc and ADP-Glc was also found and studied (Plancke et al. 2008). However, the subsequent finding of two distinct GBSS genes in the *Cyanophora* genome may indicate that these properties may be carried differently by two distinct GBSS proteins. One important biochemical finding is that of a high mass multimeric isoamylase-like enzyme. In Chloroplastida such quaternary assemblies of isoamylase subunits are thought to be important for the insolubilization of soluble pre-amylopectin into insoluble starch (Ball et al. 1996). Although all archaeplastidal ISA-GlgX-like DBEs are derived from a single GlgX-like chlamydial ancestral protein, the diversification of chloroplastidal isoamylases into three distinct subunit enzymes is clearly posterior to redirection of the pathway to plastids. Hence, the two mutlisubunit enzymes have most probably evolved independently in Chloroplastida and Glaucophyta suggesting that such elaborate organizations are essential for efficient conversion of soluble to semicrystalline polysaccharides. The presence of a GBSS activity that displayed higher affinity (but lower activity) toward ADP-Glc correlates with a mixed inhibition of starch phosphorylase with lower concentrations of ADP-Glc when compared to UDP-Glc. These properties suggest either that these properties are fossils of a previous state when both glycosyl nucleotides possibly coexisted in the cytosol of the common Archaeplastida ancestor or that *Cyanophora paradoxa* still uses ADP-Glc today despite the absence of ADP-Glc pyrophosphorylase. In favor of the extant utilization of ADP-Glc is the finding of an SSIII-SSIV candidate sequence in the *Cyanophora* genome. Such enzymes indeed define glucan-elongating enzymes that exclusively use ADP-Glc both in cyanobacteria and Chloroplastida. All the results summarized above suggest that the glaucophyte network can be viewed as a living fossil of a very ancestral and interesting situation which has truly enlightened our vision of the evolution of starch metabolism. We believe that more experimental studies deserve to address many of the unknowns raised by these first reported genomic and biochemical studies.

4.2.5 *Starch Metabolism and Structure in Chloroplastida: Why So Complex?*

Starch metabolism in green algae has been the topic of many detailed studies and has been extensively reviewed (Ball 1998, 2002; Ball and Deschamps 2008; Busi et al. 2013). Thus, we will not further review this topic in this chapter. However, we believe that a comparison of the Chloroplastida network with those of their Glaucophyta and Rhodophyceae sister lineages remains of relevance here.

The most striking feature of the green lineage storage polysaccharide metabolism network is its very high level of functional redundance. A minimum of 30 and most often over 40 genes are thought to explain the complexity of starch metabolism in plants. Yet this astonishingly high number of isoforms has been generated from a less diverse number of ancestral genes than that which is present in extant Glaucophyta. This can be quite simply explained if one considers that storage polysaccharide metabolism was exclusively cytosolic in the common ancestors of the Archaeplastida and that all the genes of the cyanobacterial enzyme network with the exception of ADP-Glc pyrophosphorylase had already been lost from the cyanobiont genome (as is the case today for the chromatophore genome of *Paulinella chromatophora* (Nowack et al. 2008). The transfer of the starch network to the plastid was triggered by some unknown selection pressure. We have proposed that resistance to additional photooxidative stresses, generated by the evolution of the LHC antennae, adapted to high light intensities, might have generated a demand for a storage polysaccharide pool to prevent ROS synthesis or ROS-induced damage, due among others to chlorophyll biosynthesis (Deschamps et al. 2008c). Anyhow, the previous loss of storage polysaccharide metabolism genes from the cyanobiont would have considerably complicated the return of such enzymes to the plastid. Indeed, it is impossible to imagine that all the required genes from the cytosolic starch network would have simultaneously duplicated and simultaneously acquired plastid-targeting signal. Hence, it is impossible to imagine that rewiring starch metabolism to plastids could occur in one step, or only a few steps, in a way compatible with such steps being retained by natural selection. Fortunately, this complexity can be broken down into a series of manageable steps, if one envisions that the return of glucan metabolism to plastids was a very long process, encompassing two intermediate states, that may have lasted thousands or millions of years. Hence, we proposed that the evolving Chloroplastida first contained a major starch pool in the cytosol and a minor MOS pool in plastids (as may be the case for extant Florideophycidae, which have been convincingly shown to contain plastidial glucan lyase and hence must contain either glycogen or more likely MOS in the rhodoplasts) (Deschamps et al. 2008d). This may have lasted for a while, until natural selection may have favored the evolution of a larger pool of plastidial glycogen. During that time, starch was still accumulated in the cytosol. Finally, plastidial starch synthesis was favored over that of cytosolic polysaccharide synthesis. The interesting thing about these intermediate stages is that we know that the biochemical properties of enzymes that have become optimized to act either on

MOS or on glycogen or starch are quite different and exacting. Hence, to move from one stage to another, it would have been easier to duplicate a gene from the cytosolic starch metabolism network and subfunctionalize it rather than accumulate mutations within a single enzyme, to optimize it, with respect to several distinct situations. We believe that this very neatly explains the specific redundancy of the enzymes of the Chloroplastida network and have previously reviewed this topic in reasonable detail (Deschamps et al. 2008a, b, c, d). It is by this mechanism that the unique glaucophyte-like chlamydial SSIII-SSIV gene has duplicated and generated several isoforms finally yielding the SSIII and SSIV enzymes. This process also explains the generation of SSI and SSII from duplicated cytosolic GBSS gene copies, the generation of BEI and several BEIIs from the unique rhodophycean BEI-II-like enzyme, the generation of ISA2 and ISA3 from a unique ISA2-3-like cytosolic chlamydial ancestor, the generation of several BAM proteins from the unique rhodophycean gene, etc. In this process, some of the archaeplastidal enzymes appear to have resisted recruitment to the plastid. This could be due possibly to problems dealing with their targeting to plastids. Such enzymes might display a low suitability to interact with chaperones and to renature correctly in the plastid stroma. This may explain why, for instance, a DPE2-like protein was not recruited from the starch cytosolic network as an α-1,4-glucanotransferase and why, consequently, an LGT of DPE1 from a foreign proteobacterial source was favored. This may also be a reason why pullulanases are polyphyletic in Rhodophyceae and Chloroplastida. Resistance of laforin to targeting to plastids may also explain why a new gene was created with reorganized CBM and dikinase domains. Notwithstanding these exceptions, the green lineage starch metabolism network can be quite simply explained from recruitment from a subset of proteins evidenced in Glaucophyta and Rhodophyceae. In addition to DPE1, the only true novelty in the chloroplastidal starch metabolism network seems to be the MEX1 transporter (Lu and Sharkey 2006). This gene may have been created de novo during Chloroplastida evolution. On the other hand it may have been recruited by LGT from a cyanobacterial source since a single but significant case of MEX-like protein was found in these organisms (Weber, personal communication). Another striking observation in Chloroplastida as a whole is the astonishing level of conservation of isoform number and functions, from the earliest diverging picophytoplanktonic Prasinophyceae microalgal species to the highly sophisticated and comparatively much more recently diverging monocot plants. This also applies to Rhodophyceae when considering the evolutionary distance between Cyanidiales and *Florideophycideae*. In the prasinophyte *Ostreococcus tauri*, small cell size correlates with the simplification of most plant pathways, the only exception being starch metabolism. These observations suggest that unlike most of the redundancies witnessed today in land plant networks which can be attributed to the acquisition of multicellularity, starch metabolism complexity reflects a very early and specific event that occurred during evolution of the very first green cells. This complexity reflects that of the biochemical rewiring of pathways in novel cellular compartments, rather that any intrinsic complexity in the biochemistry of starch synthesis and mobilization.

4.2.6 Starch Metabolism and Structure in Cryptophytes

Cryptomonads are single-cell biflagellated protists, which live both in marine and freshwater environments. At the basis of the cryptomonad clade lies heterotrophic species belonging to the genus *Goniomonas*, with no evidence of a past phototrophic history (for review see Deane et al. 2002). These phagotrophic organisms have no plastids or starch and are not reported to contain storage polysaccharides. They are close to the so-called katablepharids, another clade of poorly described heterotrophic biflagellated phagotrophs. The cryptophytes are cryptomonads which have undergone secondary endosymbiosis with a red alga. Many cryptophyte clades are known to inhabit environments with low light and they are among those photosynthetic organisms that are able to fix carbon deep into the water column thanks, among others, to the phycobilin-containing pigments they have inherited from their red algal symbionts. Cryptophytes and chlorarachniophytes are unique among secondary endosymbiosis lineages, by the fact that the periplastidial space, which corresponds to the cytoplasm and nucleus of the alga symbiont, has not undergone the dramatic reduction evidenced in stramenopiles, haptophytes, euglenids, or alveolates. In both clades, a remnant of the nucleus of the engulfed alga, called the nucleomorph, remains with 50–200 genes located in three minichromosomes. All four genomes (nuclear, nucleomorph, plastid, and mitochondria) of the model cryptophyte species *Guillardia theta* have been sequenced (Archibald and Lane 2009; Curtis et al. 2012; Douglas et al. 2001; Douglas and Penny 1999; Grosche et al. 2014). Unlike all other secondary endosymbiosis lineages, cryptophytes accumulate starch in the periplast with no evidence for the presence of any kind of storage polysaccharide in the cytosol. Deschamps et al. (2006) and Haferkamp et al. (2006) have studied the structure biosynthesis and degradation of this periplastidial starch. The starch structure resembled that reported for *Chlamydomonas*. It contained amylose and an active GBSS that could use both ADP-Glc and UDP-Glc efficiently (Deschamps et al. 2006). The GBSS had properties analogous to the *Chlamydomonas* enzyme, but with a lower specific activity. The GBSS protein displayed a phylogeny consistent with the red alga ancestry of the plastid symbiont. It had a plastid-targeting bipartite topogenic signal, with both a typical signal and transit peptide. The transit peptide however did not start with the canonical phenylalanine thought to condition import of proteins in glaucophyte muroplasts and in Rhodophyceae rhodoplasts (Steiner et al. 2005). This phenylalanine was replaced by a serine in the GBSS sequence; as a result of that the protein, that had crossed the first two membranes, thanks to the signal peptide, was trapped into the periplastidial space (Haferkamp et al. 2006). The serine-containing transit peptide was nevertheless cleaved off. No trace of ADP-Glc pyrophosphorylase activity, protein, or gene could be found in *Guillardia* through enzyme or antibody assays or bioinformatic studies. Again consistent with the red alga ancestry of the symbiont, only UDP-Glc-specific soluble glucan synthases could be evidenced biochemically. We were thus expecting to find a gene encoding the unique characteristic GT5 multipurpose rhodophycean enzyme, upon annotating the *Guillardia theta* genome.

What we found, instead, was five distinct GT5 enzymes of an entirely different type. This type consisted of enzymes phylogenetically derived from the chloroplastidal GBSS-SSI-SSII superclade. The phylogeny rejected a red alga ancestry but was consistent with either a green alga GBSS or a green alga SSI or SSII source. Because we do not see any evidence for a green alga bona fide GBSS in *Guillardia*, we believe that SSI or SSII are the most likely sources for such enzymes. It must be remembered that SSI and SSII are, most probably, derived phylogenetically from the duplication of an ancestor green alga GBSS gene, during the reconstruction of storage polysaccharide metabolism, within the evolving green chloroplast. The soluble SSI and SSII glucan synthases in green algae and terrestrial plants clearly cannot sustain glucan synthesis on their own, because of their inability to prime polysaccharide synthesis. Furthermore, these soluble enzymes are thought to be distributive in their mode of action and to differ by their ability to yield different size classes of glucans. We would like to propose a plausible mechanism for their evolution.

After secondary endosymbiosis of the red alga symbiont, the latter encoded the unique and multipurpose GT5 rhodophycean enzyme within its nucleomorph genome. This defined a truly magical protein able to prime polysaccharide synthesis, to interact with rhodophycean BE, to produce glucans with chain-length distribution compatible with their insolubilization into amylopectin-like molecules, and to be targeted by the kinases and phosphatases that regulate storage polysaccharide metabolism in the eukaryotic cytosol. Most of these functions are taken over by four and three different specialized starch synthases, respectively, in the green lineage and glaucophytes. We have outlined previously that in Rhodophyceae, there has been and still is a selection pressure for the maintenance of small-size genomes and a low total number of genes. Once trapped within its host, the selection pressure that previously applied to the rhodophycean genome no longer held for the host nuclear genome. Why did the unique gene encoding the GT5 rhodophycean enzyme not transfer to the host nucleus? There are several possibly not mutually exclusive explanations to this difficult question. First, for some unknown reason the frequency of transfer of genes from the symbiont genome to the host nucleus may be significantly lower than the rate of transfer from other sources (viruses, transformation, phagotrophy . . .). These unknown reasons might also be those that explain the maintenance of the nucleomorph in cryptophytes. It was proposed that one possible reason would be the maintenance of the symbiont, at one or two copies per cell, which would preclude autophagy as a possible source for symbiont DNA (Curtis et al. 2012). This, at first glance, seems to be supported by the observation that plastid DNA transfer to the host nucleus can be reproduced experimentally in tobacco (Stegemann et al. 2003) which harbors several plastids per cell while it cannot be reproduced in *Chlamydomonas reinhardtii* which contains a single chloroplast (Lister et al. 2003). This is also supported by the recent observations made on insect bacterial endosymbionts (Husnik et al. 2013). In this context, those that will be transferred to the host nucleus are only those rhodophycean genes that cannot, for functional reasons, be replaced by genes from other sources. But even in that case, one cannot exclude that the actual source for the transferred rhodophycean sequence

may have been a piece of DNA coming from an organism highly related to the symbiont, present in the environment and entering the organism, through the same mechanisms as those from standard lateral gene transfers. Another reason which may have prevented LGT of the unique multifunctional GT5 rhodophycean enzyme would be the possible intrinsic inability, for this enzyme, if synthesized in the host cytosol, to follow the classical route of proteins targeted to the periplast through the bipartite topogenic signal. This route entails the ability for the protein to be taken in charge by defined chaperones, through the ER and through the periplastidial membrane, and then refold in the periplast. As far as we know, GT5 UDP-Glc-specific glucan synthases are only found in the cytosol of eukaryotes, and no genes encoding such enzymes have adapted to the targeting of their gene product in other cellular compartments. Thus, is it is entirely possible that the targeting of such an enzyme to the periplast would require several mutations in the gene sequence, further complicating its recruitment. The SSI-SSII chloroplastidial sequence offered several advantages for their recruitment in this context. The gene already carried a transit peptide and was adapted to protein targeting to plastids. It thus only needed to be fused downstream of a possible signal peptide to generate the BTP-containing protein. The enzyme had already been optimized with respect to its suitability in renaturing in foreign environments and to be taken in charge by different types of protein chaperones. In addition, being derived from GBSS, which is able to accommodate both types of nucleotide sugars in all three archaeplastidal lineages, the recruited sequence could have evolved fast into a protein able to use UDP-Glc as predominant or sole glycosyl-nucleotide substrate. Nevertheless, this change of selectivity and of environment is evidenced by the presence, in phylogenies, of a long branch supporting the *Guillardia* enzymes, within the chloroplastidal GBSSI-SSI-SSII superclade. The chloroplastidan enzyme was not able to replace, at once, all the functions played by the unique rhodophycean Swiss knife-like enzyme. We believe it initially played a minor, but significant, role by supplying different lengths of glucans to the otherwise monomodal, semi-amylopectin-like structure evidenced in floridean starch of Rhodophyceae. Because the previous selection pressures that maintained simplicity in the red alga genome were no longer operating on the secondary alga nuclear genome, natural selection favored the appearance of a green alga-like multimodal distribution of amylopectin chains and the LGT of the novel SSI-II-like synthase was selected. After selection, both genes were subfunctionalized. The chloroplastidan glucan synthase specialized, for instance, in the synthesis of a subset of chains and the rhodophycean enzyme specializing in the synthesis of chains of distinct lengths. During this process, random gene duplications occurring in the nuclear genome of the host would have sent additional forms of the chloroplastidan enzyme in the periplast, opening further opportunities for mutations, followed by subfunctionalization, to take over additional specialized functions, such as synthesis priming, interaction with branching enzyme, or synthesis of glucans of distinct lengths. The rhodophycean enzyme progressively retreated into a more and more specialized role, until the last nuclear gene duplication of the SSI-SSII-related gene covered the last specialized role of the GT5 rhodophycean

enzyme. The latter could thus now be lost from the nucleomorph genome and was definitively replaced by five separate SSI-SSII-like green genes.

In the case of Guillardia's GBSS, why was this enzyme not replaced by a green alga gene? One striking feature of periplastidial starch synthesis in cryptophytes is the physical association between starch and the pyrenoid across the two inner membranes of the secondary plastid (starch being periplastidial) that separate the starch granule from the pyrenoid surface. In green algae, starch is directly in contact with the pyrenoid in the stromal space of the chloroplast. The presence of GBSS in algae seems to correlate with a need for pyrenoidal starch synthesis (see Sect. 4.3.7). The synthesis of amylose by GBSS in red and green algae is thus adapted to metabolite concentrations which could be different because of a different cellular compartmentalization. GBSS is an enzyme of cyanobacterial origin, initially present in the cytosol of the Archaeplastida common ancestor (Deschamps et al. 2008a). A GBSS sequence was recruited in the ancestor green algae to be targeted to the chloroplast. There is thus no specific reason to suppose that the rhodophycean enzyme would not likewise be recruited by the host cell to be targeted to the secondary periplast. Hence, because of functional reasons, the red alga gene was favored over a green alga version. When we look at the branching enzyme genes in the *Guillardia theta* genome, we here again are surprised to find two distinct sequences unrelated to the single rhodophycean enzyme sequence. The rationale for the recruitment of such enzymes would be the requirement of diversification of BE activity when several distinct glucan elongation enzymes had been selected. The subfunctionalization of two single novel enzymes to suit the novel SSI-SSII starch synthases was favored over the highly selected red algal enzyme adapted to function with the rhodophycean sole glucan synthase. The two BEs are phylogenetically related to either chlorophyte sequences of rather restricted distribution (present in only a few green algae) or a group of cyanobacterial and alveolate sequences. Once again, recruitment of foreign enzymes seems either to have been favored or to have happened more rapidly. As far as phosphorylases are concerned, two sequences affiliated respectively to amoebozoa and glaucophytes and alveolates have been found in the genome. The predominance of enzymes of eukaryotic phylogeny, when looking at phosphorylases and branching enzymes, begs the question of the host status prior to secondary endosymbiosis. We know nothing about the possible accumulation of small amounts of either glycogen or β-glucans or both in the putative heterotrophic host ancestors exemplified by *Goniomonas* and katablepharids. Hence, what is presently seen as a complex set of HGTs might simplify into a vertical inheritance component, if such organisms were shown to contain glycogen metabolism genes. One must also remember that nearly all phagotrophs contain some genes of α-glucan catabolism as they need to digest α-glucan containing preys. In this respect, it is striking to note that at least one beta-amylase gene of *Guillardia theta* is phylogenetically related to a *Naegleria* enzyme. *Naegleria gruberi* is known to lack glycogen metabolism and to synthesize β-glucans. These enzymes evolved in these organisms to digest their preys. Other enzymes of starch catabolism defined by the GWD-PWD dikinases and by the DPE2 amylomaltases are found in the *Guillardia* genome. It is striking

to see that in both cases some isoforms display a clear red algal phylogeny (3 out of 5 DPE2 candidate sequences) and one PWD candidate sequence, while all others make a very convincing case for a green algal phylogenetic origin. The phylogenies are unambiguous and the aforementioned enzymes can be classified as green or red with confidence. From the preceding discussion, the cryptophyte starch metabolism network appears to be defined as a very complex patchwork of genes of distinct phylogenetic origins. A significant minority of enzymes seem to have been recruited from the symbiont rhodophyte starch metabolism network, while a majority seem to have been recruited from green algae. In addition, a significant component of eukaryotic glycogen metabolism networks is present in cryptophytes, which is unrelated to the classical archaeplastidal enzymes. Whether this defines another case of recruitment of foreign genes or whether such sequences were present in the heterotrophic ancestors of cryptophytes remains to be investigated. The most surprising discovery in the *Guillardia* genome as far as starch metabolism is concerned lies no doubt in the absence of candidate gene sequence for direct or indirect starch or glycogen debranching enzymes. Because the genome is well covered, reasonably assembled, and well annotated, it seems unlikely that such sequences would have escaped us (Curtis et al. 2012). However, a candidate direct debranching enzyme activity was evidenced biochemically but not characterized in detail. It remains possible that among the GH13 α-amylase-like sequences, some activities may, in fact, define a novel kind of debranching enzyme. *Guillardia theta* and cryptophytes may thus very well define a case where crystallization of amylopectin may not rely on previously described starch debranching enzymes.

4.2.7 Starch Metabolism and Structure in Alveolates

Alveolates define a monophyletic grouping of three major eukaryotic clades: the early diverging heterotrophic ciliates, the dinoflagellates, and the obligatory intracellular apicomplexa parasites. Alveolates are always unicellular. A majority of alveolates are heterotrophic (all ciliates, all apicomplexa, and many dinoflagellates). However, many dinoflagellates and all chromerids are considered as photosynthetic algae. There is a very complex and controversial endosymbiotic history in alveolates (Baurain et al. 2010; Cavalier-Smith 1999). The presence of a true nonphotosynthetic plastid in apicomplexa parasites and of red alga-derived photosynthetic plastids in chromerids (a clade intermediate between dinoflagellates and apicomplexa) and some dinoflagellate species (known as the peridinin plastid-harboring dinoflagellates) argues for a secondary "red" endosymbiosis at least in the common ancestor of dinoflagellates and apicomplexa (Delwiche 1999). However, if such a founding event indeed exists, the exact position of secondary endosymbiosis of a red alga remains controversial and some studies suggest that this event occurred in the common ancestor of all alveolates including the early diverging ciliates (Reyes-Prieto et al. 2008). The picture is further considerably complicated by the occurrence of several tertiary endosymbioses that occurred selectively in dinoflagellates

(tertiary endosymbiosis qualifies the capture by a eukaryotic phagotroph of a secondary endosymbiosis-derived alga). During these tertiary endosymbiosis, heterotrophic dinoflagellate phagotrophs themselves resulting most probably from secondary endosymbiosis, followed by plastid loss, or loss of photosynthesis, or both, have acquired plastids from other primary or secondary eukaryotic algae, leading to an astonishing diversity of plastids (for a review of dinoflagellate evolution, see (Wisecaver and Hackett 2011)). To further complicate the issues, both dinoflagellates and ciliates are able to transiently steal and benefit from plastids or algae ingested by phagocytosis, a phenomenon known as kleptoplasty (Serodio et al. 2014). In all alveolates, storage polysaccharides are found in the cytosol in the form of either glycogen found in majority of ciliates or of starch found in most dinoflagellates and apicomplexa. There are some exceptional cases where storage polysaccharide metabolism was lost altogether. This seems to have been the case in the malaria-causing *Plasmodium* genus and in related apicomplexa. Most other apicomplexa seem however to have retained starch metabolism in their cytosol (Coppin et al. 2004). Among ciliates, a glycogen structure was convincingly reported for *Paramecium aurelia* and *Tetrahymena* (Manners and Ryley 1952). However, paraglycogen, a polysaccharide described as amylopectin like, has been reported in more ancestral (holotrich) ciliates (Forsyth and Hirst 1953). Looking at the genome contents of *Apicomplexa*, a picture of a rather simple pathway emerges, characterized by one major GT5 UDP-Glc-specific glucan synthase candidate sequence of two to three branching enzymes of a single iDBE α-amylase, GWD, and starch phosphorylase. The starch of *Toxoplasma gondii* was shown to consist only of amylopectin, with a structure recalling the semi-amylopectin structures of cyanobacteria and Rhodophyceae, because of the monomodal type of CL distribution (Coppin et al. 2004). An enzyme preferring UDP-Glc over ADP-Glc could be evidenced in *T. gondii* extracts. Ciliates are characterized by a comparable simple synthesis pathway but with a significantly higher number of enzymes involved in glycogen catabolism. Phylogenetically speaking there are both overlaps and specificities in the distinct isoforms, when apicomplexa and ciliates are compared. For instance, ciliates have one of their branching enzymes in common with apicomplexa, but also display a ciliate-specific isoform related to Rhodophyceae and Glaucophyta, while apicomplexa also have an additional specific isoform, phylogenetically related to one of the three glaucophyte BEs. One of the distinctive features of the ciliate apicomplexa network is both the absence of direct DBE and that of β-amylase and the DPE2 amylomaltases. As far as dinoflagellates are concerned, the picture, as could be expected, seems to be much more complicated with respect to the number and diversity of enzymes, probably as a consequence of tertiary endosymbioses. In *Symbiodinium*, the dinoflagellate species living in symbiosis with corals and which, supposedly, contains the classical peridinin-containing plastids, the network detected in the partial genome sequence included direct DBEs such as pullulanases and isoamylases and β-amylases (Shoguchi et al. 2013). It is not clear though that the peridinin-containing dinoflagellate plastids define the original secondary

plastids of dinoflagellates. EST sequencing of heterotrophic dinoflagellates, such as *Crypthecodinium cohnii*, has indeed been shown to contain genes encoding proteins with plastid-targeting sequences (Sanchez-Puerta et al. 2007). Hence, it is entirely possible that dinoflagellates may have lost photosynthesis very early on and adopted the phagotrophic way of life. All known plastid-containing photosynthetic dinoflagellates may thus represent tertiary endosymbiosis events, and their biochemical network of starch metabolism must be considered under this light as suggested by Petersen et al. (2014).

Deschamps et al. (2008a, b, c, d) and Dauvillee et al. (2009) have performed a detailed study of starch synthesis and structure, in the early diverging heterotrophic dinoflagellate *Crypthecodinium cohnii*. They found a classical starch, which unlike apicomplexa, contained both amylopectin and amylose. The amylopectin displayed a polymodal CL distribution similar to cryptohyte or green alga starch (Deschamps et al. 2008b). *C. cohnii* starch was crystalline and displayed an A type of diffraction pattern like that of green algae and cryptohytes and unlike starch from *Cyanophora paradoxa* which was of the classical B type. The dinoflagellate GBSS was quite different from those found in cryptophytes, Rhodophyceae, Glaucophyta, and Chloroplastida. First, it contained a 110 kDa GBSS, twice as large as standard GBSS proteins. This was mainly due to a large extension at the N-terminus, which seems to be conserved in other dinoflagellates. This GBSS differed from all other enzymes of this type by the fact that it did not use ADP-Glc at detectable rates and therefore defined an exclusively UDP-Glc-utilizing enzyme. Zymogram analysis revealed several UDP-Glc-utilizing soluble starch synthases and a set of glycosyl hydrolases and transferases. *Crypthecodinium cohnii* had been previously described as a manageable genetic system where selection of mutants and genetic crosses are feasible (Tuttle and Loeblich 1974, 1977). Deschamps et al. (2008a, b, c, d) and Dauvillee et al. (2009) have selected mutants showing various altered starch accumulation phenotypes, by screening through iodine vapors. Mutants showing a selective defect in the UDP-Glc-specific soluble starch synthase zymogram pattern were isolated with reduced starch accumulation. The low starch phenotype co-segregated in crosses with the glucan-synthase defect. This suggests strongly that the bioinformatics-based speculations proposing a UDP-Glc pathway in the cytosol of several lineages, other than the green algae and plants, are likely to be correct. Another interesting feature of starch metabolism in *Crypthecodinium cohnii* is the timing of storage polysaccharide accumulation. In this heterotrophic dinoflagellate, starch typically accumulates in log-phase cultures, as soon as glucose is added to the medium, a situation very different from what happens in most heterotrophic bacteria and fungi, where glycogen most often accumulates at the transition between log and stationary phase. This, according to Deschamps et al. and Dauvillée et al., correlates with the phagotrophic nature of these marine organisms that have to deal with sudden bursts of reduced carbon in an otherwise oligotrophic medium (Dauvillee et al. 2009).

4.2.8 Glycogen Metabolism in Excavates

Excavates represent mostly unicellular flagellated eukaryotes harboring 2–4 flagellae and most often a conspicuous ventral feeding grove (hence the term excavates). Some amoebal forms are known, as well as a few multicellular slime molds. The excavates may not define a monophyletic group and are usually thought to have diverged close to the root of the eukaryotic tree. From the point of view of storage polysaccharides, among the two major excavate divisions, the amitochondriate metamonads and the discoba that contain discoid cristae within the mitochondria, only the former seems to accumulate glycogen as storage polysaccharide (Ladeira et al. 2005; Manners and Ryley 1955), while the latter (with organisms such as *Naegleria*, *Trypanosoma*, or *Euglena*) are reported to contain either β-glucans, such as paramylon, or mannans, or both. The metamonads have a glycogen metabolism network resembling that of Amoebozoa (see Fig. 4.3). It is most probably UDP-Glc based and contains, as in Amoebozoa, both a GT3 (found exclusively in *Giardia lamblia*) and a GT5 UDP-Glc-specific glucan synthase (found exclusively in *Trichomonas vaginalis*). Surprisingly, candidate-glycogenin-like sequences have been found associated with the *Trichomonas* network and not in the GT3 GS-containing *Giardia*. Both metamonad clades contain eukaryotic iDBE phosphorylase and BE. A DPE2-like amylomaltase candidate sequence is found in both clades, but convincing β-amylase candidate sequences usually associated with DPE2 are only found in *Trichomonas*. The corresponding enzyme sequences usually yield long branches in phylogenetic trees, complicating their positioning. Because of this and their parasitic way of life, caution must be taken with respect to the presence of possible LGTs in their storage polysaccharide enzyme networks.

4.3 The Evolution of the Starch Pathway in Archaeplastida

4.3.1 Reconstructing the Ancient Network of Starch Metabolism in Archaeplastida

If we compare the storage polysaccharide metabolism networks in Archaeplastida that are evidenced in Fig. 4.3 and if we inspect the corresponding phylogenetic trees, a picture emerges that agrees with the monophyletic nature of the Archaeplastida. There is a complete set of genes encoding the full suite of enzymes required for eukaryotic glycogen metabolism in Glaucophyta. This network resembles that evidenced in *Trichomonas* and Amoebozoa. It typically includes GT5 UDP-Glc-specific glycogen (starch) synthase, eukaryotic BEs and phosphorylase, iDBE, and unlike fungi and animals (but like excavates and Amoebozoa) DPE2 and β-amylase. In addition to this, it most probably contained the laforin phosphatase, which is not found as such in the excavates and Amoebozoa but which is present in the animal lineage. As pointed out, the Rhodophyceae display a comparable network which

however lacks the iDBE which was anyhow potentially redundant with the direct ISA-GlgX-like DBEs and was probably lost because of this. The Chloroplastida also lack the iDBE but further lack the UDP-Glc-specific GT5 glycogen synthase. All other eukaryotic enzymes, with the noticeable exception of DPE2, and one of the phosphorylase isoforms are located in plastids in the green algae and land plants, while the whole network including the storage polysaccharide is entirely cytosolic in Rhodophyceae and Glaucophyta. Glaucophyta and Chloroplastida also share GT5 SSIII-SSIV-like glucan synthases. These enzymes are of bacterial origin and in all bacteria examined to date use ADP-Glc as a substrate. In *Cyanophora*, the corresponding enzyme is most probably cytosolic, while the SSIII-IV starch synthases are plastidial in the green lineage. Also found in all three lineages is GBSS an enzyme with a cyanobacterial monophyletic origin. This enzyme is always associated to starch granules and is selectively responsible for amylose synthesis. In cyanobacteria and Archaeplastida, this starch synthase is unable to synthesize glucans as a soluble enzyme. Hence, the presence of a common single GBSS gene suggests that the last common ancestor of all Archaeplastida was a starch accumulator. This last common ancestor must thus have had the possibility not only to synthesize semicrystalline polysaccharides but also and most importantly to mobilize them. The presence of the bacterial ISA-GlgX-like direct DBEs, in all Archaeplastida, is in line with this observation, since these enzymes have been demonstrated, both in Chloroplastida and cyanobacteria, to be responsible for the aggregation of precursors of amylopectin synthesis into semicrystalline starch granules. In cyanobacteria, insoluble semi-amylopectin starch-like granules can be directly digested by starch bound phosphorylases. Eukaryotic glycogen phophorylases or β-amylases and their archaeplastidal derivatives are, however, unable to digest starch. Hence, one wonders why the cyanobacterial phosphorylase gene was not transferred to the eukaryotic nucleus for expression in the cytosol. In the last common archaeplastidal ancestor, a novel pathway of starch catabolism emerged (see Cenci et al. 2014 for detailed discussion). This made use of a preexisting pathway of glycogen phosphorylation that included an unknown glycogen kinase and the corresponding laforin phosphatase (see Sect. 4.2.1). The transition of glycogen to starch metabolism in the last common archaeplastidal ancestor was made possible by the replacement of the preexisting glycogen kinase, that was ineffective toward crystalline starch, by a novel enzyme resulting of a rather simple gene fusion (Cenci et al. 2014). This consisted of a fusion of a CBM (carbohydrate binding module) to a preexisting dikinase domain. This protein was able to drive the phosphorylation of amylopectin crystallites, thereby rendering the inaccessible hydrophobic crystals more hydrophilic and thus accessible to the eukaryotic enzymes of glycogen catabolism that were poorly adapted to deal with such structures. This gene fusion possibly generated the first glucan-phosphoglucan water dikinase, which defines the only true novelty that was made in the common ancestor of Archaeplastida (Cenci et al. 2014). Hence, the minimum number of genes that must have been contained in the last common ancestor of Archaeplastida, to explain, through vertical inheritance, the extant distribution of genes in Archaeplastida, consists of one GT5 UDP-Glc-specific glycogen (starch) synthase, one SSIII-SSIV ADP-Glc-specific glucan

synthase, one ADP-Glc pyrophosphorylase subunit, two branching enzymes, one phosphorylase, one β-amylase, two DPE2-like amylomaltase (one supplying the DPE1-like function), two ISA-GlgX (one supplying the isoamylase function), one iDBE, one GBSS, one PWD-GWD-like protein, and one laforin (Deschamps et al. 2008c). ADP-Glc pyrophosphorylase defined the only protein which we suspect to have been active in the evolving plastids. Indeed, this enzyme of cyanobacterial phylogenetic origin is tightly coupled through its substrates and allosteric effectors to the Calvin cycle. Since the Calvin cycle never left the plastid stroma, it is reasonable to assume that the enzyme remained in this compartment. We have proposed that, apart the maintenance of ADP-Glc pyrophosphorylase, most other cyanobacterial enzymes of storage polysaccharide metabolism were lost in the common Archaeplastida ancestor. This proposal is justified through several distinct observations:

1. Today in both Glaucophyta and Rhodophyceae, there are no traces of storage polysaccharides in the plastid.
2. In general a correlation has been established between the obligate intracellular nature of bacteria and the absence of storage polysaccharide metabolism (Henrissat et al. 2002). A remarkable exception to this rule is defined by the Chlamydiales intracellular pathogens. The cyanobiont would thus be expected to lose its ability to synthesize storage polysaccharides early on. This has in fact happened in the chromatophores of *Paulinella* where no traces of starch, glycogen, or glycogen metabolism genes can be found (Nowack et al. 2008), while the α-cyanobacterial free-living ancestors of chromatophores contain the full suite of cyanobacterial glycogen metabolism genes.
3. The complexity of starch metabolism in Chloroplastida can be easily explained, if the cyanobiont had lost its glycogen metabolism genes. Had the evolving cyanobacteria kept its glycogen metabolism pathway, the replacement of the cyanobiont enzymes by eukaryotic versions would not have been as systematic, nor would it have entailed such dramatic duplications and subfunctionalizations. One would have expected at least partial conservation of some cyanobacterial enzymes through HGT to the nucleus, as well as occasional replacement of cyanobacterial enzyme versions, by enzymes from other bacteria, which would be better adapted to the prokaryotic pathway of storage polysaccharide metabolism.
4. The efflux of carbon from the cyanobiont and cyanobacterial storage polysaccharide synthesis are competing processes. The early transfer of the chlamydial ATP translocator to the three archaeplastidal lineages, by HGT, argues that there was an early need for ATP in darkness in the evolving cyanobiont. Such a need would not have existed, if the latter had substantial storage polysaccharide pools.

From all these considerations, we hypothesize the presence of a cytosolic network of starch metabolism (Fig. 4.4b) in the last common ancestor of Archaeplastida with no significant pool of polysaccharides synthesized by the cyanobiont. The presence of the bacterial SSIII-SSIV glucan synthase in the cytosol of the last common ancestor suggests that this ancestor fed both ADP-Glc and UDP-Glc into starch metabolism. This implies that, to be used in the ancestor's cytosol, this bacterial specific glycosyl nucleotide synthesized by the cyanobiont must be

exported to the host cytoplasm by a suitable transporter. For this reason, we have inferred the presence of a nucleotide-sugar translocator in the inner membrane of the evolving plastid.

The evidence that we have concerning the existence of an ancient ADP-Glc translocator in the last common ancestor of the Archaeplastida comes from the phylogenetic analyses performed by Weber et al. (2006) and the functional studies reported by Colleoni et al. (2010). Indeed, Weber et al. (2006) reported that all extant carbon translocators of the major pPT family that are present on the inner membrane of plastids of Rhodophyceae and Chloroplastida can be traced back to a unique ancestral protein, which is derived from a eukaryotic (host) *n*ucleotide-*s*ugar *t*ranslocators (Weber et al. 2006). These purine sugar nucleotide transporters that include Golgi GDP-mannose translocators were demonstrated to transport ADP-Glc, as fast as their normal substrate, but with lower affinity (Colleoni et al. 2010). In addition to this, the most studied extant pPT carbon translocator (the triose-phosphate translocator) has been previously demonstrated to be able to reach the mitochondria's inner membrane when expressed in yeast without its transit peptide (Loddenkotter et al. 1993). This remarkable property if it was displayed by the ancestral protein would have greatly facilitated its recruitment at a time where protein targeting machineries to the evolving plastid had not *yet appeared*.

4.3.2 In the Beginning There Was . . . Glycogen?

The reconstruction proposed in Fig. 4.4b defines a status of storage polysaccharide metabolism compatible only with the presence of starch in the eukaryotic cytosol. This does not define the status of storage polysaccharide at the onset of plastid endosymbiosis, but rather that which characterized the last common ancestor of the three Archaeplastida clades. This last common ancestor already had common elements of the machinery targeting proteins to plastids and defined a significantly advanced stage of metabolic integration of the evolving cyanobiont. By looking at the extant diversity of heterotrophic eukaryotes, there is no evidence for the accumulation of starch in clades other than those derived from primary or secondary endosymbiosis of the plastid. Hence, at the onset of plastid endosymbiosis, we can safely assume the presence of a slightly more simple network than that displayed in Fig. 4.4b. This network yielded glycogen rather than starch in the cytosol. Hence, there was no immediate need for glucan nor phosphoglucan water dikinases, no need for GBSS which is only active when associated to starch granules, and no need for multiple forms of ISA-GlgX-like proteins. This status of storage polysaccharide is displayed in Fig. 4.4a. Nevertheless, we believe that the transition from glycogen to starch metabolism was rapidly favored because of the huge amounts of photosynthetic carbon that became available to the Archaeplastida ancestor. This transition was probably facilitated by the presence, at the onset of plastid endosymbiosis, in the cytosol of an ancestral GlgX direct DBE. This activity as discussed in this review is not a natural component of the cytosolic glycogen metabolism machinery, and its presence in the eukaryotic cytosol at the onset

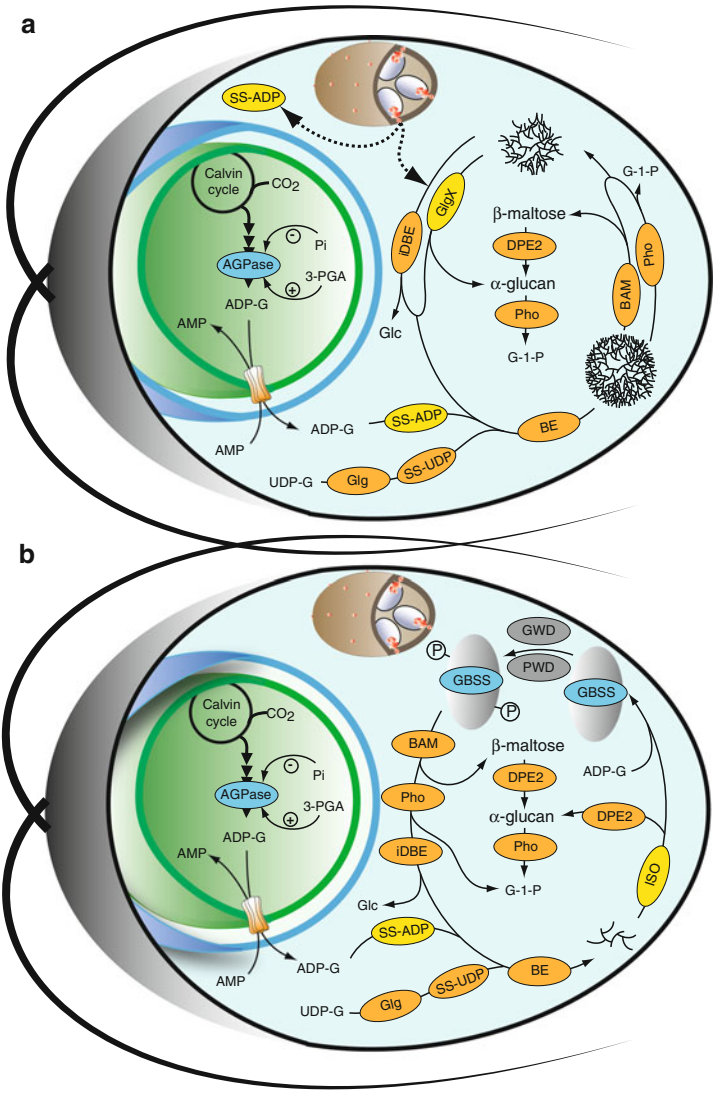

Fig. 4.4 Metabolic reconstruction of storage polysaccharide metabolism at the onset of plastid endosymbiosis (**a**) and in the last common Archaeplastida ancestor (**b**). The reconstruction was built by assuming vertical inheritance of those enzymes displaying a common phylogenetic origin and shared between the three Archaeplastida clades (Deschamps et al. 2008a). The number of enzymes was minimized according to the suspected gene duplications and subfunctionalizations which occurred selectively in Chloroplastida. ADP-glucose pyrophosphosphorylase and SSIII-SSIV (displayed as SS-ADP) are the only enzymes which are respectively found only in Chloroplastida and Glaucophyta. The reconstruction yields the network (**b**) as it was possibly operating in the last common ancestor of Archaeplastida before the three lineages diverged. This network produced starch as evidenced by the common presence of GBSS. To derive the glycogen metabolism network that was operating at the onset of endosymbiosis (**a**), we deleted those enzymes that are specific to crystalline polysaccharide metabolism (GBSS, GWD-PWD).

of plastid endosymbiosis can be deduced by the cytosolic effector nature of this protein that is phylogenetically derived from Chlamydiales intracellular pathogens. We have detailed in the preceeding paragraphs the observations that establish that GlgX proteins had been recruited prior to plastid endosymbiosis to enable some cyanobacteria to synthesize starch-like polymers in a polyphyletic fashion. Hence, this subfamily of the large GH13 family of glycosylhydrolases defines the perfect candidate to provide the switch from glycogen to starch metabolism. In a similar fashion, duplication within chlamydial symbionts of the gene encoding the GlgX chlamydial effector protein would have enabled novel effector subunits to provide the multimeric isoamylase function. Cenci et al. (2014) have shown through phylogenetic analysis that both in Glaucophyta and Rhodophyceae, the gene encoding the ISA-GlgX protein had indeed been duplicated. It is tempting to speculate that this happened, as the gene was still encoded by the chlamydial symbiont, further strengthening the interdependence of the ménage à trois (see Sect. 4.3.5 for definition) at plastidial endosymbiosis. The duplication and evolution of isoamylase, from the ancestral *GlgX* chlamydial gene, would have required the dikinase CBM fusion to optimize the mobilization of the novel form of carbohydrate stores. LGT of the cyanobacterial GBSS gene would have quickly followed, as amylose synthesis provides a convenient additional storage mechanism, in case the amylopectin synthesis machinery is overflowed by carbon. The successful EGT of the GBSS cyanobacterial gene, for expression in the host cytosol, suggests all this happened fast after the onset of plastid endosymbiosis, before all the genes from the cyanobacterial network of glycogen-starch metabolism had been lost. The invention of the glucan-phosphoglucan water dikinases offered, in addition, the possibility to integrate a novel first step of starch mobilization which replaced the direct access of

Fig. 4.4 (continued) The GWD-PWD are shaded in grey because they do not define pre-existing proteins of eukaryotic or bacterial phylogeny but rather a post-endosymbiosis gene fusion involving two domains of undefined origin (see text). Those enzymes phosphorylate starch granules by adding phosphate residues to the C6 and C3 positions of a glucosyl unit. Most of the enzymes are phylogenetically derived from the eukaryotic glycogen metabolism network (displayed in beige) which is essentially complete and operating in the host cytosol as is the case in all heterotrophic eukaryotes documented. Enzymes of cyanobacterial phylogenetic origin are displayed in blue, while those of chlamydial origin are shown in *yellow*. The host-derived NST ADP-glucose transporter (see text) is displayed in beige on the cyanobiont's inner membrane. Abbreviations are as follows: *AGPase* ADP-glucose pyrophosphorylase, *Glg* glycogenin, *SS-UDP* UDP-glucose specific GT5 glycogen(starch) synthase, *BE* branching enzyme, *BAM* β-amylase, *Pho*phosphorylase, *GlgX* GlgX-type of direct DBE (debranching enzyme), *iDBE* indirect debranching enzyme, *DPE2* amylomaltase or disproportionating enzyme (according to specificity), *Pho* glycogen or maltodextrin phosphorylase (according to specificity), *SS-ADP* SSIII-SSIV-type of ADP-glucose-specific soluble starch (glycogen) synthase, *GBSS* granule-bound starch synthase, *PWD-GWD* phosphoglucan or glucan water dikinases. Starch granules are displayed in white. Glycogen particles and oligosaccharide breakdown products are displayed in *black*. The chlamydial inclusion vesicle is displayed in *light brown* and the reticulate chlamydial bodies are shown in *gray* attached through the TTS (type three secretion system) to the inclusion vesicle membrane. Figure 4.5a displays the secretion of the chlamydial GlgX and GlgA (SS-ADP ancestor of SSIII-SSIV) effector proteins

the enzymes of glycogen catabolism to the glycogen pools (see Cenci et al. 2014 for a detailed discussion). We believe the evolving Archaeplastida capitalized on this to initiate a novel regulation of carbon breakdown that took into account the activity of the cyanobiont. In addition, this switch prevented direct access of intracellular Chlamydiales pathogens to the carbon pools, through their effector phosphorylases.

4.3.3 The Astonishing Features of the Ancient Glycogen Metabolism Network

As was pointed out in this chapter, storage polysaccharides and storage lipids define ideal buffers to circumvent all physiological problems, due to the disconnection of supply and demand for carbon, between the two unrelated partners of plastid endosymbiosis. The flux of carbon that became apparent in our first efforts of reconstruction of storage polysaccharide metabolism (Fig. 4.4a, b) convinced us that this pathway could have defined the biochemical flux through which photosynthetic carbon was exported at plastid endosymbiosis. This defines the symbiotic flux at the heart of symbiosis: the "raison d'être" of the cyanobiont-host partnership. It strokes us that the flux was optimal, at the very onset of the event. Indeed, according to Fig. 4.4a the substrate that flows out of the cyanobiont consists of ADP-Glc, which is the glycosyl-nucleotide sugar committed to storage in bacteria. Hence, the carbon that escapes the cyanobiont physically corresponds to that which was committed to storage anyhow. There is no penalty for the cyanobiont in the light, since this was committed to leave temporarily cyanobacterial metabolism. However, the loss of carbohydrate stores would have to be compensated for in darkness. There are several options here that include either the early targeting of the chlamydial ATP import protein to the cyanobiont inner membrane, or alternatively the cyanobiont could have maintained the synthesis of smaller amounts of glycogen by losing its former ability to synthesize large amounts of starch. This would have entailed the maintenance of a minimal amount of carbohydrate stores required in darkness and nevertheless the export of the large surplus of ADP-Glc. This would have further delayed the requirement for an immediate targeting of the chlamydial ATP import protein to the cyanobiont inner membrane, in the absence of the TOC machinery. In the cytosol of the host, ADP-Glc would not have interfered with any of the host's pathways since this metabolite is neither produced nor used by eukaryotes. The presence of the SSIII-SSIV glucan synthase in the host cytosol is inferred to have fed this carbon initially in the glycogen pool. The presence of this enzyme, in the cytosol, was required in the reconstruction of the ancient storage polysaccharide metabolism, to explain the presence of the SSIII-SSIV enzymes in Chloroplastida. It is this enzyme which, in the reconstruction, is responsible for establishing the biochemical link between the cyanobiont and its host. Five years after hypothesizing with such speculations the ancient presence of this activity, the corresponding enzyme sequence was found in the genome of the glaucophyte *Cyanophora paradoxa* genome sequence thereby considerably strengthening this hypothesis.

4.3.4 The Ancient Biochemical Link of Plastid Endoysmbiosis

One of the major findings of the *Cyanophora paradoxa* genome sequencing project consisted in the detection of a candidate UhpC-like transporter harboring clear-cut plastidial targeting sequences (Curtis et al. 2012). This correlated with the absence of candidate plastidial pPT transporters sequences, while several host NSTs of the eukaryotic endomembrane system were nevertheless detected. These results from bioinformatic studies were confirmed by the proteomic analysis carried on the glaucophyte plastids (the muroplast) by Facchinelli et al. (2013). In these studies the authors were surprised to find only 14 distinct candidate transporter proteins (including both thylakoid and envelope fractions), 3 of which were presumed to be of chlamydial descent (2 UhpC transporters and one ATP transporter) (Facchinelli et al. 2013). The figure of less than 14 candidate transporters, including the thylakoids, pales by comparison to the five- to tenfold higher figures reported for Rhodophyceae and Chloroplastida. Hence, it seems that the maintenance of the thick peptidoglycan layer prevented massive exploration of novel transport possibilities freezing the muroplast in an ancestral state of metabolite exchange. This at first glance agrees with the presence of the full glycolytic sequence observed in the muroplast stroma of *Cyanophora paradoxa* (Facchinelli et al. 2013). UhpC in proteobacteria has been proved to be a glucose-6-P sensor, regulating the import of hexose-P, but not a transporter (Schwoppe et al. 2003). *Chlamydia pneumoniae* UhpC was convincingly demonstrated to have lost the glucose-6-P sensor activity, upon gaining the transport properties (Schwoppe et al. 2002). This transport is in fact an exchange reaction between glucose-6-P and orthophosphate. This type of transport had been previously demonstrated to operate on purified *Cyanophora paradoxa* muroplasts (Schlichting and Bothe 1993). The UhpC gene was acquired by all three Archaeplastida clades, and displays a monophyletic origin in Archaeplastida. However, in both Rhodophyceae and Chloroplastida, the gene product is located elsewhere and not in plastids. In addition, only the *Chlorophyceae* and *Prasinophyceae* seem to have retained the gene which was apparently lost from all land plants and possibly also from the streptophyte green alga lineage.

Hence, the pressing question arises which came in first at the onset of endosymbiosis: UhpC or NST (nucleotide-sugar translocator) ? There is, as yet, no clear answer to this question, and both scenarios are equally viable. In a first scenario, NST successfully reached the inner membrane of the cyanobiont, at the onset of plastid endosymbiosis, as we initially proposed. This scenario is supported by the demonstrated ability of some members of this family of transporters to reach inner membranes of organelles in the absence of a protein-targeting machinery (Loddenkotter et al. 1993). In this scenario, the UhpC gene was transferred to the last common ancestor of Archaeplastida to fulfill a yet to be investigated function within the host but outside of the evolving plastid. The transporter was then recruited to the muroplast, from a duplicated copy that had gained a plastid-targeting sequence after divergence of the Glaucophyta from the other Archaeplastida lineages. Exchange of glucose-6-P for orthophosphate substituted for ADP-Glc export, thereby explaining

the loss of the ADP-Glc pyrophosphorylase, together with the NST gene, from the glaucophyte lineage. In a second scenario, UhpC came in first and the pPTs were acquired selectively by the Rhodophyceae and Chloroplastida lineages, from an ancient NST, after these lineages diverged from the Glaucophyta. Duplicated copies of the UhpC gene were then recruited by the common ancestor of the red and green algae to work elsewhere in the cell. If we assume that UhpC reached the inner membrane of the cyanobiont at the onset of the event, there would be no need for storage polysaccharide synthesis in the host cytosol from ADP-Glc. Yet, the existence of an SSIII-SSIV-like glucan synthase sequence in the *Cyanophora paradoxa* genome sequence clearly contradicts this conclusion. There is no need in eukaryotes for such an enzyme to sustain glycogen synthesis in the cytosol. Hence, even if the enzyme present in extant glaucophytes is likely to fulfill an important function in starch synthesis that is likely to be independent from the presence of ADP-Glc, this was clearly not the case for the ancestral gene product that was transferred from the Chlamydiales pathogens to the common ancestor of Archaeplastida. Another problem with a simplistic view of the UhpC first scenario comes from its phylogenetic origin. How could a hydrophobic transporter protein present on the membranes of Chlamydiales reach the inner membrane of the cyanobiont at the onset of plastid endosymbiosis? UhpC is not documented as an effector protein in Chlamydiales. It is very difficult to reconcile this observation with an early targeting of this transporter to the cyanobiont inner membrane. A way around this problem would be to imagine that the gene encoding this transporter had been transferred selectively to the common ancestor of Archaeplastida prior to plastid endosymbiosis to fulfill an unknown function in the heterotrophic phagotrophic ancestor of Archaeplastida. The UhpC transporter then reached the cyanobiont inner membrane, in the absence of the protein-targeting machinery. Under this hypothesis, the chlamydial phylogenetic origin of the protein is purely coincidental. This hypothesis has three major disadvantages. First it does not allow to understand why the glaucophytes should have lost the ancestral UhpC function of the heterotrophic eukaryotic ancestor while the latter should have been maintained in green and red algae. Second, it does not allow to understand what the benefits would have been for the transfer of the SSIII-SSIV glucan synthase gene for expression in the cytosol of the common Archaeplastida ancestor. Yet, this prokaryotic gene was indeed transferred and is evidenced today in the *Cyanophora* genome. Third, UhpC has not been documented to display an innate ability to reach the inner membranes of organelles. Hence, there is truly no reason to prefer this simplistic version over the comparable version of the NST first scenario. Nevertheless, we do not believe that the Chlamydiales origin of this transporter can be coincidental as this protein would have participated in the establishment of the metabolic link between the cyanobiont and its host. It is therefore very tempting to favor the UhpC first scenario, but in a fashion that explains the targeting of this protein at the onset of plastid endosymbiosis, and not in the simplistic version detailed above. In the meantime, it is presently very difficult to prefer any of the aforementioned scenarios and more information is needed. In particular, we need to know what function is fulfilled by UhpC transporters in red and green algae and what are the biochemical properties

of these transporters with respect to the targeting of organelle membranes in the absence of protein translocation machineries. We also need to further study the eukaryotic ancestors of the pPTs in more detail.

4.3.5 *Two Models for a Successful "Ménage à Trois"*

A growing number of phylogenetic studies demonstrate that Archaeplastida are characterized by the presence of 30–50 genes that display a common phylogenetic origin with intracellular bacterial pathogens of the order Chlamydiales (Moustafa et al. 2008). These obligatory intracellular pathogens replicate within inclusion vesicles, derived from endocytic membranes inside eukaryotic cells such as amoebas or animal cells (for review see Collingro et al. 2011). The presence of a continuous cell wall precludes infection, explaining why fungi plants and nonphagotrophic algae are immune to chlamydial infections. Nevertheless, if comparable studies are performed in the animal or fungal kingdoms, the signals decrease to 50 or 20 % respectively, of their intensity seen in Archaeplastida (Ball et al. 2013). This is especially relevant, because the fungi and animals have presently been more extensively studied and sampled than the Archaeplastida. So, how can we explain the presence of such a phylogenomic signal in clades that are immune to *Chlamydia* infections? Part of the answer came with the realization that over one-third of the LGTs, uniting Chlamydiales to Archaeplastida, were common to at least two of the three Archaeplastida lineages (Becker et al. 2008; Collingro et al. 2011; Huang and Gogarten 2007; Moustafa et al. 2008). Hence, a significant part of these LGTs can be dated back to a time where these lineages had not yet diverged. This time coincides with that of plastid endosymbiosis. Indeed, the common ancestor of Archaeplastida is not thought to have been surrounded by a continuous cell wall and is highly suspected to have been an active phagotroph, rendering it highly susceptible to chlamydial infections. While this explains the presence of chlamydial LGTs in the Archaeplastida genomes, it does not explain the selective enrichment of Archaeplastida, relatively to the animals, which unlike the latter were continuously exposed to chlamydial infections, ever since they diverged from their unicellular choanozoan-like ancestors. Several authors have explained this by hypothesizing a specific role of Chlamydiales intracellular pathogens in plastid endosymbiosis (Huang and Gogarten 2007; Becker et al. 2008; Moustafa et al. 2008; Collingro et al. 2011). In 2013, Ball et al. demonstrated that the SSIII-SSIV ancestral glucan synthase present in the cytosol of *Cyanophora paradoxa* and in the plastids of the green algae and land plants was an enzyme of chlamydial affiliation. In our reconstruction of the ancient storage polysaccharide metabolism network, we had proposed the existence of two key proteins that established, according to us, the efflux of photosynthetic carbon from the cyanobiont and its assimilation in the host glycogen pools. These are respectively the NST ADP-Glc transporter and the cytosolic ancestor of the ADP-Glc-specific SSIII-SSIV glucan synthase. To explain the presence of the SSIII-SSIV enzyme in the cytosol of the

common Archaeplastida ancestor, Ball et al. proposed that this gene in Chlamydiales encoded an effector protein that was secreted by these intracellular pathogens through the TTS (type three secretion system) in the cytosol of their host. By using a semi-in vitro TTS system in *Shigella*, Ball et al. demonstrated that genes from Parachlamydiaceae (the closest relatives to the archaeplastidal enzymes) encoding GlgC (ADP-Glc pyrophosphorylase), GlgA (ancestor of the archaeplastidal SSIII-SSIV glucan synthase), GlgX (debranching enzyme ancestor of the archaeplastidal isoamylase), and GlgP (glycogen phosphorylase) were indeed effector proteins secreted by the pathogens into their host's cytosol. From the Chlamydiales point of view, these effectors would induce cytosolic ADP-Glc synthesis and induction of glycogen synthesis at the beginning of the infection. Toward the end of the infection, the pathogen's unregulated glycogen phosphorylase would induce massive breakdown of the glycogen pool and its conversion into glucose-1-P and maltotetraose. Quite significantly, maltotetraose does not define a suitable substrate for cytosolic degradation by host enzymes. Hence, the pathogens induce formation of a metabolite, for which only they can find further use. The effector nature of the GlgA protein of *Chlamydia trachomatis* was further proven by others using an in vivo approach in a homologous system (Lu et al. 2013). If we assume that a heterotrophic phagotroph ancestor, infected by such pathogens, internalized a cyanobiont and managed to have an NST targeted to its inner membrane, then in effect, plastid endosymbiosis would have rescued the infected cell, by feeding unlimited amounts of ADP-Glc into the cytosolic glycogen pools, thanks to the presence of the pathogen's GlgA (SSIII-SSIV-like) glucan synthase. Without a single mutation in the pathogen's genome, this instantly changed the nature of the GlgA effector from potentially a virulence effector to a symbiogenic effector at the heart of one of the most important symbioses associations to have appeared on our planet. Ball et al. (2013) further outlined that plastid endosymbiosis was, in effect, a "ménage à trois," where three partners participated in a common carbon assimilation flux (Fig. 4.5a). This affords for a straightforward explanation of all observations, including the phylogenomic signal of Chlamydiales in Archaeplastida. Because, unlike the cyanobiont, the pathogen's compartment did not offer any other purpose than encoding useful genes, the chlamydial partner eventually disappeared, after it had exhausted all possibilities of evolving useful effectors to the tripartite symbiosis. Hence, the selection pressure that maintained the chlamydial compartment was the constant evolution of beneficial effectors to keep ahead of the LGT of its genes to the host nucleus. The ISA-GlgX effector, in this context, might define another example of beneficial cytosolic effector. Indeed, the switch to starch, in the common Archaeplastida ancestor, was most probably triggered by a gene duplication of GlgX. This, as we outlined above, happened fast after the onset of plastid endosymbiosis. We believe this could have happened before EGT of the ISA-GlgX effector genes to the nucleus of the host. Hence, maintenance of the pathogen would have been required for continued starch synthesis.

How does the "ménage à trois" stand in face of the UhpC first hypothesis? As we outlined above, we do not believe that the fact that UhpC defines a gene of Chlamydial phylogeny can be coincidental in the context of plastid endosymbiosis.

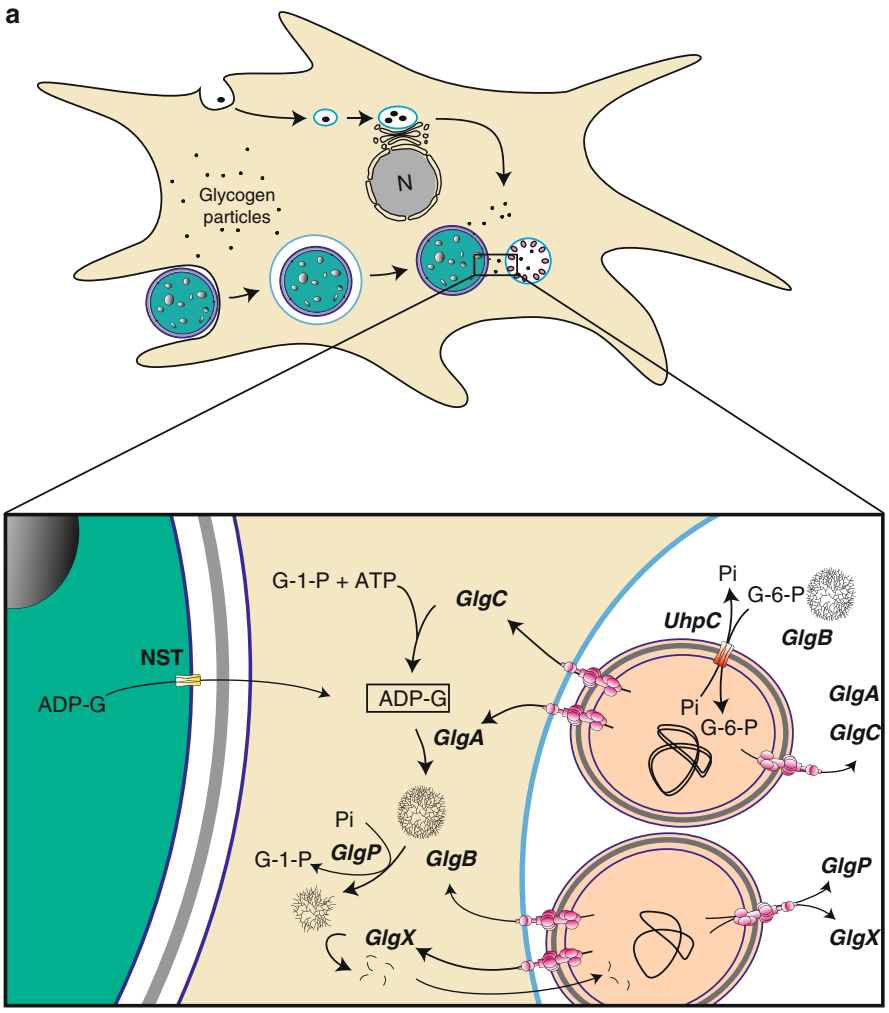

Fig. 4.5 Two alternative scenarios explaining the ménage à trois hypothesis. (**a**) The classical ménage à trois hypothesis (NST first scenario). Panel (**a**) recapitulates the scenario previously detailed in Fig. 4.4a. The large-size cyanobiont (in blue-green with starch granules displayed) is shown entering the host independently from the chlamydial pathogen (to scale) a mere *thick black dot*. A section displaying the tripartite interaction is *enlarged* and *boxed*. The chlamydial reticulate bodies are displayed attached through their syringe-like TTS (type three secretion system, displayed in *pink*) to the inclusion vesicle membrane (in *light blue*). Glycogen particles are depicted in *black* within the inclusion vesicle and in the host cytosol. Only the enzymes of chlamydial provenance are displayed and abbreviated by their gene symbols: *GlgA* glycogen synthase, *GlgB* branching enzyme, *GlgC* ADP-glucose pyrophosphorylase, *GlgP* glycogen (maltodextrin) phosphorylase, and *GlgX* GlgX-type of direct DBE. The NST (ADP-glucose transporter) transporter is highlighted on the cyanobiont membrane which was reached independently of an inexisting TIC-TOC protein import machinery. (**b**) The modified (chlamydioplast) ménage à trois (UhpC first scenario) hypothesis. The cyanobiont is shown entering together with a chlamydial elementary body.

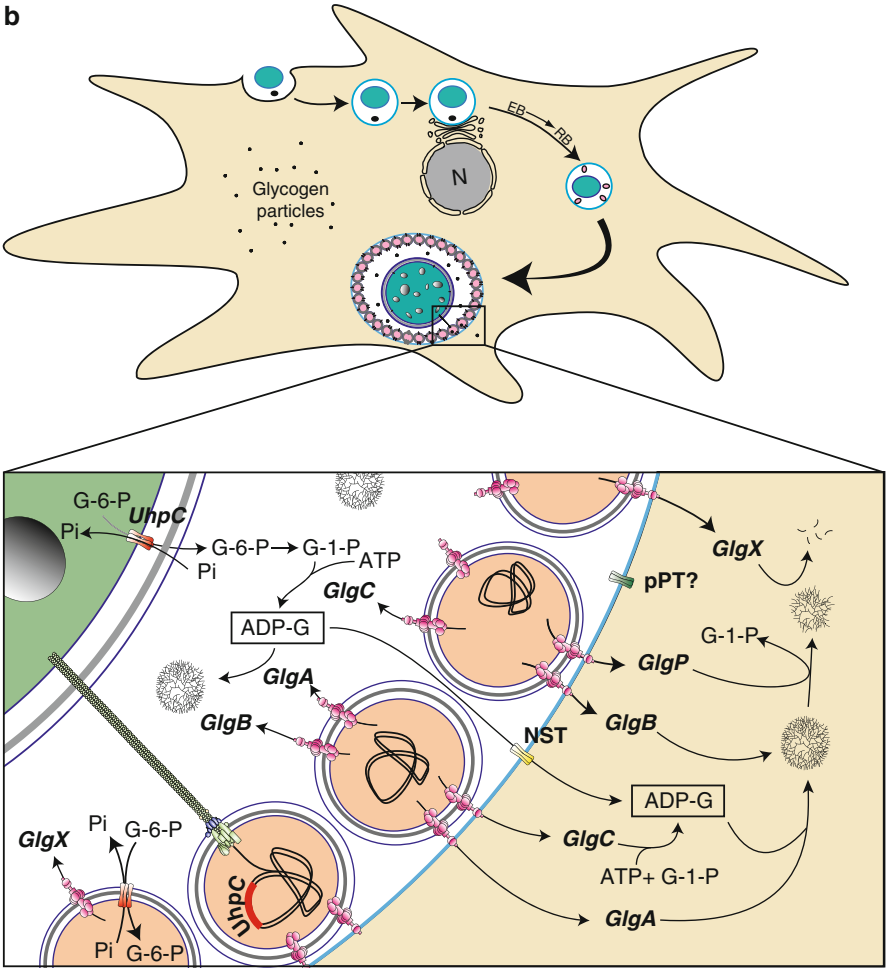

Fig. 4.5 (continued) The elementary body differentiates into reticulate bodies attached through their TTS to the modified phagocytic vacuole thereby preventing phagocytosis of the cyanobiont. The T4SS (type four secretion system displayed in *green*, *blue,* and *black*) responsible for conjugative DNA transfer transfers ad minima the UhpC and NTT (ATP import protein) chlamydial genes for expression within the cyanobiont. UhpC is displayed in red on the cyanobiont inner membrane, while NST is displayed in yellow on the inclusion vesicle. The low affinity ADP-glucose transporter (NST) will transport the overflow of carbon assimilation to the cytosol where it will be metabolized as depicted in Figs. 4.4b and 4.5a

Hence, we do not believe that UhpC had already been transferred to the host ancestor genome prior to plastid endosymbiosis. Rather, we believe that UhpC was transferred, post endosymbiosis, to the nucleus of the host. If UhpC defines the original carbon translocator, we must assume a very early targeting of this protein to the inner envelope of the cyanobiont. If we assume that UhpC was encoded by a Chlamydiales

intracellular pathogen that infected the host of plastid endosymbiosis, it is very hard to imagine that such a hydrophobic transporter could have been secreted through the TTS and targeted to the inner membrane of the cyanobiont. In order to solve this problem, Facchinelli et al. (2013) have proposed that the cyanobiont entered by phagocytosis together with the Chlamydiales pathogen (Fig. 4.5b). We know that Parachlamydiaceae which are the pathogens that are closest by sequence to the genes transferred by EGT to Archaeplastida possess a fully functional type IV secretion system (Collingro et al. 2011) that would enabled the conjugative transfer of chlamydial DNA to the cyanobiont. This would have facilitated the very early transfer of the UhpC and the NTT (ATP import protein) genes to the cyanobiont genome for expression of these hydrophobic transporter proteins on the cyanobiont inner membrane. Possibly, other useful chlamydial genes could have been transmitted in this fashion to stabilize the cyanobiont in the inclusion vesicle environment. To be able to benefit from the cyanobiont's export of glucose-6-P, there was a need for the pathogen to have a biochemical buffer between the unsynchronized supply and demand of carbon within the inclusion vesicle for similar reasons as those that we have outlined concerning cytosolic storage polysaccharide synthesis. In *Chlamydia trachomatis*, massive intravesicular glycogen accumulation has been documented to the point where iodine staining of infected tissues constitutes a diagnostic feature of infection of human tissues by *C. trachomatis*. In addition to this, we know that at least in *C. trachomatis*, the GlgA effector enzyme is secreted in the host cytosol and that glycogen synthesis is also induced in the host cytosol upon infection with Chlamydiae (Ojcius et al. 1998). Hence, it very likely that Chlamydiales can trigger glycogen synthesis and degradation in all three compartments: in the host cytosol, in the inclusion vesicle, and within the bacterial pathogens. In this context, if divisions of the pathogens within the inclusion vesicle slowed down and if the cyanobiont supplied an excess of carbon, the latter would have been redirected to the intravesicular glycogen pool. It is plain to see that the cyanobiont is enslaved by the pathogen and that the latter gets most of the photosynthate (Fig. 4.5b). Hence, the Chlamydiales are on top of the "ménage à trois." But what's in there for the eukaryotic host? Facchinelli et al. (2013) proposed that the Golgi NST was targeted as a low affinity ADP-Glc translocator to the inclusion vesicle membrane (Fig. 4.5b). This would have been very simple, as this inclusion vesicle is in continuity and derived from the host endomembrane system. Colleoni et al. have previously documented that host GDP-mannose translocators, phylogenetically related to the pPT translocators, transport ADP-Glc as efficiently as the natural GDP-mannose substrate but with lower affinity. Hence, whenever the sink of carbon toward the Chlamydiales decreased, the concentration of ADP-Glc within the inclusion vesicle would rise leading to an efflux of the nucleotide sugar toward the host cytosol. In the cytosol, the situation would be exactly the same as that detailed in the NST first scenario and the ménage à trois would have been operating in a similar fashion. The UhpC scenario detailed above (Fig. 4.5b) displays a number of very attractive features. It explains how phagocytosis aborted and how the cyanobiont benefitted from the remodeling of the phagocytosis vesicle into a stable inclusion vesicle by the Chlamydiales pathogens. The environment

offered by the inclusion vesicle is less dense than the cytosol and more similar to the external medium to which free-living cyanobacteria are better adapted. The chlamydial genes, required for carbon efflux and ATP import, would have been readily transmitted to the cyanobiont at the onset of plastid endosymbiosis. Other chlamydial genes stabilizing the cyanobiont in an intracellular environment could also have been directly transferred by conjugative transfer. We would like to refer to this stage as the "chlamydioplast" stage of plastid endosymbiosis. At this stage, the presence of the outer LPS-like polysaccharide shell, surrounding the cyanobiont, would not have prevented the transfer and expression of UhpC. The efflux of Glucose-6-P, on the other hand, would have minimized the synthesis of all polysaccharides within the cyanobiont. Because of the presence of the inclusion membrane, it is unlikely that the events that led to the appearance of the TIC-TOC protein import machinery were initiated at the chlamydioplast stage. However, this stage remains compatible with massive gene losses and conjugative transfers of chlamydial genes, in effect priming the cyanobiont for its future escape to the cytosol. It also was compatible with EGT to the nuclear genome of cyanobacterial genes for expression of the products in the host cytosol but not for their targeting back to the cyanobiont. Hence, the GBSS gene could have been transferred to the host nucleus at this stage and in the host cytosol starch would already have replaced glycogen thanks to the chlamydial GlgX effectors. We believe that the outer polysaccharide layers disappeared at this stage, thereby exposing the outer cyanobacterial membrane. Because of the efflux of glucose-6-P, we believe that the massive synthesis of cyanobacterial starch was replaced by that of minimal amount of glycogen through the loss of the genes required for starch crystallization. This would have increased the resistance to oxidative stress that would otherwise have been generated by the complete loss of cyanobacterial starch. The presence of a significant glycogen pool in the cyanobiont together with the chlamydial ATP import protein would have also allowed the loss of the cyanobacterial respiratory complexes. The presence in the TIC-TOC apparatus of a common core set of genes in the three Archaeplastida lineages suggests that the cyanobiont escaped from the chlamydioplast, very early on, before divergence of the 3 Archaeplastida lineages. This allowed the common evolution of the TIC-TOC machinery. Exposure of the cyanobiont membranes to the cytosol could have enabled the direct targeting of NST to the cyanobiont inner membrane. The fast evolution of TOC would have stabilized this targeting of the NST protein and would have facilitated EGT to the host nucleus of both chlamydial transporter genes present in the cyanobiont genome. Before divergence of the glaucophytes, the last common ancestor achieved complete loss of all glycogen metabolism genes with the exception of ADP-Glc pyrophosphorylase. Upon divergence from the Rhodophyceae and Chloroplastida, the Glaucophyta lost both the NST transporter and ADP-Glc pyrophosphorylase. On the other hand, the common ancestor of both Rhodophyceae and Chloroplastida kept the UhpC nuclear gene but redirected the gene product to other compartments. Hence, UhpC disappeared from the inner cyanobiont membrane in the common ancestor of the evolving red and green algae. The peptidoglycan layer also disappeared in this ancestor. This disappearance enabled the pyrenoid to physically interact

with cytosolic starch across the two plastidial membranes, an interaction which was previously prevented by the peptidoglycan layer. This interaction allowed the exploration of new transporter properties facilitated by the duplication of the NST gene which led to the pPT transporters. The green algae further evolved to redirect cytosolic starch metabolism to the evolving chloroplasts, while the red algae kept the pPTs but lost both the original NSTs, ADP-Glc pyrophosphorylase and the cytosolic ADP-Glc-specific SSIII-SSIV chlamydial glucan synthase.

Both "ménage à trois scenarios" detailed above are presently compatible with all recorded observations. Future studies on the NST ancestors or on the UhpC transporters are required to help us distinguish between these alternatives.

4.3.6 Polyphyletic Transitions of Glycogen to Starch Metabolism in the Eukaryote and Bacterial Domains

When we initially tackled the study of the evolution of storage polysaccharide metabolism, we had misinterpreted the phylogeny of the plant isoamylases and initially believed these enzymes to be of cyanobacterial origin (Deschamps et al. 2008a). The finding of starch-like structures in cyanobacteria initially suggested to us that there may have been one single acquisition of starch from glycogen metabolism that happened over two billion years ago in single-cell diazotrophic cyanobacteria, possibly at a time when nitrogenase became exposed to critical levels of oxygen. This would have been followed by transmission of this ability at primary plastid endosymbiosis to Archaeplastida through EGT of the required cyanobacterial isoamylase. The Archaeplastida would have in turn transmitted this ability through secondary endosymbiosis to the cryptophytes and the alveolates. This simple and attractive scenario turned out to be completely wrong! First, acquisition of starch (or semi-amylopectin) in cyanobacteria turns out to be distinctively polyphyletic. Indeed, while genetic approaches carried out in *Cyanobacterium* CLg1 demonstrated that an isoamylase-like enzyme was responsible for generating semi-amylopectin (Cenci et al. 2013), genome sequences of other starch or semi-amylopectin-accumulating clades suggest that at in some clades, GH13 direct DBEs are apparently absent. If the corresponding genomes are complete, this observation implies that, in these clades, the enzymes responsible for turning glycogen metabolism into starch are evidently different and that starch acquisition has therefore occurred several times in cyanobacteria. In addition to this, we now know that the ancestor of the plant isa1 gene is of chlamydial origin. So, not only did starch appear, possibly many times, from glycogen metabolism in cyanobacteria, but the appearance of starch metabolism in Archaeplastida was also independent of these events. From our present knowledge of starch metabolism evolution, it remains plausible that transition to starch occurred only once in the cytosol of the last common Archaeplastida ancestors. Yet, because the starch metabolism pathway was probably redirected stepwise to the chloroplast, the network went through

another sequence of glycogen to starch transition in the green algae and land plants evolving chloroplasts (Deschamps et al. 2008c). Hence, Archaeplastida experienced at minima two such transitions. In starch-storing secondary endosymbiosis lineages (cryptophytes and alveolates), the secondary eukaryotic host of endosymbiosis is generally thought to be distinct. This correlates with a distinct localization of starch and a different nature of the enzyme network. To be more specific, bona fide debranching enzyme candidate sequences are absent altogether from cryptophytes, while direct DBEs were apparently replaced by indirect DBE in alveolates (not taking into account the tertiary endosymbiosis dinoflagellate lineages). This again suggests a distinct transition in the alveolate cytosol. The cryptophytes possibly never experienced such a transition, since starch was present in the periplastidial space at the very beginning of the event. Evidently in this case the rhodophycean isoamylase-like enzyme that we highly suspect to be responsible for amylopectin crystallization was replaced by an enzyme from another CAZy subfamily. The only activity which in eukaryotes was never replaced by distinct families of enzyme sequences is defined by the glucan and phosphoglucan water dikinases. These enzyme sequences correlate 100 % with the presence of starch-like structures in eukaryotes and can be taken as a diagnostic of the presence of semicrystalline α-glucan polysaccharides. We believe that in cyanobacteria, the absence of such dikinases stems from the absence of a preexisting pathway of glycogen phosphorylation. This absence must have required coevolution and adaptation of the glycogen catabolism machinery to the presence of more hydrophobic polysaccharides. Quite significantly, in eukaryotes, the glucan dikinase genes disappear whenever a clade reverts from starch to glycogen metabolism. Going back from starch to glycogen is documented in at least two cases in Cyanidiales (*Galdieria sp.* and *Cyanidium sp*) and possibly also in ciliates, in at least one case, in alveolates (the ciliates). We believe that in cyanobacteria this would also have been frequent. Hence, we can conclude that transition of glycogen to starch metabolism and reversion to glycogen seem to have occurred frequently both in cyanobacteria and eukaryotes from distinct enzyme networks. Such transitions would have been dictated by polysaccharide structure optimization with respect to distinct types of cell physiologies.

4.3.7 Why Amylose?

Amylose (defined as chains exceeding DP100 with less than 1 % branches) does not define a universal feature of starches accumulating in both the bacterial and eukaryotic domains. However in both domains, amylose synthesis always requires the presence of GBSS (granule-bound starch synthase) an enzyme belonging to the GT5 bacterial CAZy family of glycogen(starch)synthases. We believe that the cyanobacterial origin of this enzyme is presently beyond doubt (Ball et al. 2013). The reasoning is as follows. First the GBSS-SSI-SSII supergroup of starch synthases in all organisms is monophyletic and displays only bacteria and archaea at the root of this superclade. Second the phylogeny of the GBSS subgroup places

the cyanobacterial enzymes as sisters to the Chloroplastida, glaucophytes, and Rhodophyceae and not within either the green or red algae. Had the cyanobacterial enzyme been acquired by LGT from green or red algae, the *Cyanobacterium* CLg1 and *Crocosphaera watsonii* GBSS sequences would have been expected to lie within the red or green algae possibly in a polyphyletic fashion. Hence, the GBSS phylogeny is consistent with the enzyme having evolved in cyanobacteria and been transmitted by EGT to Archaeplastida. Robust grouping of SSI-SSII enzymes of green algae with the GBSS sequences is interpreted as originating from gene duplications that happened during the process of the complex rewiring of the storage polysaccharide network to the evolving chloroplast. Outgroup rooting of the GBSS-SSI-SSII superclade displays an origin at midpoint between the GBSS and SSI-SSII subgroup. The very high bootstrap support evidenced for this node reflects more the fact that this supergroup is unquestionably monophyletic, rather than correctly positioned with respect to the general GT5 phylogenetic tree. This can be attributed to the fact that the GBSS-SSI-SSII superclade is distantly related to other GT5 starch(glycogen) synthases and that the SSI-SSII subgroup has itself substantially diverged from the GBSS subgroup. Hence, as outlined in Ball et al. (2013), we believe that the true root position of the GBSS-SSI-SSII superclade lies at the base of the cyanobacterial GBSS sequences.

Very few cyanobacteria and Rhodophyceae display the presence of amylose, while the latter is universally present in Chloroplastida and possibly also in Glaucophyta. One can speculate relatively to the benefits of amylose synthesis. A universal property of GBSS is defined by its apparent low affinity for its glycosyl-nucleotide substrates and its apparent low activity in the unbound form. The processivity of the Chloroplastida enzyme seems confirmed by those evidenced in semi-in vitro systems analyzed with *Cyanobacteria*, Glaucophyta, cryptophytes, and dinoflagellates (Deschamps et al. 2006, 2008a, b; Plancke et al. 2008). In all these systems, amylose synthesis is triggered by very high nucleotide-sugar concentrations by processive elongation. A substantial part of this synthesis escapes branching, quite simply because the polysaccharide product is sheltered within the granule and therefore escapes the action of the hydrosoluble branching enzymes. In wild-type cells, amylose accounts to up to 30 % maximum of the weight of the granule and many times less. Amylose synthesis typically acts as an overflow mechanism, when the amylopectin synthesis machinery is insufficient to deal with the rising nucleotide-sugar concentrations. This will enable amylose containing starch to incorporate an additional 10–30 % glucose within the granules. However, overflow is not the only possible function of amylose synthesis. Wattebled et al. showed that intensive de novo synthesis of amylose, in semi-in vitro systems, results in a very strong alteration of starch granule morphogenesis that is accompanied by a switch from a classical lenticular shape to a tubular fused network, probably resulting from fusion of distinct altered starch granules (Wattebled et al. 2002). Hence, GBSS affords for localized polysaccharide synthesis that triggers starch granule fusion, thereby acting as a molecular glue. If we couple this property, to the very low apparent affinity of this enzyme toward its glycosyl-nucleotide substrate, we can predict that GBSS will respond dramatically to localized substrate gradients. In

effect, localized nucleotide-sugar gradients will induce a morphogenesis response, whereupon the granules will surround the glycosyl-nucleotide source. This very conveniently explains the surrounding of the pyrenoids, which will contain most of the active RuBisCO in normal CO_2-limited conditions. It is striking to note that GBSS and amylose are present in all cases where a localized pyrenoidal starch sheet is evidenced (all green algae, cryptophytes, Porphyridiales). In other red algae, where starch is uncoupled from the pyrenoid, floridean starch appears devoid of amylose. This also appears to be the case in heterotrophic alveolates. Pyrenoids are certainly involved in the CO_2 concentration mechanism, operating in aqueous environments, where they are considered to define carboxysome-like structures devoid of carbonic anhydrase. Quite strikingly, in Glaucophyta and green algae GBSS transcription is massively induced in low CO_2 conditions. Pyrenoidal starch is not required for the CCM to operate normally but results from the generation through the pyrenoids of localized metabolite concentration gradients (Izumo et al. 2011; Villarejo et al. 1996). The correlation between pyrenoidal starch and the presence of amylose is nevertheless insufficient to account for the presence of GBSS in cyanobacteria glaucophytes and land plants which are all devoid of pyrenoids (except for the hornworts). Nevertheless, we would like to suggest that in all these cases selection for amylose-containing starches can also be explained by a combination of morphogenetic and metabolic reasons.

4.4 Conclusions

Investigating the evolution of starch metabolism has yielded a number of insights that have impacted both our understanding of the biochemistry of starch metabolism and our understanding of the metabolic integration of novel organelles.

The major conclusions that we can deduce from these approaches on our understanding of the biochemistry of starch metabolism are as follows:

1. Comparative biochemistry of starch metabolism proves that starch synthesis and degradation is not inherently more complex than those of glycogen. The presence of 30–40 genes in the green lineage that display specific and only partly redundant functions for the synthesis or degradation of starch has mislead researchers to believe that the apparent complexity of the polysaccharide structure had to be matched by an equally complex metabolic pathway. This is clearly not the case, and this complexity results more from the history of a pathway, which was redirected to plastids, at a time when the genes of storage polysaccharide metabolism had already been lost from the ancestral plastid genome. This history testifies to the apparent low costs that entail gene duplications and subfunctionalizations. The simplest networks that are sufficient to yield starch are those displayed by the red algae and the apicomplexa parasites. Presently, researchers are trying to achieve starch granule synthesis in vitro. We therefore

strongly believe that the enzymes that should be used for these approaches are those that are evidenced in the red alga or apicomplexa networks.

2. The generation of cyanobacterial mutants that overaccumulate glycogen and are defective for starch synthesis mirrors analogous results obtained in green algae and plants. The finding of a defect for an isoamylase type of direct DBE emphasizes a remarkable case of convergent evolution, where at two independent times nature has recruited a similar enzyme, from its molecular toolbox, to achieve the crystallization of amylopectin and the insolubilization of starch granules. This result strongly suggests that polysaccharide debranching may define a universal feature required to generate amylopectin from a preexisting glycogen synthesis machinery. Nevertheless, bioinformatic analysis clearly demonstrates the absence of isoamylase candidate sequences, in some starch-accumulating cryptophytes and alveolates. We believe that other CAZymes have been recruited to replace the classical DBE isoamylase function in these clades, but this, however, remains to be demonstrated. If this is the case, then the comparison of the 3D structures and organization of all these very diverse enzymes will facilitate the finding of those common features, which are responsible for amylopectin crystallization.
3. It is generally believed that GT3 glycogen synthases require glycogenin for glycogen synthesis priming, and GT5 glycogen and starch synthases do not. The study of the distribution of glycogenin among eukaryotes and its correlation to that of the GT5 and GT3 enzymes seems at first glance to invalidate this idea. Polysaccharide synthesis priming could thus define a built-in general property of glycogen synthases that could be occasionally lost in a quite unpredictable clade-specific fashion.

Our studies on the evolution of the starch metabolism pathway have also impacted our current understanding of the establishment of plastid endosymbiosis as follows:

1. The reconstruction of a hypothetical ancient storage polysaccharide network in the common ancestor of Archaeplastida suggests that cytosolic storage polysaccharides may have acted as buffers between the unsynchronized supply and demand of carbon at the onset of plastid endosymbiosis.
2. The finding that both the SSIII-SSIV ancestral enzyme and isoamylase are phylogenetically derived from chlamydial effector proteins secreted by these pathogens in the cytosol of eukaryotes provides a detailed scenario, explaining the early phases of plastid endosymbiosis.

References

Albrecht T, Haebel S, Koch A et al (2004) Yeast glycogenin (Glg2p) produced in Escherichia coli is simultaneously glucosylated at two vicinal tyrosine residues but results in a reduced bacterial glycogen accumulation. Eur J Biochem 271:3978–3989

Alonso-Casajus N, Dauvillee D, Viale AM et al (2006) Glycogen phosphorylase, the product of the glgP Gene, catalyzes glycogen breakdown by removing glucose units from the nonreducing ends in Escherichia coli. J Bacteriol 188:5266–5272

Archibald JM, Lane CE (2009) Going, going, not quite gone: nucleomorphs as a case study in nuclear genome reduction. J Hered 100:582–590

Arias MC, Danchin EG, Coutinho P et al (2012) Eukaryote to gut bacteria transfer of a glycoside hydrolase gene essential for starch breakdown in plants. Mob Genet Elem 2:81–87

Badger MR, Price GD (2003) CO2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot 54:609–622

Ball S (1998) Regulation of starch biosynthesis. Kluwer Academic Publishers, Dordrecht

Ball S (2002) The intricate pathway of starch biosynthesis and degradation in the monocellular alga Chlamydomonas reinhardtii. Aust J Chem 55:1–11

Ball S, Deschamps P (2008) Starch metabolism. In: Stem DB (ed) The chlamydomonas sourcebook: 2. Organellar and metabolic processes. Academic, London, pp 1–40

Ball S, Guan HP, James M et al (1996) From glycogen to amylopectin: a model for the biogenesis of the plant starch granule. Cell 86:349–352

Ball S, Colleoni C, Cenci U et al (2011) The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. J Exp Bot 62:1775–1801

Ball SG, Subtil A, Bhattacharya D et al (2013) Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis? Plant Cell 25:7–21

Bandyopadhyay A, Elvitigala T, Liberton M et al (2013) Variations in the rhythms of respiration and nitrogen fixation in members of the unicellular diazotrophic cyanobacterial genus Cyanothece. Plant Physiol 161:1334–1346

Bäumer D, Preisfeld A, Ruppel HG (2001) Isolation and characterization of paramylon synthase from *Euglena gracilis* (Euglenophyceae). J Phycol 37:38–46

Baurain D, Brinkmann H, Petersen J et al (2010) Phylogenomic evidence for separate acquisition of plastids in cryptophytes, haptophytes, and stramenopiles. Mol Biol Evol 27:1698–1709

Becker B, Hoef-Emden K, Melkonian M (2008) Chlamydial genes shed light on the evolution of photoautotrophic eukaryotes. BMC Evol Biol 8:203

Beller M, Thiel K, Thul P et al (2010) Lipid droplets: a dynamic organelle moves into focus. FEBS Lett 584:2176–2182

Bergman B, Gallon JR, Rai AN et al (1997) N2 fixation by non-heterocystous cyanobacteria. FEMS Microbiol Rev 19:139–185

Bhattacharya D, Archibald JM, Weber AP et al (2007) How do endosymbionts become organelles? Understanding early events in plastid evolution. Bioessays 29:1239–1246

Bhattacharya D, Price DC, Chan CX et al (2013) Genome of the red alga Porphyridium purpureum. Nat Commun 4:1941

Blank CE, Sanchez-Baracaldo P (2010) Timing of morphological and ecological innovations in the cyanobacteria – a key to understanding the rise in atmospheric oxygen. Geobiology 8:1–23

Borowitza MA (1978) Plastid development and floridean starch grain formation during carposporogenesis in the coralline red alga Lithothrix aspergillum gray. Protoplasma 95:217–228

Brautigan DL (2013) Protein Ser/Thr phosphatases – the ugly ducklings of cell signalling. FEBS J 280:324–345

Burris RH (1991) Nitrogenase. J Biol Chem 266:9339–9342

Busi MV, Barchiesi J, Martin M et al (2013) Starch metabolism in green algae. Starch/Stärke 66:28–40

Cavalier-Smith T (1999) Principles of protein and lipid targeting in secondary symbiogenesis: euglenoid, dinoflagellate, and sporozoan plastid origins and the eukaryote family tree. J Eukaryot Microbiol 46:347–366

Cavalier-Smith T, Chao EE, Snell EA et al (2014) Multigene eukaryote phylogeny reveals the likely protozoan ancestors of opisthokonts (animals, fungi, choanozoans) and Amoebozoa. Mol Phylogenet Evol 81C:71–85

Cenci U, Chabi M, Ducatez M et al (2013) Convergent evolution of polysaccharide debranching defines a common mechanism for starch accumulation in cyanobacteria and plants. Plant Cell 25:3961–3975

Cenci U, Nitschke F, Steup M et al (2014) Transition from glycogen to starch metabolism in Archaeplastida. Trends Plant Sci 19:18–28

Chandra G, Chater KF, Bornemann S (2011) Unexpected and widespread connections between bacterial glycogen and trehalose metabolism. Microbiology 157:1565–1572

Charng YY, Kakefuda G, Iglesias AA et al (1992) Molecular cloning and expression of the gene encoding ADP-glucose pyrophosphorylase from the cyanobacterium Anabaena sp. strain PCC 7120. Plant Mol Biol 20:37–47

Choi JH, Lee H, Kim YW et al (2009) Characterization of a novel debranching enzyme from Nostoc punctiforme possessing a high specificity for long branched chains. Biochem Biophys Res Commun 378:224–229

Collen J, Porcel B, Carre W et al (2013) Genome structure and metabolic features in the red seaweed Chondrus crispus shed light on evolution of the Archaeplastida. Proc Natl Acad Sci U S A 110:5247–5252

Colleoni C, Suzuki E (2012) Chapter 5: Storage polysaccharide metabolism in Cyanobacteria. In: Tetlow IJ (ed) Starch: origins, structure and metabolism, vol 5, Essential reviews in experimental biology. Society for Experimental Biology, London

Colleoni C, Linka M, Deschamps P et al (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

Collingro A, Tischler P, Weinmaier T et al (2011) Unity in variety – the pan-genome of the Chlamydiae. Mol Biol Evol 28:3253–3270

Compaore J, Stal LJ (2010) Oxygen and the light–dark cycle of nitrogenase activity in two unicellular cyanobacteria. Environ Microbiol 12:54–62

Coppin A, Varre JS, Lienard L et al (2004) 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

Coutinho PM, Henrissat B (1999) Carbohydrate-active enzymes: an integrated database approach. Royal Society of Chemistry, Cambridge

Crowe SA, Dossing LN, Beukes NJ et al (2013) Atmospheric oxygenation three billion years ago. Nature 501:535–538

Curatti L, Giarrocco LE, Cumino AC et al (2008) Sucrose synthase is involved in the conversion of sucrose to polysaccharides in filamentous nitrogen-fixing cyanobacteria. Planta 228:617–625

Curtis BA, Tanifuji G, Burki F et al (2012) Cryptophyte and chlorarachniophyte nuclear genomes reveal evolutionary mosaicism and fate of nucleomorphs. Nature 492:59–65

Dauvillée D, Kinderf IS, Li Z et al (2005) Role of the Escherichia coli glgX gene in glycogen metabolism. J Bacteriol 187:1465–1473

Dauvillee D, Deschamps P, Ral JP et al (2009) Genetic dissection of floridean starch synthesis in the cytosol of the model dinoflagellate Crypthecodinium cohnii. Proc Natl Acad Sci U S A 106:21126–21130

Deane JA, Strachan IM, Saunders GW et al (2002) Cryptomonad evolution: nuclear 18S rDNA phylogeny versus cell morphology and pigmentation. J Phycol 38:1236–1244

Delrue B, Fontaine T, Routier F et al (1992) Waxy Chlamydomonas reinhardtii: monocellular algal mutants defective in amylose biosynthesis and granule-bound starch synthase activity accumulate a structurally modified amylopectin. J Bacteriol 174:3612–3620

Delwiche CF (1999) Tracing the thread of plastid diversity through the tapestry of life. Am Nat 154:S164–S177

Deschamps P, Haferkamp I, Dauvillée D et al (2006) Nature of the periplastidial pathway of starch synthesis in the cryptophyte Guillardia theta. Eukaryot Cell 5:954–963

Deschamps P, Colleoni C, Nakamura Y et al (2008a) Metabolic symbiosis and the birth of the plant kingdom. Mol Biol Evol 25:536–548

Deschamps P, Guillebeault D, Devassine J et al (2008b) The heterotrophic dinoflagellate Crypthecodinium cohnii defines a model genetic system to investigate cytoplasmic starch synthesis. Eukaryot Cell 7:872–880

Deschamps P, Haferkamp I, d'Hulst C et al (2008c) The relocation of starch metabolism to chloroplasts: when, why and how. Trends Plant Sci 13:574–582

Deschamps P, Moreau H, Worden AZ et al (2008d) Early gene duplication within chloroplastida and its correspondence with relocation of starch metabolism to chloroplasts. Genetics 178:2373–2387

Devillers CH, Piper ME, Ballicora MA et al (2003) Characterization of the branching patterns of glycogen branching enzyme truncated on the N-terminus. Arch Biochem Biophys 418:34–38

Douglas SE, Penny SL (1999) The plastid genome of the cryptophyte alga, Guillardia theta: complete sequence and conserved synteny groups confirm its common ancestry with red algae. J Mol Evol 48:236–244

Douglas S, Zauner S, Fraunholz M et al (2001) The highly reduced genome of an enslaved algal nucleus. Nature 410:1091–1096

Ekman P, Yu S, Pedersén M (1991) Effects of altered salinity, darkness and algal nutrient status on floridoside and starch content, alpha-galactosidase activity and agar yield of cultivated *Gracilaria sordida*. Br Phycol J 26:123–131

Facchinelli F, Pribil M, Oster U et al (2013) Proteomic analysis of the Cyanophora paradoxa muroplast provides clues on early events in plastid endosymbiosis. Planta 237:637–651

Falcon LI, Cipriano F, Chistoserdov AY et al (2002) Diversity of diazotrophic unicellular cyanobacteria in the tropical North Atlantic Ocean. Appl Environ Microbiol 68:5760–5764

Falkowski P, Raven JA (2007) Aquatic photosynthesis. Princeton University Press, Princeton

Fathinejad S, Steiner JM, Reipert S et al (2008) A carboxysomal carbon-concentrating mechanism in the cyanelles of the 'coelacanth' of the algal world, Cyanophora paradoxa? Physiol Plant 133:27–32

Fettke J, Hejazi M, Smirnova J et al (2009) Eukaryotic starch degradation: integration of plastidial and cytosolic pathways. J Exp Bot 60:2907–2922

Forsyth G, Hirst EL, Oxford AE (1953) Protozoal polysaccharides. Structure of a polysaccharide produced by Cycloposthium. J Chem Soc 2030–2033

Fredrick JF (1968) Biochemical evolution of glucosyl transferase isozymes in algae. Ann N Y Acad Sci 151:413–423

Frueauf JB, Ballicora MA, Preiss J (2002) Alteration of inhibitor selectivity by site-directed mutagenesis of Arg(294) in the ADP-glucose pyrophosphorylase from Anabaena PCC 7120. Arch Biochem Biophys 400:208–214

Fu J, Xu X (2006) The functional divergence of two glgP homologues in Synechocystis sp. PCC 6803. FEMS Microbiol Lett 260:201–209

Glöckner G, Noegel A (2013) Comparative genomics in the Amoebozoa clade. Biol Rev 88:215–225

Goldemberg SH, Marechal LR (1963) Biosynthesis of paramylon in Euglena gracilis. Biochim Biophys Acta 71:743–744

Grosche C, Hempel F, Bolte K et al (2014) The periplastidal compartment: a naturally minimized eukaryotic cytoplasm. Curr Opin Microbiol 22C:88–93

Grundel M, Scheunemann R, Lockau W et al (2012) Impaired glycogen synthesis causes metabolic overflow reactions and affects stress responses in the Cyanobacterium Synechocystis sp. PCC 6803. Microbiology 158:3032–3043

Guerra LT, Xu Y, Bennette N et al (2013) Natural osmolytes are much less effective substrates than glycogen for catabolic energy production in the marine cyanobacterium Synechococcus sp. strain PCC 7002. J Biotechnol 166:65–75

Haferkamp I, Deschamps P, Ast M et al (2006) Molecular and biochemical analysis of periplastidial starch metabolism in the cryptophyte Guillardia theta. Eukaryot Cell 5:964–971

Henrissat B, Deleury E, Coutinho PM (2002) Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? Trends Genet 18:437–440

Higo A, Katoh H, Ohmori K et al (2006) The role of a gene cluster for trehalose metabolism in dehydration tolerance of the filamentous cyanobacterium Anabaena sp. PCC 7120. Microbiology 152:979–987

Hirabaru C, Izumo A, Fujiwara S et al (2010) The primitive rhodophyte Cyanidioschyzon merolae contains a semiamylopectin-type, but not an amylose-type, alpha-glucan. Plant Cell Physiol 51:682–693

Huang J, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biol 8:R99

Husnik F, Nikoh N, Koga R et al (2013) Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell 153:1567–1578

Izumo A, Fujiwara S, Sakurai T et al (2011) Effects of granule-bound starch synthase I-defective mutation on the morphology and structure of pyrenoidal starch in Chlamydomonas. Plant Sci 180:238–245

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

Kiel JA, Boels JM, Beldman G et al (1990) Nucleotide sequence of the Synechococcus sp. PCC7942 branching enzyme gene (glgB): expression in Bacillus subtilis. Gene 89:77–84

Kiss JK, Roberts EM, Brown RM et al (1988) X-ray and dissolution studies of paramylon storage granules from Euglena. Protoplasma 146:150–156

Kopp RE, Kirschvink JL, Hilburn IA et al (2005) The Paleoproterozoic snowball Earth: a climate disaster triggered by the evolution of oxygenic photosynthesis. Proc Natl Acad Sci U S A 102:11131–11136

Kotting O, Pusch K, Tiessen A et al (2005) Identification of a novel enzyme required for starch metabolism in Arabidopsis leaves. The phosphoglucan, water dikinase. Plant Physiol 137:242–252

Kotting O, Santelia D, Edner C et al (2009) STARCH-EXCESS4 is a laforin-like Phosphoglucan phosphatase required for starch degradation in Arabidopsis thaliana. Plant Cell 21:334–346

Kubo A, Fujita N, Harada K et al (1999) The starch-debranching enzymes isoamylase and pullulanase are both involved in amylopectin biosynthesis in rice endosperm. Plant Physiol 121:399–410

Ladeira RB, Freitas MA, Silva EF et al (2005) Glycogen as a carbohydrate energy reserve in trophozoites of Giardia lamblia. Parasitol Res 96:418–421

Lafora GR, Glueck B (1911) Beitrag zur histopathologie der myoklonischen epilepsie. Z Gesamte Neurol Psychiatr 6:1–14

Larsson J, Nylander JA, Bergman B (2011) Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. BMC Evol Biol 11:187

Latysheva N, Junker VL, Palmer WJ et al (2012) The evolution of nitrogen fixation in cyanobacteria. Bioinformatics 28:603–606

Le Gall L, Saunders GW (2007) A nuclear phylogeny of the Florideophyceae (Rhodophyta) inferred from combined EF2, small subunit and large subunit ribosomal DNA: establishing the new red algal subclass Corallinophycidae. Mol Phylogenet Evol 43:1118–1130

Linka M, Jamai A, Weber AP (2008) Functional characterization of the plastidic phosphate translocator gene family from the thermo-acidophilic red alga Galdieria sulphuraria reveals specific adaptations of primary carbon partitioning in green plants and red algae. Plant Physiol 148:1487–1496

Lister DL, Bateman JM, Purton S et al (2003) DNA transfer from chloroplast to nucleus is much rarer in Chlamydomonas than in tobacco. Gene 316:33–38

Lluisma AO, Ragan MA (1998) Characterization of a galactose-1-phosphate uridylyltransferase gene from the marine red alga Gracilaria gracilis. Curr Genet 34:112–119

Loddenkotter B, Kammerer B, Fischer K et al (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 U S A 90:2155–2159

Lorenzo-Morales J, Kliescikova J, Martinez-Carretero E et al (2008) Glycogen phosphorylase in Acanthamoeba spp.: determining the role of the enzyme during the encystment process using RNA interference. Eukaryot Cell 7:509–517

Lou J, Dawson KA, Strobel HJ (1997) Glycogen biosynthesis via UDP-glucose in the ruminal bacterium Prevotella bryantii B1(4). Appl Environ Microbiol 63:4355–4359

Lu Y, Sharkey TD (2006) The importance of maltose in transitory starch breakdown. Plant Cell Environ 29:353–366

Lu C, Lei L, Peng B et al (2013) Chlamydia trachomatis GlgA is secreted into host cell cytoplasm. PLoS One 8:e68764

Maddelein ML, Libessart N, Bellanger F et al (1994) Toward an understanding of the biogenesis of the starch granule. Determination of granule-bound and soluble starch synthase functions in amylopectin synthesis. J Biol Chem 269:25150–25157

Manners DJ, Ryley JF (1952) Studies on the metabolism of the Protozoa. II. The glycogen of the ciliate Tetrahymena pyriformis (Glaucoma piriformis). Biochem J 52:480–482

Manners DJ, Ryley JF (1955) Studies on the metabolism of the protozoa. 6. The glycogens of the parasitic flagellates Trichomonas foetus and Trichomonas gallinae. Biochem J 59:369–372

Marin B, Nowack EC, Melkonian M (2005) A plastid in the making: evidence for a second primary endosymbiosis. Protist 156:425–432

Martijn J, Ettema TJ (2013) From archaeon to eukaryote: the evolutionary dark ages of the eukaryotic cell. Biochem Soc Trans 41:451–457

Matsuzaki M, Misumi O, Shin IT et al (2004) Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. Nature 428:653–657

McCracken DA, Cain JR (1981) Amylose in Floridean starch. New Phytol 88

Meeuse BJ, Kreger DR (1954) On the nature of floridean starch and Ulva starch. Biochim Biophys Acta 13:593–595

Meeuse BJ, Andries M, Wood JA (1960) Floridean starch. J Exp Bot 11:129–140

Miao X, Wu Q, Wu G et al (2003) Sucrose accumulation in salt-stressed cells of agp gene deletion-mutant in cyanobacterium Synechocystis sp PCC 6803. FEMS Microbiol Lett 218:71–77

Moran-Zorzano MT, Alonso-Casajus N, Munoz FJ et al (2007) Occurrence of more than one important source of ADPglucose linked to glycogen biosynthesis in Escherichia coli and Salmonella. FEBS Lett 581:4423–4429

Moriyama T, Sakurai K, Sekine K et al (2014) Subcellular distribution of central carbohydrate metabolism pathways in the red alga Cyanidioschyzon merolae. Planta 240:585–598

Mouille G, Maddelein ML, Libessart N et al (1996) Preamylopectin processing: a mandatory step for starch biosynthesis in plants. Plant Cell 8:1353–1366

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

Murakami T, Kanai T, Takata H et al (2006) A novel branching enzyme of the GH-57 family in the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1. J Bacteriol 188:5915–5924

Nagashima H, Nakamura S, Nisizawa K et al (1971) Enzymatic synthesis of floridean starch in red alga Serraticardia maxima. Plant Cell Physiol 12:243–253

Nakamura Y, Takahashi J, Sakurai A et al (2005) Some Cyanobacteria synthesize semi-amylopectin type alpha-polyglucans instead of glycogen. Plant Cell Physiol 46:539–545

Nakayama A, Yamamoto K, Tabata S (2001) Identification of the catalytic residues of bifunctional glycogen debranching enzyme. J Biol Chem 276:28824–28828

Nelson OE, Rines HW (1962) The enzymatic deficiency in the waxy mutant of maize. Biochem Biophys Res Commun 9:297–300

Nowack EC, Melkonian M, Glockner 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 et al (1999) Purification and characterisation of a novel starch synthase selective for uridine 5′-diphosphate glucose from the red alga Gracilaria tenuistipitata. Planta 209:143–152

Nyvall P, Pedersen M, Kenne L et al (2000) Enzyme kinetics and chemical modification of alpha-1,4-glucan lyase from Gracilariopsis sp. Phytochemistry 54:139–145

Ojcius DM, Degani H, Mispelter J et al (1998) Enhancement of ATP levels and glucose metabolism during an infection by Chlamydia. NMR studies of living cells. J Biol Chem 273:7052–7058

Pade N, Linka N, Ruth W et al (2014) Floridoside and isofloridoside are synthesized by trehalose 6-phosphate synthase-like enzymes in the red alga Galdieria sulphuraria. New Phytol. doi:10.1111/nph.13108

Page-Sharp M, Behm CA, Smith GD (1999) Involvement of the compatible solutes trehalose and sucrose in the response to salt stress of a cyanobacterial Scytonema species isolated from desert soils. Biochim Biophys Acta 1472:519–528

Palmer TN, Ryman BE, Whelan WJ (1976) The action pattern of amylomaltase from Escherichia coli. Eur J Biochem 69:105–115

Park JT, Shim JH, Tran PL et al (2011) Role of maltose enzymes in glycogen synthesis by Escherichia coli. J Bacteriol 193:2517–2526

Petersen J, Ludewig AK, Michael V et al (2014) Chromera velia, endosymbioses and the rhodoplex hypothesis – plastid evolution in cryptophytes, alveolates, stramenopiles, and haptophytes (CASH lineages). Genome Biol Evol 6:666–684

Pittenauer E, Schmid ER, Allmaier G et al (1993) Structural characterization of the cyanelle peptidoglycan of Cyanophora paradoxa by 252Cf plasma desorption mass spectrometry and fast atom bombardment/tandem mass spectrometry. Biol Mass Spectrom 22:524–536

Plancke C, Colleoni C, Deschamps P et al (2008) Pathway of cytosolic starch synthesis in the model glaucophyte Cyanophora paradoxa. Eukaryot Cell 7:247–257

Porchia AC, Curatti L, Salerno GL (1999) Sucrose metabolism in cyanobacteria: sucrose synthase from Anabaena sp. strain PCC 7119 is remarkably different from the plant enzymes with respect to substrate affinity and amino-terminal sequence. Planta 210:34–40

Preiss J (1984) Bacterial glycogen synthesis and its regulation. Annu Rev Microbiol 38:419–458

Price DC, Chan CX, Yoon HS et al (2012) Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants. Science 335:843–847

Ragan MA, Bird CJ, Rice EL et al (1994) A molecular phylogeny of the marine red algae (Rhodophyta) based on the nuclear small-subunit rRNA gene. Proc Natl Acad Sci U S A 91:7276–7280

Reyes-Prieto A, Moustafa A, Bhattacharya D (2008) Multiple genes of apparent algal origin suggest ciliates may once have been photosynthetic. Curr Biol 18:956–962

Rodriguez-Ezpeleta N, Brinkmann H, Burey SC et al (2005) Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. Curr Biol 15:1325–1330

Rogers PV, Luo S, Sucic JF et al (1992) Characterization and cloning of glycogen phosphorylase 1 from Dictyostelium discoideum. Biochim Biophys Acta 1129:262–272

Rogers PV, Sucic JF, Yin Y et al (1994) Disruption of glycogen phosphorylase gene expression in Dictyostelium: evidence for altered glycogen metabolism and developmental coregulation of the gene products. Differentiation 56:1–12

Rutherford CL, Peery RB, Sucic JF et al (1992) Cloning, structural analysis, and expression of the glycogen phosphorylase-2 gene in Dictyostelium. J Biol Chem 267:2294–2302

Sanchez-Baracaldo P, Ridgwell A, Raven JA (2014) A neoproterozoic transition in the marine nitrogen cycle. Curr Biol 24:652–657

Sanchez-Puerta MV, Lippmeier JC, Apt KE et al (2007) Plastid genes in a non-photosynthetic dinoflagellate. Protist 158:105–117

Santos CR, Tonoli CC, Trindade DM et al (2011) Structural basis for branching-enzyme activity of glycoside hydrolase family 57: structure and stability studies of a novel branching enzyme from the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1. Proteins 79:547–557

Sawada T, Nakamura Y, Ohdan T et al (2014) Diversity of reaction characteristics of glucan branching enzymes and the fine structure of alpha-glucan from various sources. Arch Biochem Biophys 562C:9–21

Schlichting R, Bothe H (1993) The cyanelle (Organelles of a low evolutionary scale) possess a phosphate-translocator and a glucose-carrier in Cyanophora paradoxa. Botanica Acta 106:428–434

Schneegurt MA, Sherman DM, Nayar S et al (1994) Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium Cyanothece sp. strain ATCC 51142. J Bacteriol 176:1586–1597

Schonknecht G, Chen WH, Ternes CM et al (2013) Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339:1207–1210

Schwoppe C, Winkler HH, Neuhaus HE (2002) Properties of the glucose-6-phosphate transporter from Chlamydia pneumoniae (HPTcp) and the glucose-6-phosphate sensor from Escherichia coli (UhpC). J Bacteriol 184:2108–2115

Schwoppe C, Winkler HH, Neuhaus HE (2003) Connection of transport and sensing by UhpC, the sensor for external glucose-6-phosphate in Escherichia coli. Eur J Biochem 270:1450–1457

Serodio J, Cruz S, Cartaxana P et al (2014) Photophysiology of kleptoplasts: photosynthetic use of light by chloroplasts living in animal cells. Philos Trans R Soc Lond B Biol Sci 369:20130242

Sesma JI, Iglesias AA (1998) Synthesis of floridean starch in the red alga Gracilaria gracilis occurs via ADP-glucose. In: Garab G (ed) Photosynthesis: mechanisms and effects. Kluwer Academic Publishers, Dordrecht, pp 3537–3540

Sheath RG, Hellebust JA, Sawa T (1979) Floridean starch metabolism of Porphyridium purpureum Rhophyta I. Changes during ageing of batch culture. Phycologia 18:292–293

Sheng J, Preiss J (1997) Arginine294 is essential for the inhibition of Anabaena PCC 7120 ADP-glucose pyrophosphorylase by phosphate. Biochemistry 36:13077–13084

Shih PM, Matzke NJ (2013) Primary endosymbiosis events date to the later Proterozoic with cross-calibrated phylogenetic dating of duplicated ATPase proteins. Proc Natl Acad Sci U S A 110:12355–12360

Shimonaga T, Fujiwara S, Kaneko M et al (2007) Variation in storage alpha-polyglucans of red algae: amylose and semi-amylopectin types in Porphyridium and glycogen type in Cyanidium. Mar Biotechnol (NY) 9:192–202

Shimonaga T, Konishi M, Oyama Y et al (2008) Variation in storage alpha-glucans of the Porphyridiales (Rhodophyta). Plant Cell Physiol 49:103–116

Shoguchi E, Shinzato C, Kawashima T et al (2013) Draft assembly of the Symbiodinium minutum nuclear genome reveals dinoflagellate gene structure. Curr Biol 23:1399–1408

Singh PK (1973) Nitrogen fixation by the unicellular blue-green alga *Aphanothece*. Arch Mikrobiol 92:59–62

Stadnichuk IN, Semenova LR, Smirnova GP et al (2007) A highly branched storage polyglucan in the thermoacidophilic red microalga Galdieria maxima cells. Prikl Biokhim Mikrobiol 43:88–93

Stegemann S, Hartmann S, Ruf S et al (2003) High-frequency gene transfer from the chloroplast genome to the nucleus. Proc Natl Acad Sci U S A 100:8828–8833

Steiner JM, Yusa F, Pompe JA et al (2005) Homologous protein import machineries in chloroplasts and cyanelles. Plant J 44:646–652

Streb S, Delatte T, Umhang M et al (2008) Starch granule biosynthesis in Arabidopsis is abolished by removal of all debranching enzymes but restored by the subsequent removal of an endoamylase. Plant Cell 20:3448–3466

Summons RE, Jahnke LL, Hope JM et al (1999) 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. Nature 400:554–557

Suzuki E, Umeda K, Nihei S, Moriya K, Ohkawa H, Fujiwara S, Tsuzuki M, Nakamura Y (2007) Role of the GlgX protein in glycogen metabolism of the cyanobacterium, Synechococcus elongatus PCC 7942. Biochim Biophys Acta 1770:763–773

Suzuki E, Ohkawa H, Moriya K et al (2010) Carbohydrate metabolism in mutants of the cyanobacterium Synechococcus elongatus PCC 7942 defective in glycogen synthesis. Appl Environ Microbiol 76:3153–3159

Suzuki E, Onoda M, Colleoni C et al (2013) Physicochemical variation of cyanobacterial starch, the insoluble alpha-Glucans in cyanobacteria. Plant Cell Physiol 54:465–473

Szydlowski N, Ragel P, Raynaud S et al (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, Turnbull J, Wang W et al (2007) Laforin is a glycogen phosphatase, deficiency of which leads to elevated phosphorylation of glycogen in vivo. Proc Natl Acad Sci U S A 104:19262–19266

Tomos AD, Northcote DH (1978) A protein-glucan intermediate during paramylon synthesis. Biochem J 174:283–290

Torija MJ, Novo M, Lemassu A et al (2005) Glycogen synthesis in the absence of glycogenin in the yeast Saccharomyces cerevisiae. FEBS Lett 579:3999–4004

Tresse E, Kosta A, Giusti C et al (2008) A UDP-glucose derivative is required for vacuolar autophagic cell death. Autophagy 4:680–691

Tuttle RC, Loeblich AR 3rd (1974) Genetic recombination in the dinoflagellate Crypthecodinium cohnii. Science 185:1061–1062

Tuttle RC, Loeblich AR 3rd (1977) N-methyl-N′-nitro-N-nitrosoguanidine and UV induced mutants of the dinoflagellate Crypthecodinium cohnii. J Protozool 24:313–316

Ugalde JE, Parodi AJ, Ugalde RA (2003) De novo synthesis of bacterial glycogen: Agrobacterium tumefaciens glycogen synthase is involved in glucan initiation and elongation. Proc Natl Acad Sci U S A 100:10659–10663

Villarejo A, Martinez F, del Pino Plumed M et al (1996) The induction of the CO2 concentrating mechanism in a starch-less mutant of Chlamydomonas reinhardtii. Physiol Plant 98:798–802

Viola R, Nyvall P, Pedersen M (2001) The unique features of starch metabolism in red algae. Proc Biol Sci 268:1417–1422

Vogel K, Barber AA (1968) Degradation of paramylon by Euglena gracilis. J Protozool 15:657–662

Wattebled F, Buleon A, Bouchet B et al (2002) Granule-bound starch synthase I. A major enzyme involved in the biogenesis of B-crystallites in starch granules. Eur J Biochem 269:3810–3820

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 Physiol 138:184–195

Weber AP, 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

Williams TA, Foster PG, Nye TM et al (2012) A congruent phylogenomic signal places eukaryotes within the Archaea. Proc Biol Sci 279:4870–4879

Williamson BD, Favis R, Brickey DA et al (1996) Isolation and characterization of glycogen synthase in Dictyostelium discoideum. Dev Genet 19:350–364

Wilson WA, Roach PJ, Montero M et al (2010) Regulation of glycogen metabolism in yeast and bacteria. FEMS Microbiol Rev 34:952–985

Wisecaver JH, Hackett JD (2011) Dinoflagellate genome evolution. Annu Rev Microbiol 65:369–387

Wolf A, Kramer R, Morbach S (2003) Three pathways for trehalose metabolism in Corynebacterium glutamicum ATCC13032 and their significance in response to osmotic stress. Mol Microbiol 49:1119–1134

Wyatt JT, Silvey JK (1969) Nitrogen fixation by gloeocapsa. Science 165:908–909

Xu Y, Guerra LT, Li Z et al (2013) Altered carbohydrate metabolism in glycogen synthase mutants of Synechococcus sp. strain PCC 7002: cell factories for soluble sugars. Metab Eng 16:56–67

Yoo SH, Spalding MH, Jane JL (2002) Characterization of cyanobacterial glycogen isolated from the wild type and from a mutant lacking of branching enzyme. Carbohydr Res 337:2195–2203

Yoo SH, Lee BH, Moon Y et al (2014) Glycogen synthase isoforms in Synechocystis sp. PCC6803: identification of different roles to produce glycogen by targeted mutagenesis. PLoS One 9:e91524

Yoon HS, Hackett JD, Pinto G et al (2002) The single, ancient origin of chromist plastids. Proc Natl Acad Sci U S A 99:15507–15512

Yoon HS, Hackett JD, Ciniglia C et al (2004) A molecular timeline for the origin of photosynthetic eukaryotes. Mol Biol Evol 21:809–818

Yoon HS, Müller KM, Stheath RG et al (2006) Defining the major lineages of red algae (Rhodophyta). J Phycol 42:482–492

Yoshikawa H, Nagashima M, Morimoto K et al (2003) Freeze-fracture and cytochemical studies on the in vitro cyst form of reptilian Blastocystis pythoni. J Eukaryot Microbiol 50:70–75

Yu S, Pedersen M (1993) Alpha-1,4-glucan lyase, a new class of starch/glycogen-degrading enzyme. II. Subcellular localization and partial amino-acid sequence. Planta 191:137–142

Yu S, Bojsen K, Svensson B et al (1999) alpha-1,4-glucan lyases producing 1,5-anhydro-D-fructose from starch and glycogen have sequence similarity to alpha-glucosidases. Biochim Biophys Acta 1433:1–15

Yu TS, Kofler H, Hausler RE et al (2001) The Arabidopsis sex1 mutant is defective in the R1 protein, a general regulator of starch degradation in plants, and not in the chloroplast hexose transporter. Plant Cell 13:1907–1918

Zea CJ, Pohl NL (2004) General assay for sugar nucleotidyltransferases using electrospray ionization mass spectrometry. Anal Biochem 328:196–202

Zona R, Chang-Pi-Hin F, O'Donohue MJ et al (2004) Bioinformatics of the glycoside hydrolase family 57 and identification of catalytic residues in amylopullulanase from Thermococcus hydrothermalis. Eur J Biochem 271:2863–2872