

Metabolic Symbiosis and the Birth of the Plant Kingdom

Philippe Deschamps,^{*1} Christophe Colleoni,^{*1} Yasunori Nakamura,[†] Eiji Suzuki,[‡] Jean-Luc Putaux,[‡] Alain Buléon,[§] Sophie Haebel,^{||} Gerhard Ritte,[¶] Martin Steup,[¶] Luisa I. Falcón,[#] David Moreira,^{**} Wolfgang Löffelhardt,^{††} Jenifer Nirmal Raj,^{*} Charlotte Plancke,^{*} Christophe d'Hulst,^{*} David Dauvillée,^{*} and Steven Ball^{*}

^{*}Université des Sciences et Technologies de Lille, Unité Mixte de Recherches 8576 du Centre National de la Recherche Scientifique, Cité Scientifique, Villeneuve d'Ascq, France; [†]Faculty of Bioresource Sciences, Akita Prefectural University, Shimoshinjo-Nakano, Akita, Japan; [‡]Centre de Recherches sur les Macromolécules Végétales (CERMAV-CNRS), Grenoble, France—; Center of Mass Spectrometry of Biopolymers of the University of Potsdam, Golm, Germany; [§]Institut National de la Recherche Agronomique, Centre de Recherches Agroalimentaires, Rue de la Géraudière, Nantes, France; ^{||}Center of Mass Spectrometry of Biopolymers of the University of Potsdam; [¶]Plant Physiology, Institute of Biochemistry and Biology, University of Potsdam, Golm, Germany; [#]Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, México; ^{**}Unité d'Ecologie, Systematique et Evolution, UMR CNRS 8079, Université Paris-Sud, Orsay, France; and ^{††}Max F. Perutz Laboratories, Department of Biochemistry, University of Vienna, Vienna, Austria

Eukaryotic cells are composed of a variety of membrane-bound organelles that are thought to derive from symbiotic associations involving bacteria, archaea, or other eukaryotes. In addition to acquiring the plastid, all Archaeplastida and some of their endosymbiotic derivatives can be distinguished from other organisms by the fact that they accumulate starch, a semicrystalline-storage polysaccharide distantly related to glycogen and never found elsewhere. We now provide the first evidence for the existence of starch in a particular species of single-cell diazotrophic cyanobacterium. We provide evidence for the existence in the eukaryotic host cell at the time of primary endosymbiosis of an uridine diphosphoglucose (UDP-glucose)-based pathway similar to that characterized in amoebas. Because of the monophyletic origin of plants, we can define the genetic makeup of the Archaeplastida ancestor with respect to storage polysaccharide metabolism. The most likely enzyme-partitioning scenario between the plastid's ancestor and its eukaryotic host immediately suggests the precise nature of the ancient metabolic symbiotic relationship. The latter consisted in the export of adenosine diphosphoglucose (ADP-glucose) from the cyanobiont in exchange for the import of reduced nitrogen from the host. We further speculate that the monophyletic origin of plastids may lie in an organism with close relatedness to present-day group V cyanobacteria.

Introduction

Archaeplastida (previously kingdom Plantae according to Cavalier-Smith 1998) are composed of 3 eukaryotic lines: the Glaucophyta (freshwater algae harboring cyanelles [peptidoglycan-containing plastids]), the Rhodophyceae (red algae), and the Chloroplastida (green algae and land plants) (Adl et al. 2005). All 3 lines originate from a single endosymbiotic event involving a cyanobiont related to present-day cyanobacteria and an unicellular heterotrophic eukaryotic host. This unique event brought the ability to perform oxygenic photosynthesis into the eukaryotic world. Simultaneously, all 3 lines gained the ability to synthesize starch, a novel form of storage polysaccharide related to glycogen, the otherwise most widespread form of storage polysaccharide found in living cells (for a review of starch structure see Buléon et al. 1998; for reviews of starch metabolism see Ball and Morell 2003; Zeeman et al. 2007). Importantly, Rhodophyceae and Glaucophytes accumulate starch in their cytoplasm, whereas the latter is exclusively found in the plastids of the Chloroplastida.

Glycogen and starch both consist of α -1,4 linked and α -1,6 branched polymers but differ vastly in their structures and ensuing physicochemical properties. Typically, glycogen consists of tiny homogeneous hydrosoluble particles whose diameter is limited to a maximum of 40 nm. Starch can be defined as a heterogeneous mixture of 2 distinct frac-

tions: the high mass moderately branched amylopectin and the lower mass and less branched amylose. These polysaccharides aggregate in semicrystalline granules of unlimited size. Unlike glycogen, where the α -1,6 linkages are uniformly distributed, the branches of amylopectin are concentrated in certain regions of the molecule. For this reason, the chains resulting from these branches are organized in clusters. Within each cluster, the chains are so tightly packed that they intertwine to form double-helical structures. These double helices align and fall out of solution to aggregate into huge semicrystalline granules (for a detailed review of starch structure see Buléon et al. 1998; otherwise, if needed, a summary of starch structure is provided in supplementary fig. 1 [Supplementary Material online]).

Recent studies have proven that starch synthesis in Rhodophyceae, Glaucophyta, and Chloroplastida consists of a mosaic of genes that originated from cyanobacterial and eukaryotic glycogen metabolisms, suggesting that endosymbiosis involved partners able to synthesize similar types of storage polysaccharides (Coppin et al. 2005; Patron and Keeling 2005; Plancke et al. 2008). Glycogen synthesis is very well conserved among both heterotrophic eukaryotes and most bacteria including cyanobacteria. The pathways of bacterial and eukaryotic glycogen metabolism and that of starch in the green algae and land plants have been intensively studied (for reviews of glycogen metabolism in bacteria and eukaryotes see Preiss 1984; Roach 2002). Briefly, all pathways use a glucosyl-nucleotide substrate to synthesize novel α -1,4 linkages from oligosaccharide or polysaccharide primers. This elongation reaction is catalyzed through the action of the glycogen (starch) synthases. Introduction of the α -1,6 branch through the glycogen (starch)-branching enzymes proceeds in all cases by hydrolysis of a preexisting α -1,4 linkage from an oligosaccharide

¹ These authors contributed equally to this work.

Key words: Endosymbiosis, starch, plastid, glycogen.

E-mail: steven.ball@univ-lille1.fr.

Mol. Biol. Evol. 25(3):536–548. 2008

doi:10.1093/molbev/msm280

Advance Access publication December 18, 2007

or polysaccharide donor chain, followed by transfer and linkage of a segment of this chain in the α -1,6 position. Polysaccharide degradation occurs, in the presence of orthophosphate, through release of glucose-1-P from the terminal glucose of each external chain. However, the glycogen (starch) phosphorylases that catalyze this release are unable to digest the α -1,6 branch. Complete degradation of the polysaccharides thus requires the coordinated action of phosphorylase and glycogen (starch)-debranching enzyme. In addition to this phosphorolytic pathway, a complex hydrolytic pathway also operates. It requires a variable combination of amylases, glucosidases, and α -1,4 glucanotransferases.

Glycogen metabolism in eukaryotes and bacteria differs essentially by 3 criteria. First, UDP-glucose, a glucosyl-nucleotide shared by many other pathways, is the only substrate used by heterotrophic eukaryotes. On the other hand, ADP-glucose is devoted to glycogen synthesis in most bacteria and in all cyanobacteria reported. Second, because of this, ADP-glucose synthesis defines the first committed step of glycogen synthesis, whereas, in eukaryotes, the latter resides in the elongation of the polymers through the glycogen synthases. As a consequence, bacteria control the flux to glycogen through ADP-glucose production by the regulation of ADP-glucose pyrophosphorylase, whereas heterotrophic eukaryotes control their fluxes through posttranslational modifications of glycogen synthases and glycogen phosphorylases. A third lesser emphasized, but important, difference between the bacterial and eukaryotic pathways, can be found in the mechanism used by debranching enzymes to hydrolyze the α -1,6 branch. In eukaryotes a bifunctional enzyme, called indirect debranching enzyme, with 2 distinct active sites, releases glucose from the glycogen particle. In bacteria a more classical direct debranching enzyme directly hydrolyzes the α -1,6 branch and therefore releases an oligosaccharide from glycogen. Because of this, enzymes of oligosaccharide metabolism are required to complete polysaccharide degradation in bacteria, whereas the latter is not required, and indeed not found, as part of glycogen metabolism in heterotrophic eukaryotes (detailed summaries of glycogen metabolism can be found, if needed, in supplementary figs. 2 and 3 [Supplementary Material online]).

Starch metabolism results from the merging of the cyanobacterial and eukaryote host pathways of storage polysaccharide metabolism (for a review see Ball and Morell 2003). Functional studies carried out in green algae and land plants suggest that the major differences between glycogen and starch metabolism consist of a novel function assumed by a subset of direct debranching enzymes during amylopectin synthesis. Isoamylase, a form of direct debranching enzyme unable to digest tightly spaced branches, is thought to be responsible for removing loosely spaced misplaced branches that are randomly formed during synthesis by the branching enzymes (Ball et al. 2006; Myers et al. 2000). These α -1,6 linkages prevent formation of the asymmetric distribution of tightly spaced branches that are required for the aggregation and packaging of amylopectin into insoluble semicrystalline granules. Indeed, mutants defective for isoamylase in green algae and plants revert to the synthesis of glycogen (James et al. 1995;

Mouille et al. 1996; Nakamura et al. 1997; Zeeman et al. 1998; Wattedled et al. 2005). A second difference between starch and glycogen metabolisms consists of the presence of a starch synthase that selectively binds to the semicrystalline granules and catalyzes within the insoluble polysaccharide matrix, the processive elongation of an unbranched glucan that defines amylose. The product of this reaction catalyzed by the granule-bound starch synthase (GBSS), the sole enzyme required for amylose synthesis, escapes the action of branching enzymes which are chiefly present in the soluble phase, as is the case for all other starch metabolism enzymes (for review of amylose synthesis see Ball et al. 1998). The third difference consists in the presence of a very different pathway of polysaccharide degradation (reviewed in Zeeman et al. 2007). The latter is initiated through phosphorylation of the starch granule through the action of glucan water dikinases (GWDs) and phosphoglucan water dikinases (PWDs). This phosphorylation loosens the semicrystalline structure that otherwise resists enzymatic attack (Edner et al. 2007). The phosphorylated starch particle is then digested through the coordinated action of β -amylase (Scheidig et al. 2002) and another form of direct debranching enzyme (Edner et al. 2007). β -amylase produces β -maltose disaccharides by recessing outer chains directly within starch. β -maltose is selectively digested through a form of α -1,4 glucanotransferase called transglucosidase. Transglucosidase releases 1 glucose from β -maltose and transfers the other in α -1,4 position to soluble glycogen-like acceptors (reviewed in Zeeman et al. 2007, a summary of starch metabolism in the green lineage can be found, if needed, in supplementary fig. 4 [Supplementary Material online]).

Whereas both GBSS and isoamylase display phylogenies that group these enzymes with cyanobacterial glycogen synthases and debranching enzymes, the bacterial enzymes do not play similar functions in glycogen synthesis and display very different biochemical properties and substrate preferences (Dauvillée et al. 2005; Suzuki et al. 2007). In addition, the phylogenetic origin of β -amylases, GWDs, PWDs, and transglucosidases are unclear and no comparable enzymes have been reported in bacteria and eukaryotes. They were therefore thought to define plant-specific enzymes.

In this study, we report the presence in cyanobacteria of true starch including amylopectin, amylose packaged with a "bona fide" GBSS in large semicrystalline granules. This enables us to propose the status of polysaccharide metabolism for the cyanobiont before endosymbiosis. Because our present knowledge of eukaryotic glycogen metabolism is confined to opisthokonts (fungi, animals, and choanoflagellates), we then studied the pathway of glycogen metabolism in previously uninvestigated eukaryotic genomes. This approach enabled us to define transglucosidases, β -amylases, and the form of starch synthase found in both Rhodophyceae and glaucophytes, as enzymes that were active for glycogen metabolism within the host cell, well before endosymbiosis. With this knowledge, we are in the position to propose the status of polysaccharide metabolism in the host cell before endosymbiosis. By looking at the distribution of the genes in the 3 Archaeplastida lineages, we were able to reconstruct the minimal set of genes contained in their common ancestor after endosymbiosis.

Unexpectedly, the most plausible enzyme-partitioning scenario for this minimal ancestral set reveals a precise mechanism by which metabolic symbiosis was achieved between the cyanobacterial ancestor of plastids and its host.

Materials and Methods

Strains and Culture Conditions

The axenic nitrogen-fixing cyanobacteria, *Cyanotheca* ATCC51142 was kindly given by Louis A. Sherman of Purdue University. Strain Clg1 was recently isolated by one of us from the North Atlantic (Falcón et al. 2004) and axenized through repeated cloning on solid ASNIII medium (0.9% agar). Both strains were cultured in ASNIII medium (Rippka et al. 1979) at 20 °C at a day/night cycle (12H/12H) under 30- $\mu\text{mol photons}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$.

Starch Purification and Quantification

Cells were harvested after 5–6 days of culture by centrifugation 10 min at $3,000 \times g$ at 4 °C. The pellets were resuspended in water, and cells disrupted using a French press (10,000 Psi). Starch and cell debris were collected by centrifugation ($10,000 \times g$, 15 min) and resuspended in 90% Percoll (GE Healthcare, formerly Amersham Bioscience, Little Chalfont, UK). Due to density differences, starch can be pelleted by centrifugation ($10,000 \times g$, 30 min) away from the bulk of cell debris. The gradient step was repeated once to insure a complete removal of cell debris from the starch pellet. Starch was then washed twice in sterile water. Clean white starch pellets were stored at 4 °C. Starch or α -1,4 glucans amounts were assayed using the Diffchamb Enzyplus Starch kit (Diffchamb, Lyon, France).

In Vitro Synthesis of Amylose

The enzyme activity of GBSS can be assayed as follows: 2 mg of native starch granules were incubated for 12 h at 30 °C in 100 μl of 50 mM Tris/HCl pH 7.5, 0.47% mercaptoethanol, 5.5 mM MgCl_2 , 3.2 mM ADP-Glc, and 2.2 μM ^{14}C radiolabelled ADP-glucose (GE Healthcare, formerly Amersham Bioscience) at 304 mCi/mmol. Tubes were vigorously shaken during incubation, and the reaction was stopped by adding 900 μl 70% ethanol. Granules were washed twice with 1 ml 70% ethanol and dried. Incubated starch granules were boiled in 50 μl of 30% dimethyl sulfoxide (DMSO) at 98 °C during 10 min. The dispersed starch samples were debranched by *Pseudomonas amyloidermosa* isoamylase as described above. Linear glucon chains were subjected to TSK HW50 chromatography running at 10 ml/min in 10% DMSO ($D = 1 \text{ cm H} = 47 \text{ cm}$). A total of 30 μl of each fraction (200 μl) was used to determine the equivalent of glucose following miniaturized procedures (Fox and Robyt 1991). The radioactivity was measured by liquid scintillation counting as described previously (van de Wal et al. 1998).

Extraction and Visualization of Granule-Bound Proteins

A complete procedure for starch granule-bound protein extraction by sodium dodecyl sulfate (SDS) boiling and their separation by SDS page is described in Deschamps et al. (2006). Major proteins were subjected to trypsin digestion and analyzed by QSTAR-QqTOF tandem hybrid system to build a peptide map and sequence the major peptides.

Transmission Electron Microscopy Observation

Starch granules were treated by 2.2 N hydrochloric acid during 48 h, at 37 °C. After washing to neutrality by repeated centrifugation in distilled water, drops of suspensions were deposited onto glow-discharged carbon-coated transmission electron microscopy (TEM) grids and negatively stained with 2% uranyl acetate. The specimens were observed using a Philips CM200 microscope operating at 80 kV. The images were recorded on Kodak SO163 films.

Protein Sequence Alignments and Phylogenetic Trees Construction

Amino acid sequences were aligned using ClustalW (Thompson et al. 1994), and alignment gaps were manually removed. Unrooted Maximum likelihood trees were inferred for 500 bootstrap replicates using ProML (PHYMLIP package, <http://evolution.genetics.washington.edu/phymlip.html>) with the Jones-Taylor-Thornton amino acid change model and a constant rate of site variation. Trees were edited using Retree (PHYMLIP package) and Treeview (Page 1996).

Results

True Starch Can Be Found in Cyanobacteria

Isoamylase, the enzyme explaining the basic difference between starch and glycogen metabolisms displays a distinctive bacterial phylogeny (Patron and Keeling 2005; Suzuki et al. 2007). In addition, no eukaryotic species not suspected to derive directly or indirectly by secondary endosymbiosis from Archaeplastida were ever documented to contain starch-like polymers. Because of this, we suspected that starch might have originated in cyanobacteria.

In cyanobacteria, anomalous glycogen was first described in *Cyanotheca* sp. strain ATCC 51142, a group V single-cell diazotrophic cyanobacterium (Reddy et al. 1993). However, the initial reports focused on the granule morphology of this material and concluded to the presence of glycogen organized in unusually large granules (Schneegurt et al. 1994). Recently, several diazotrophic single-cell cyanobacteria were reported to contain a storage polysaccharide that differed from glycogen by a lower branching ratio, a significantly higher polysaccharide mass and a chain-length distribution that closely resembles amylopectin and was therefore called semiamylopectin (Nakamura et al. 2005). All these bacteria belonged to the same taxonomic subgroup and in addition share the same

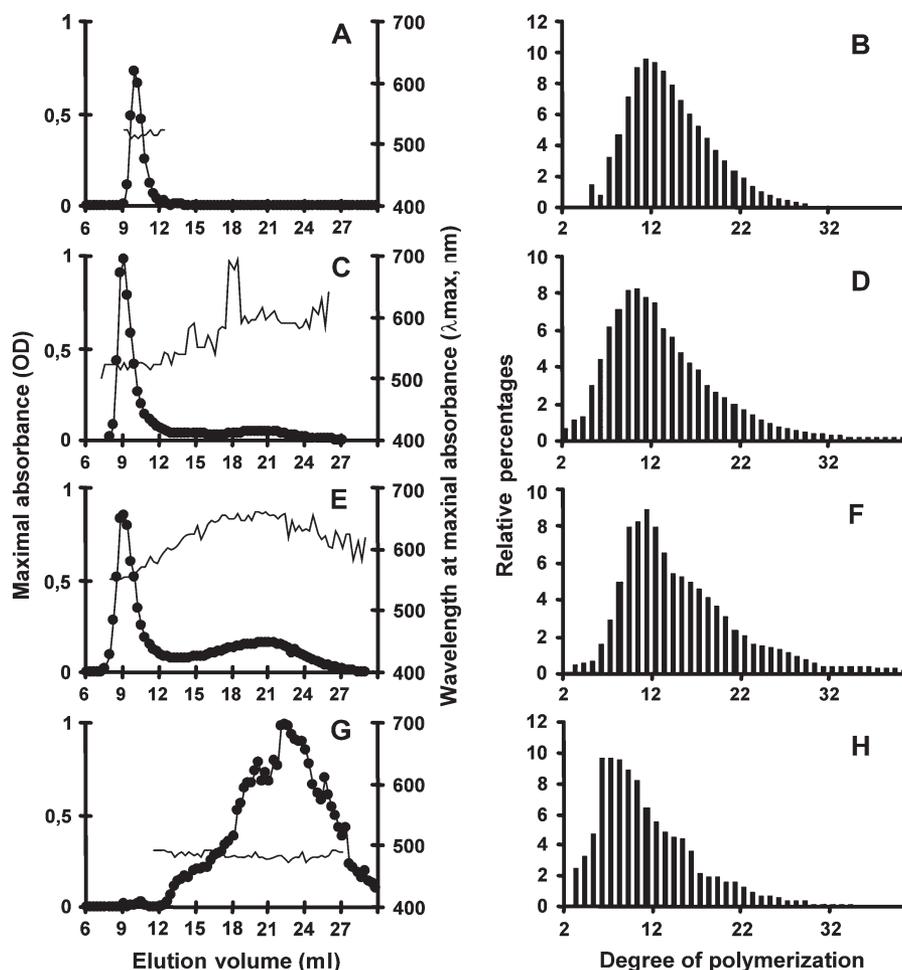


FIG. 1.—Starch structure analysis through size exclusion chromatography and chain-length distribution of amylopectin. Constituent fractions of starch granules were separated using gel permeation chromatography (left column) (see experimental procedures). Amylopectin is a massive molecule that is excluded from the gel, linear amylose, or glycogen molecules are eluted later. Glucans eluted in each fraction were detected by their interaction with iodine. The elution volume and the maximum absorbance (Optical Density) of the iodine–polysaccharide complex (λ_{\max}) of each fraction are indicated by the x axis and the y axis, respectively. Branched glucans of the purified amylopectin fraction are then analyzed by isoamylase-mediated debranching and determination of the chain-length distribution (right column) (see supplementary figures, Supplementary Material online). Chain lengths (Degree of Polymerization) are indicated by the x axis. The percentage of each class of glucans is presented by the y axis. The amylopectin molecule shows a multimodal chain-length distribution with longer chains compared with glycogen. Results of size exclusion chromatography and chain-length distributions are presented, respectively, for *Cyanothece* semiamylopectin (A and B), Clg1 starch (C and D), *Chlamydomonas reinhardtii* starch (E and F), and bovine liver glycogen (G and H), respectively. The chromatography methods are detailed in the methods section.

physiology with respect to nitrogen fixation (see Discussion for further details). This subgroup is called subgroup V, according to the classification of cyanobacteria based on 16S rRNA phylogeny proposed by Honda et al. (1999).

In addition granules significantly larger than those reported for *Cyanothece* were reported in another subgroup V diazotrophic cyanobacterium (Clg1) that was recently isolated by one of us (Falcón et al. 2004). However, the detailed structure of the polysaccharide was not investigated. This strain behaves like *Cyanothece* with respect to nitrogen fixation and also selectively induced nitrogenase during the night (supplementary fig. 5, Supplementary Material online). We investigated in greater detail, the structures of these polysaccharides and ascertained their relationship to starch and glycogen. In a first series of experiments, we separated amylopectin amylose and glycogen through gel permeation chromatography (fig. 1A, C, E,

and G). We then purified the high mass amylopectin-like from figure 1A, C and E and the glycogen from figure 1G and subjected these polysaccharides to enzymatic debranching followed by separation of the debranched chains by capillary electrophoresis, thereby yielding their chain-length distributions (figs 1B, D, F, and H). Results show that *Cyanothece* and strain Clg1 synthesize a large mass storage polysaccharide (fig. 1A and C) related to those recently described for other group V cyanobacteria (Nakamura et al. 2005). However, strain Clg1 contained in addition a small but significant amount of another polysaccharide fraction (fig. 1C) eluting later on the gel permeation column with an iodine–polysaccharide interaction similar to that of the amylose from *Chlamydomonas reinhardtii* (fig. 1E). Plant starch is defined not only by the presence of both amylose and amylopectin but also by the typical semicrystalline organization of its amylopectin fraction

(supplementary fig. 6, Supplementary Material online). Wide-angle X-ray diffraction and small-angle X-ray scattering analysis (supplementary fig. 6, Supplementary Material online) demonstrate the presence of patterns and structures within the cyanobacterial polysaccharides that are identical to those of plant starch. Finally, we were able to directly visualize the shape and organization of the crystallites within the granules by TEM (fig. 2). The crystallites were revealed after partial acid hydrolysis that selectively digests the amorphous material. Crystalline platelets can be seen in all samples to aggregate in concentric layers of similar organization. No such platelets are ever seen when glycogen is subjected to similar treatments.

The presence of amylose in plants is correlated to the presence of an enzyme that was thought as a plant specific: GBSS. In all cases documented so far, GBSS defines the major protein found associated to starch granules and the sole enzyme responsible for the processive synthesis of the long glucans found in amylose. We investigated the identity of the starch-bound proteins in the amylose-containing strain Clg1 and compared these with those associated with the amylose-less polysaccharide from *Cyanothece* sp. strain ATCC 51142. Results displayed in figure 3A demonstrate that strain Clg1 contains a 57 kDa GBSS-like protein. We investigated the amylose-synthesizing ability of polysaccharide granules purified from the 2 cyanobacteria by studying in vitro the incorporation of radioactive ADP-glucose within the granules (fig. 3B). Results shown in figure 3 prove that Clg1 granules only synthesize long glucan chains in vitro and therefore contain a bona fide GBSS.

From all these results, we conclude that strain Clg1 accumulates true starch whereas an amylose-less form of the polysaccharide is synthesized by *Cyanothece* sp. strain ATCC 51142.

Phylogenetic Analysis and Distribution of Storage Polysaccharide Metabolism Genes Enables to Define the Minimal Set of Enzymes Present in the Ancestor of All Plants

Despite the finding of starch in group V cyanobacteria and despite the previously published detailed phylogenies for ADP-glucose pyrophosphorylases, starch (glycogen) synthases, branching enzymes, GWD, and isoamylases (Coppin et al. 2005; Patron et al. 2005; Suzuki et al. 2007), a number of very important enzymes of starch metabolism (GWD, PWD, rhodophycean starch synthase, β -amylase, and transglucosidase) have not yet been assigned a clear host or endosymbiont origin. We argue that this is chiefly because the heterotrophic eukaryotic genomes probed in these studies were restricted to fungi and animals.

We have thus readdressed the trees built for starch (glycogen) synthases (fig. 4) and branching enzymes (fig. 5) to include the sequences obtained for cryptophytes (Deschamps et al. 2006), glaucophytes, rhodophytes, group V diazotrophic cyanobacteria, and most importantly from the recently reported heterotrophic eukaryote genome sequences of *Dictyostelium discoideum* (Eichinger et al. 2005), *Entamoeba histolytica* (Loftus et al. 2005), *Parame-*

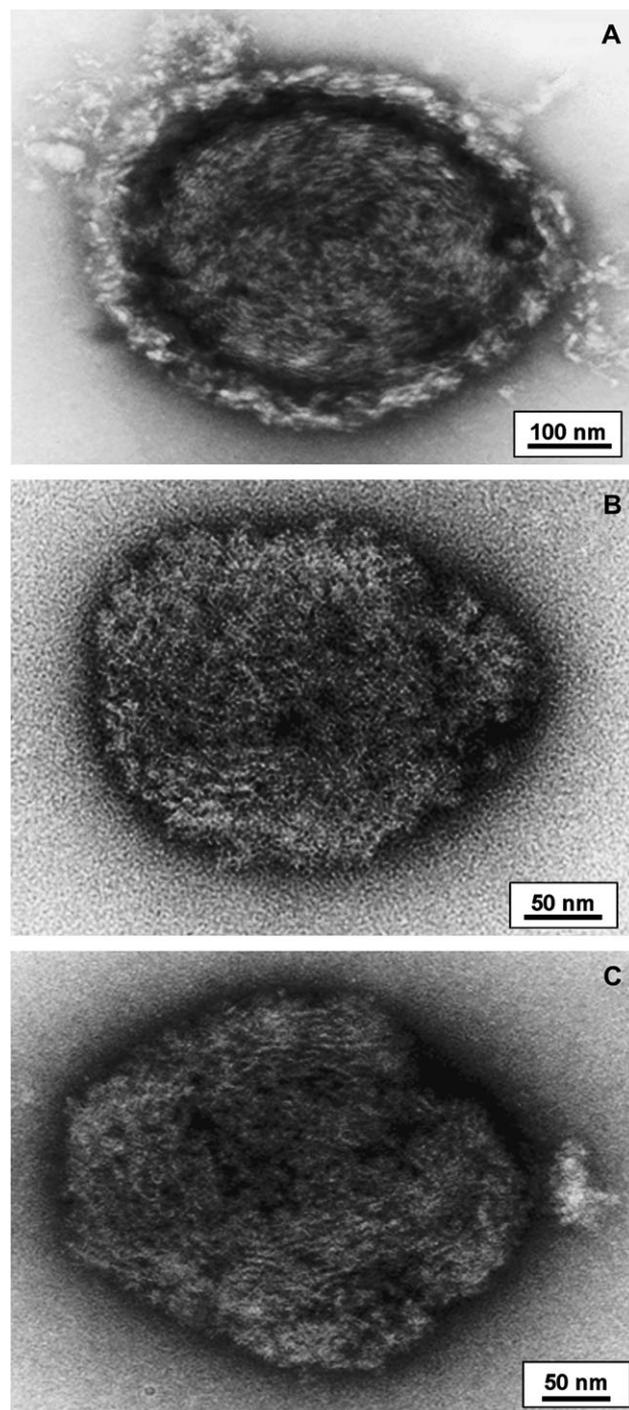


FIG. 2.—TEM of partially acid-hydrolyzed starch granules. Native starch granules were partially hydrolyzed in HCl and observed by TEM. Under such circumstances, amorphous polysaccharides (such as glycogen) or sections of amorphous materials in otherwise crystalline zones would be completely hydrolyzed. The images thus reveal the lamellar radial organization of amylopectin crystallites: (A) a typical amylose free *Chlamydomonas* starch granule, (B) a *Cyanothece* semiamylopectin granule, (C) and a starch granule from strain Clg1.

cium tetraurelia (Aury et al. 2006), *Tetrahymena thermophila* (Eisen et al. 2006), and *Trichomonas vaginalis* (Carlton et al. 2007). Amoebas such as *D. discoideum* are indeed thought to have diverged earlier than the

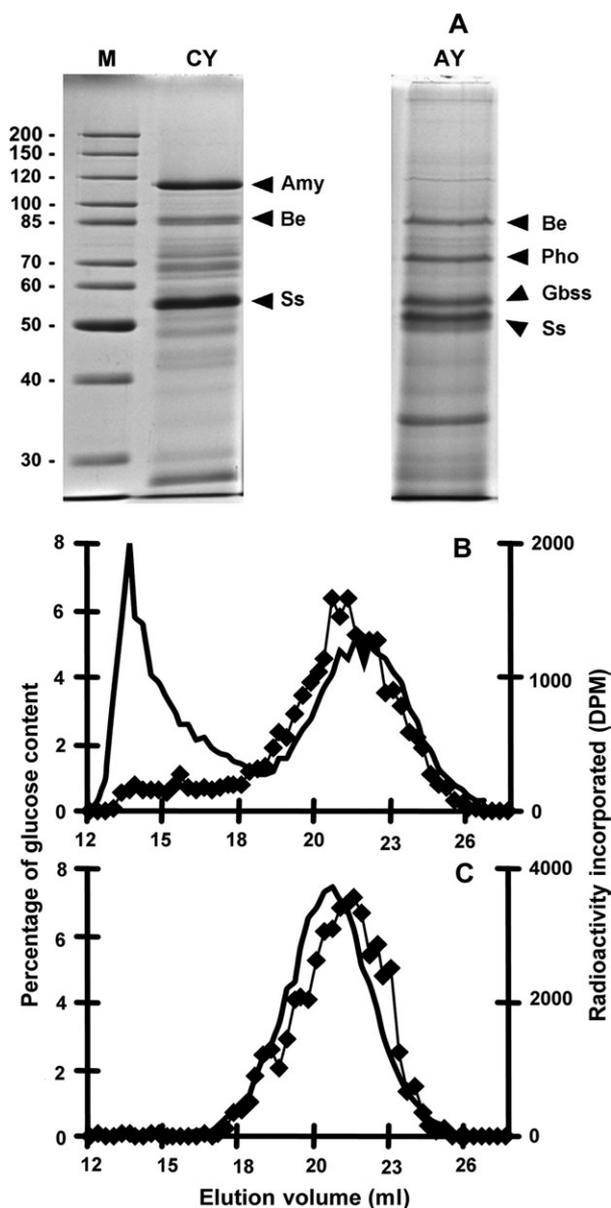


FIG. 3.—Observation and assay of enzymes bound to cyanobacteria polysaccharide granules. (3A) Proteins specifically bound to polysaccharide granules from *Cyanothecce* (CY) and *Clg1* (AY) were extracted and observed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (see experimental procedures). Major proteins were identified by trypsin digestion followed by Matrix Assisted Laser Desorption Ionisation–Time Of Flight and Tandem Mass Spectrometry peptide sequencing. Three enzymes were identified in *Cyanothecce* semiamylopectin: an amylase (Amy), a branching enzyme (Be), and a glycogen synthase (Ss). Four enzymes were identified in *Clg1* starch granules: a glucan-phosphorylase (Pho), a branching enzyme (Be), and 2 glycogen synthase (Ss) with one more related to the GBSS-like family (Gbss). (see supplementary fig. 9 [Supplementary Material online] for peptide sequences and homologies). (3B and 3C) Granule-bound starch/glycogen synthases are still active inside the purified starch granules. By incubation with radiolabeled ADP-glucose, these enzymes will elongate glucans from the polysaccharide matrix. The nature of the glucans produced by granule-bound glycogen/starch synthases was assayed for strain *Clg1* (B) and for *Cyanothecce* (C). After *in vitro* synthesis, granules were fractionated by Cl-2B gel permeation chromatography, purified amylopectin was debranched by isoamylase, and linear glucans were separated by TSK–HW50 gel permeation chromatography (see experimental procedures). Longer chains are eluted first. Glucans in fractions were detected by iodine interaction, assay of total glucose and by scintillation counting. The elution volume is indicated

opisthokonts (animals and fungi) and are observed to contain a larger set of eukaryotic genes (Loftus et al. 2005; Song et al. 2005). In addition to the aforementioned genes, we also report phylogenies for α -1,4 glucanotransferases (including D-enzyme and transglucosidase) (fig. 6), starch phosphorylases (supplementary fig. 7, Supplementary Material online), and for β -amylase (supplementary fig. 8, Supplementary Material online) despite the rather low number of sequences available for the latter enzymes. These phylogenetic trees enabled us to propose the endosymbiotic or host origin for all genes but one (table 1). Indeed, figure 4 confirms that the ADP-glucose-specific starch synthase from the Chloroplastida group with cyanobacterial enzymes. However, the tree equally shows that the UDP-glucose-specific starch synthases present in Rhodophyceae and Glaucophyta (Plancke et al. 2008) do not group with the latter but rather with enzymes of the same family found in various heterotrophic eukaryotes such as amoebas, parabasalids, and ciliates. It is worth stressing that amoebas and parabasalids are not suspected to result from secondary endosymbiosis of an Archaeplastida symbiont. The status of ciliates is unclear in this respect but nevertheless and in stark contrast with apicomplexan parasites, the sequencing of both *Tetrahymena* and *Paramecium* genomes failed to recover genes displaying phylogenies grouping these organisms with Rhodophyceae. Figure 5 confirms the eukaryotic (host) origin for all the starch-branching enzymes found in Archaeplastida. Figure 6 clearly shows that the *dpe1* (also called D-enzyme) type of α -1,4 glucanotransferase groups with cyanobacterial enzymes, whereas the *dpe2* (also called transglucosidase) groups with enzyme sequences found in amoebas and parabasalids.

Finally, both phosphorylase and β -amylase sequences found in Archaeplastida group with heterotrophic eukaryote sequences in trees that therefore support a host origin for the corresponding genes. (supplementary figs. 7 and 8, Supplementary Material online).

Because of the finding of true starch in cyanobacteria and because of the finding of additional genes involved in glycogen metabolism absent from opisthokonts, we chose the completed genomes of the group V diazotrophic cyanobacterium *Crocospaera watsonii* and *E. histolytica* as our paradigms of cyanobacterial starch and heterotrophic eukaryote glycogen synthesis, respectively, and the *Ostreococcus tauri* genome as a paradigm for starch synthesis occurring in the green plastids. Indeed, the number and function of the starch metabolism genes are conserved with very little variation from the prasinophycean *O. tauri* to the chlorophycean *C. reinhardtii* and angiosperms such as *Arabidopsis thaliana* (Ral et al. 2004; Derelle et al. 2006). As a paradigm for the red lineage, we compiled results obtained with the apicomplexan parasite genome of *Toxoplasma gondii* (Coppin et al. 2005) or with the cryptophyte

←

by the *x* axis, the left *y* axis corresponds to the relative percentage of glucose present in each fraction (\blacklozenge), and the right *y* axis presents the radioactivity (Dots per minute) incorporated into the corresponding glucans (—). Whereas short chains produced by a regular starch/glycogen synthase were found in both species, long-radiolabeled glucans were only observed for *Clg1*. Such chain lengths are characteristic of GBSS-like enzyme products.

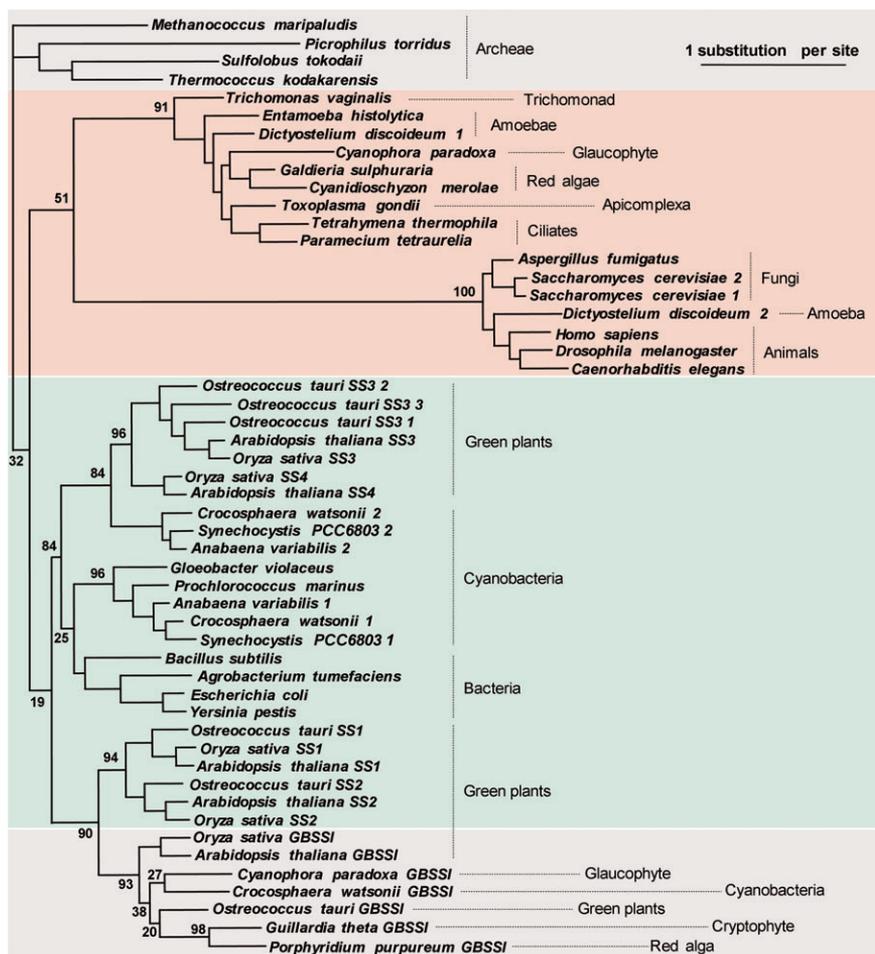


FIG. 4.—Maximum of likelihood unrooted tree inferred for starch and glycogen synthases. The scalebar represents the branch length corresponding to 0.5 substitution per site. Bootstrap values calculated for 500 replicates are indicated at corresponding nodes. Colored underlines define the nature of the substrate used by these enzymes as follow: red for UDP-glucose, blue for ADP-glucose, and gray for a mixed use of both. Three groups can be clearly distinguished: « yeast-like » UDP-glucose-specific glycogen synthases, UDP-glucose using glycogen/starch synthase found in amoebae and the red algae lineage, and a last group containing bacterial, cyanobacterial, archeal glycogen synthases together with plants starch synthases. All GBSS proteins, whatever the nature of their substrate (ADP-glucose or UDP-glucose), group together, and denote a common origin.

Guillardia theta (Deschamps et al. 2006) or the cyanidiales *Cyanidioschyzon merolae* and *Galdieria sulphuraria* (Coppin et al. 2005). Indeed, we have previously shown that most *Toxoplasma* genes listed in table 1 (with the exception of branching enzymes) are phylogenetically related to those of cyanidiales such as *C. merolae* and *G. sulphuraria* (Coppin et al. 2005). In addition, *Guillardia theta* was reported to contain an UDP-glucose-utilizing GBSSI responsible for amylose synthesis (Deschamps et al. 2006). Although cyanidiales or apicomplexan polysaccharides clearly do not contain amylose, other unicellular red algae such as *Porphyridium* or *Rhodella* were reported to contain a substantial amount of this polysaccharide fraction, and the complete sequence for a *Porphyridium* GBSSI gene was obtained by one of us (GenBank accession number AB274917).

From table 1 and all our detailed phylogenomics studies, we found that 3 enzymes correlated perfectly with the presence of starch in eukaryotes: direct debranching enzyme, D-enzyme, and GWD.

Direct debranching enzymes (see fig. 1 for a definition) of bacterial origin were present in all starch-storing organ-

isms and were never found in glycogen-accumulating eukaryotes. Indeed, mutants of plants and green algae defective for a specific form of direct debranching enzyme known as isoamylase substitute starch by glycogen synthesis. Because of this, we had initially proposed that isoamylase was responsible for generating the aggregated semicrystalline form of amylopectin (Ball et al. 1996; Mouille et al. 1996). This proposal seems largely confirmed by the distribution of the isoamylase-related gene within eukaryotes (Coppin et al. 2005). The clear cyanobacterial origin of isoamylase suggests that starch may have appeared first in these organisms. This conclusion is now supported by our finding of starch in group V diazotrophic single-cell cyanobacteria.

GWDs are enzymes known to phosphorylate amylopectin at the C3 and C6 position (reviewed in Zeeman et al. 2007). Functional analysis in *Arabidopsis* and potato proves that the reaction catalyzed by GWD is required to initiate starch breakdown that then proceeds through the combined action of β -amylases and transglucosidase. The origin of GWD is unclear, the protein being absent

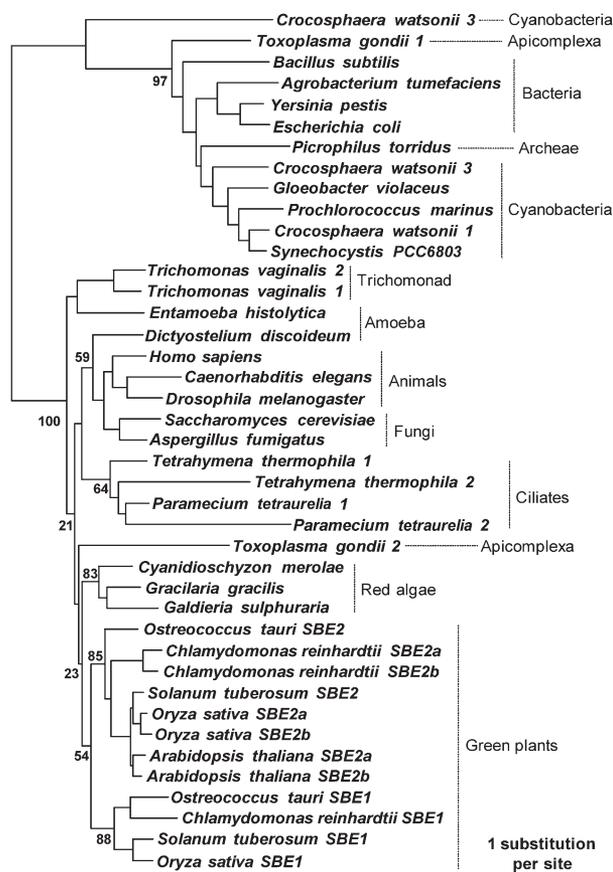


FIG. 5.—Maximum likelihood unrooted tree inferred for alpha-glucan-branching enzymes. The scalebar represents the branch length corresponding to 1 substitution per site. Bootstrap values calculated for 500 replicates are indicated at corresponding nodes. The tree strongly supports that all plants branching enzymes are related to nonphotosynthetic eukaryotes branching enzymes.

from all cyanobacteria and from all heterotrophic eukaryote genomes presently available.

Interestingly, the red algae and glaucophytes contain a GT5 UDP-glucose-specific soluble starch synthase. These sequences markedly differ from the glycogen synthase of fungi and animals that belong to another family (GT3) of glycosyltransferases according to the Carbohydrate Active Enzymes classification (Coutinho and Henrissat 1999). For this reason, it was assumed that this synthase evolved through mutation from a cyanobacterial GT5 ADP-glucose-specific soluble starch synthase after endosymbiosis (Patron et al. 2005). However, the GT5 sequences are now shown to be present in both Archamoebas and Mycetozoa and also in other very distant heterotrophic eukaryotes such as *T. vaginalis* (fig. 4). This strongly suggests that the gene encoding this enzyme is of ancient eukaryote origin. Indeed, *D. discoideum* contains both the GT3 and the GT5 glycogen synthases, whereas the latter was clearly lost in opisthokonts.

From table 1, starch synthesis in the red lineage can be seen as very similar to glycogen metabolism in amoebas. It contains all enzyme types present in *E. histolytica* in addition to GWD and to direct debranching enzyme and GBSSI, the latter 2 being clearly of cyanobacterial origin. Table 1 clearly suggests that a complete (for Rhodophyceae and pos-

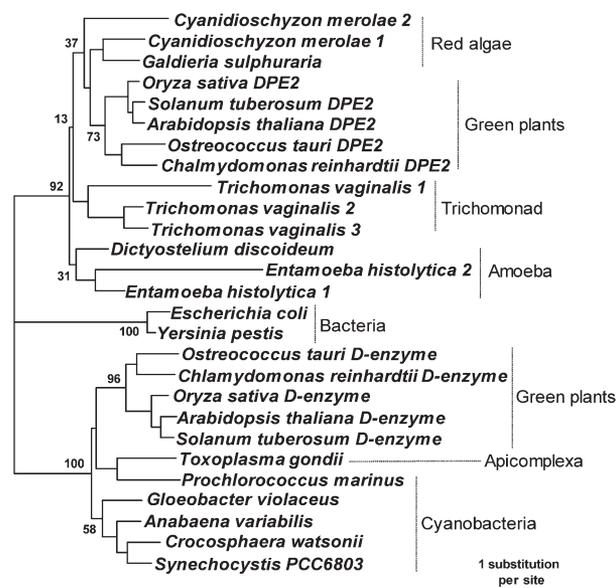


FIG. 6.—Maximum likelihood unrooted tree inferred for α -1,4-glucanotransferases. The scalebar represents the branch length corresponding to 1 substitution per site. Bootstrap values calculated for 500 replicates are indicated at corresponding nodes. Two groups can be clearly observed. The first one contains amyloamylase-like proteins from amoeba and plants. The second one shows that plants disproportionating enzymes are related to cyanobacterial enzymes.

sibly glaucophytes) or near to complete pathway (for Chloroplastida, who are only missing the GT5 UDP-glucose-utilizing glycogen [starch] synthase) of heterotrophic eukaryotic glycogen metabolism (as exemplified by *E. histolytica*) was maintained in all lines derived from primary endosymbiosis. Interestingly, both amoebas and the parabasalids not only contain an additional GT5 UDP-glucose-utilizing glycogen (starch) synthase not found in opisthokonts but they also contain β -amylase and *dpe2*, enzymes that have been demonstrated to be responsible for the degradation of starch in green plants into β -maltose and for the degradation of this disaccharide into glucose in the presence of a soluble glycogen type of acceptor. The fact that the distantly related Archamoebas, Mycetozoa, and parabasalids all contain *dpe2* and GT5 UDP-glucose-utilizing glycogen (starch) synthase argues that these genes have not been acquired through recent lateral transfers from plants where they have been first found and characterized. Indeed, the Archamoeba and Mycetozoa split is ancient whereas that of the amoebas and parabasalids is expected to be far older, possibly before endosymbiosis of the plastid.

Because of the established monophyly of Archaeplastida (Rodríguez-Ezpeleta et al. 2005), we are in the position to define the minimal gene content of their common ancestor (table 1). This was deduced by restricting the number of genes encoding enzymes of the green lineage to the number of their putative ancestral genes nowadays still contained in group V cyanobacteria or in amoebas. Indeed, we suspect that duplications of all the green algae genes occurred early after their divergence from the red algae. The mechanisms underlying the selective duplications of the Chloroplastidal sequences will be detailed elsewhere. The minimal set numbers listed in table I were derived as follows:

Table 1
Storage Polysaccharides Metabolism Enzyme Sets

	Cyanobacteria (<i>Crocospheara watsonii</i>)	Eukaryotes (<i>Entamoeba histolytica</i>)	Minimal Set for the Common Ancestor	Green lineage (<i>Ostreococcus tauri</i>)	Red Lineage Compiled Minimum set	Glaucophytes Minimal Set	Literature Cited
ADP-glucose pyrophosphorylase	1	—	1	2	—	?	Patron and Keeling (2005)
Soluble starch synthase (ADPG)	2	—	2	5	—	?	This work
Soluble starch synthase (UDPG)	—	1	1	—	1	1	This work
GBSS I	1	—	1	1	1	1	This work
Branching enzyme	3	1	1	2	1	1	Coppin et al. (2005); Patron and Keeling (2005), This work
Isoamylase	1	—	1	3	1	1	Patron and Keeling (2005)
Indirect debranching enzyme	—	1	1	—	1	?	This work
Phosphorylase	2	2	1	2	1	1	This work
Glucanotransferase	1	—	1	1	1	?	This work
Transglucosidase	—	2	1	1	1	?	This work
Beta-amylase	—	4	1	2	1	?	This work
GWD	—	—	(1)	(4)	(1)	?	—

NOTE.—The number of isoforms found for each class of glycogen/starch metabolism enzymes was listed. Using phylogenetics, we could determine the origin of each isoform in the red and green lineages except for GWDs. Enzymes of cyanobacterial phylogeny are highlighted in blue. Enzymes of eukaryotic origin are highlighted in beige. Enzymes of uncertain origin are listed between brackets. The cyanobacterial eukaryotes and green plants display highly conserved sets of enzymes. We chose *Crocospheara watsonii*, *Entamoeba histolytica*, and *Ostreococcus* as paradigm genomes for, respectively, cyanobacteria, heterotrophic eukaryotes, and green plants. The information concerning rhodophytes was compiled from several genomes as explained in the text. The information concerning glaucophytes was provided from Plancke et al. 2008. The absence of sequenced genomes in glaucophytes prevented us to give clear negative answers for the presence of some enzymes. This was symbolized by a question mark.

Briefly, the enzymes of ADP-glucose metabolism are unquestionably of cyanobacterial origin. The minimal set comprises 2 soluble starch synthase genes, 1 GBSS and 1 ADP-glucose pyrophosphorylase subunit. Figure 4 shows that one of the cyanobacterial soluble starch synthase family of sequences clearly groups with the SSIII, SSIV, and SSV Chloroplastida sequences. In the case of SSI and SSII, the phylogenetic tree does not show more or less relatedness between these enzymes and other specific subgroups of GT5 bacterial glycogen synthase. They are, however, much more related to these enzymes than to the other UDP-glucose-utilizing enzymes. This result is thus perfectly consistent with a unique cyanobacterial origin for both the SSI and SSII sequences that was followed by more sequence divergence than in the case of SSIII, SSIV, and SSV. Because many cyanobacterial genomes and certainly all available group V cyanobacterial genomes contain 2 distinct soluble glycogen (starch) soluble types, the results displayed in figure 4 are in good agreement with the transfer to the ancestor of Archaeplastida of the 2 genes encoding each of the cyanobacterial soluble starch synthase. Figure 4 also convincingly shows that the cyanobacterial GBSS sequence groups with those of Archaeplastida and shows more relatedness to the Glaucophyta UDP-glucose-preferring enzyme and to Chloroplastida ADP-glucose-preferring GBSS sequences than to those of UDP-glucose-preferring Rhodophyceae or cryptophytes. One such GBSS sequence has thus to be added to the minimal gene list. In addition to the ADP-glucose-specific soluble starch synthase, figure 4 also clearly supports the existence in the ancestor of all Archaeplastida of one host-derived GT5 UDP-glucose-specific soluble starch

synthase found in the genomes of glaucophytes and Rhodophyceae. The obvious relatedness between the cyanobacterial and Archaeplastida ADP-glucose pyrophosphorylase subunits has been documented elsewhere (Patron and Keeling 2005). The complexity of the multiple ADP-glucose pyrophosphorylase subunits seen in plants and green algae was probably generated after separation of the Chloroplastida from the Rhodophyceae by mechanisms similar to those yielding the multiple forms of soluble starch synthase. Only one gene-encoding ADP-glucose pyrophosphorylase subunit was thus added to the list.

Both figure 5 and supplementary figure 7 (Supplementary Material online) show trees that support the grouping of, respectively, starch-branching enzyme and starch phosphorylase with their heterotrophic eukaryote-corresponding sequence. As mentioned above, we can safely conclude that these genes are of host origin. In addition, once again, the BEI and BEII families have apparently been generated after separation of the green algae from Rhodophyceae. Therefore, a minimum of 1 branching enzyme and 1 phosphorylase of host origin only must be added to our minimum gene set in table I.

As to the debranching enzymes, the presence of isoamylase was revealed in all Archaeplastida lineages either by bioinformatics study (Coppin et al. 2005) or by activity detection (Plancke et al. 2008). We suspect that the diversification of isoamylase into distinct subunits also came after separation of Chloroplastida from the Rhodophyceae. Indirect debranching enzyme-related sequences were found in apicomplexans (Coppin et al. 2005) and may also have been present in the common ancestor. However, caution is needed here because there is no evidence for the

proteins were actually located shortly after endosymbiosis and if additional genes of cyanobacterial origin were actually maintained in the ancestor. We now propose and strongly support that the most likely of all possible compartmentalization scenarios consist of storage polysaccharide metabolism loss by the cyanobiont at a very early stage, possibly right at the beginning, well before sophisticated machineries allowing for the targeting of proteins from the cytoplasm to the cyanobiont had evolved.

This assumption is based on several distinct observations. First, Henrissat et al. (2002) established a clear correlation between loss of glycogen metabolism in bacteria and the degree of parasitism displayed by pathogens. Becoming an obligatory endosymbiont would, according to this view, automatically lead to the loss of storage polysaccharide synthesis by the symbiont.

Second, loss of storage polysaccharide can indeed be evidenced in all available genomes of obligatory endosymbionts with no known extracellular stage such as those recently sequenced occurring in insect cells (Gil et al. 2004).

Third, the concomitant loss of the multiple forms of cyanobacterial-branching enzyme (3 forms in all 3 genomes available for group V cyanobacteria) and phosphorylase genes in the green and red algae as well as in glaucophytes (all these enzymes display eukaryotic phylogenies in all plant lines) argues that these losses occurred at a very early stage. This observation is certainly in line with storage polysaccharide metabolism loss in the cyanobiont.

Finally, the sole presence of the eukaryotic phosphorylases and β -amylase-DPE2 pathways of glycogen degradation is surprising when faced with the problems inherent to degradation of crystalline structures. This is also in line with a very early loss of the cyanobacterial pathway of starch mobilization. Indeed, these cyanobacterial pathways were certainly better adapted than the host glycogen degradation machinery for this purpose.

The Proposed Pathway of Storage Polysaccharide Synthesis in the Ancestor of All Archaeplastida Defines the Nature of the Ancient Symbiotic Fluxes

Successful organelle evolution is a 2-step process. In a first step, some kind of metabolic symbiosis must be immediately established after internalization. However, this metabolic symbiosis must function effectively in the absence of an organelle protein targeting machinery that will take time to evolve. Nevertheless, this step turns the cyanobacterium into an obligatory symbiont. It involves for the future plastid, a first wave of both gene losses and gene transfers from the organelle to the host nucleus and their expression in host compartments under the control of host promoters. However, targeting of proteins back to the symbiont is not possible and no significant transfer of host genes to the organelle DNA has been evidenced. Therefore, gene transfers would very quickly establish novel functions outside the future plastid.

The second step of organelle evolution consists of the appearance of a sophisticated and effective plastid targeting machinery allowing for the import of proteins into the future organelle. This will allow for a complete reshuffling of host and symbiont functions probably accompanied by the

diversification of the 3 major lines of Archaeplastida. At this stage, the symbiont can be considered as an organelle, although some authors do not consider this a requirement to distinguish endosymbionts from organelles (Bodyl et al. 2007). This second step leads to a second wave of gene transfers and gene losses from the endosymbiont, thereby generating the plastid genomes that we observe today.

In accordance with Henrissat et al. (2002), we propose that storage polysaccharide synthesis was lost from stage I endosymbionts. In this view, to be maintained by natural selection, the cyanobacterial enzymes of storage polysaccharide synthesis must have been expressed after transfer to the host nucleus and used immediately in the cytoplasm for starch synthesis. This suggests that ADP-glucose must have been present in the cytosol. Indeed, ADP-glucose pyrophosphorylase is present in the green lineage plastids and must have been maintained in the ancestor of all plants. That ADP-glucose must have been present in the cytosol is further suggested by the presence of GBSSI in the cytoplasm of both Rhodophyceae and glaucophytes. Indeed in these cases, the successful transfer of this cyanobacterial enzyme to the cytoplasm suggests that the corresponding gene must have been maintained in the nuclear genome by natural selection for expression in the cytosol. This must have been long enough to allow for the successive mutations turning GBSSI from an ADP-glucose-specific activity to an enzyme able to use UDP-glucose more efficiently as is the case for Rhodophyceae and glaucophytes. We argue therefore that ADP-glucose must have been immediately present in the cytosol for natural selection to allow the maintenance of cytosolic GBSSI. We infer that among the functions listed in table 1, ADP-glucose pyrophosphorylase was the only enzyme that was not transferred at this stage. Keeping this enzyme expressed within the endosymbiont would have allowed to produce a carbon committed to storage, whereas keeping the enzyme tuned to photosynthate availability through its original cyanobacterial allosteric effectors, a regulation that was indeed conserved throughout the green lineage.

The maintenance of ADP-glucose synthesis in the endosymbiont requires the presence of a nucleotide sugar translocator on the symbiont envelopes that would export ADP-glucose in the cytosol in exchange for ADP. Interestingly, a recent report tracks the monophyletic origin of a diversity of plastid translocators from red and green algae to a nucleotide-sugar/triose phosphate translocator gene family that originates from the host endomembrane system (Weber et al. 2006). In effect, with this compartmentalization scenario, the ancient symbiosis metabolic fluxes immediately come to the light (fig. 7). The latter consisted in the export of ADP-glucose from the cyanobiont for cytoplasmic starch synthesis in exchange for the import of reduced nitrogen.

Cytosolic Starch Synthesis through ADP-Glucose Export from the Symbiont Defines an Efficient Mean to Establish Endosymbiosis

In effect, export of ADP-glucose for storage polysaccharide synthesis in the host cytosol would have been a very efficient means to establish metabolic symbiosis. We suggest that export of carbon by the cyanobacterium in

exchange for the import of reduced nitrogen formed the basis of plastid endosymbiosis. Such a symbiotic relation is astonishingly easy to set up through storage polysaccharide metabolism. Indeed, both the host and the cyanobacterium had an UDP-glucose-based glycogen and ADP-glucose-based starch synthesis pathway. We now know that the transfer of only a few genes of the ADP-glucose (a minimum of 1 isoamylase, 1 ADP-glucose-utilizing starch synthase, and 1 dpe1-like α -1,4 glucanotransferase)-based pathway to the nucleus of the host and their expression under control of eukaryotic promoters would have been sufficient to turn cytosolic host glycogen into cytoplasmic starch (Mouille et al. 1996; Colleoni et al. 1999). The presence of starch in the cytoplasm and the concomitant loss of storage polysaccharide synthesis in the endosymbiont would have created a very strong and permanent sink for carbon in the former compartment. Fueling photosynthate into storage material is a very convenient way to establish endosymbiosis because the host and future organelle metabolic networks are not connected to start with. Storage of carbon into starch in the cytoplasm would thus occur whenever the endosymbiont metabolism allows it independently of the cytoplasmic concentration of UDP-glucose and hexoses regulated by metabolism of the host. This would leave time for the later development of adapted and optimized metabolic networks. If ADP-glucose that is only used for storage polysaccharide synthesis by bacteria and that is not recognized or degraded at high rates by the host leaks out the symbiont, then the latter would become unable of synthesizing starch and therefore fueling nitrogen fixation. The host would therefore need to supply the symbiont with reduced nitrogen. This would not be a problem because the bacterium probably already contained suitable transporters to start with. In addition, the presence of the symbiont in a nitrogen-rich environment defined by the host cell cytoplasm would also have favored the loss of diazotrophy.

Lowering the ADP-glucose pools even to the point of their disappearance is known in present-day bacteria, cyanobacteria, or green algae to have either no or minimal impact on survival or growth (Preiss 1984; Zabawinski et al. 2001; Miao et al. 2003). Another advantage of the loss of diazotrophy would be to synchronize the future plastid divisions with the supply of nitrogen by the host. Increased storage polysaccharide synthesis defines a universal response to nitrogen limitation or starvation in microorganisms. By controlling and limiting this supply to the cyanobiont, the host had a powerful handle on cell cycle control of its endosymbiont. In addition, it further enslaved it to perform increased ADP-glucose synthesis to feed the host's own metabolic pathways through transient storage into cytoplasmic starch. We suggest that storage materials in general (lipids or carbohydrates) may be the ideal substrates for establishing symbiotic links between unrelated metabolic networks.

Supplementary Material

Supplementary figures 1–9 are available at *Molecular Biology and Evolution* online (<http://www.mbe.oxfordjournals.org/>).

Acknowledgments

The authors would like to thank D. Dupeyre (CER-MAV, Grenoble) for TEM imaging and B. Pontoire (INRA, Nantes) for X-ray diffraction analyses. The authors would like to thank the CNRS, the ANR, the INRA, and the French Ministry of Education for support. Immunolocalization of nitrogenase was conducted by Luisa Folcón in Birgitta Bergman's laboratory (Stockholm University). This work is dedicated to Cathy, queen of the plant kingdom for continuing encouragement and support.

Literature Cited

- Adl SM, Simpson AGB, Farmer MA, et al. (30 co-authors). 2005. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. *J Euk Microbiol.* 52:399–451.
- Aury JM, Jaillon O, Duret L, et al. (43 co-authors). 2006. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature.* 444:171–178.
- Ball SG, Guan HP, James M, Myers A, Keeling P, Mouille G, Buléon A, Colonna P, Preiss J. 1996. From glycogen to amylopectin: a model explaining the biogenesis of the plant starch granule. *Cell.* 86:349–352.
- Ball SG, van de Wal MHB, Visser RGF. 1998. Progress in understanding the biosynthesis of amylose. *Trends Plant Sci.* 3:462–467.
- Ball SG, Morell MK. 2003. Starch biosynthesis. *Annu Rev Plant Biol.* 54:207–233.
- Bodyl A, Makiewicz P, Stiller JW. 2007. The intracellular cyanobacterial of *Paulinella chromatophora*: endosymbionts or organelles? *Trends Microbiol.* 15:295–296.
- Buléon A, Colonna P, Planchot V, Ball S. 1998. Starch granules: structure and biosynthesis. *Int J Biol Macromol.* 23:85–112.
- Carlton JM, Hirt RP, Silva JC, et al. (69 co-authors). 2007. Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science.* 315:207–212.
- Cavalier-Smith T. 1998. A revised six-kingdom system of life. *Biol Rev Camb Philos Soc.* 73:203–266.
- Colleoni C, Dauvillée D, Mouille G, Morell M, Samuel M, Slomiany MC, Liénard L, Wattedled F, D'Hulst C, Ball S. 1999. Biochemical characterization of the *Chlamydomonas reinhardtii* alpha-1,4 glucanotransferase supports a direct function in amylopectin biosynthesis. *Plant Physiol.* 120:1005–1014.
- Coppin A, Varre JS, Lienard L, Dauvillée D, Guerardel Y, Soyer-Gobillard MO, Buleon A, Ball S, Tomavo S. 2005. Evolution of plant-like crystalline storage polysaccharide in the protozoan parasite *Toxoplasma gondii* argues for a red alga ancestry. *J Mol Evol.* 60:257–267.
- Coutinho PM, Henrissat B. 1999. Carbohydrate-active enzymes: an integrated database approach. In: Gilbert HJ, Davies G, Henrissat B, Svensson B, editors. "Recent advances in carbohydrate bioengineering". Cambridge: The Royal Society of Chemistry. p. 3–12.
- Dauvillée D, Kinderf IS, Li Z, Kosar-Hashemi B, Samuel MS, Rampling L, Ball S, Morell MK. 2005. Role of the *Escherichia coli* *glgX* gene in glycogen metabolism. *J Bacteriol.* 187:1465–1473.
- Derelle E, Ferraz C, Rombault S, et al. (26 co-authors). 2006. Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc Natl Acad Sci USA.* 103:11647–11652.
- Edner C, Li J, Albrecht T, et al. (11 co-authors). 2007. Glucan, water dikinase activity stimulates breakdown of starch granules by plastidial beta-amylases. *Plant Physiol.* 145:17–28.

- Eichinger L, Pachebat JA, Glöckner G, et al. (97 co-authors). 2005. The genome of the social amoeba *Dictyostelium discoideum*. *Nature*. 435:43–57.
- Eisen JA, Coyne RS, Wu M, et al. (53 co-authors). 2006. Macronuclear genome sequence of the ciliate *Tetrahymena thermophila*, a model eukaryote. *PLoS Biol*. 4:1620–1642.
- Falcón L, Lindwall I, Bauer K, Bergman B, Carpenter EJ. 2004. Ultrastructure of unicellular N-2 fixing cyanobacteria from the tropical North Atlantic and subtropical North Pacific Oceans. *J Phycol*. 40:1074–1078.
- Fox J, Robyt J. 1991. Miniaturization of three carbohydrate analyses using a microsample plate reader. *Anal Biochem*. 195:93–96.
- Gil R, Latorre A, Moya A. 2004. Bacterial endosymbionts of insects: insights from comparative genomics. *Environ Microbiol*. 6:1109–1122.
- Henrissat B, Deleury E, Coutinho PM. 2002. Glycogen metabolism loss: a common marker of parasitic behaviour in bacteria? *Trends Genet*. 18:437–440.
- Honda D, Yokota A, Sugiyama J. 1999. Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. *J Mol Evol*. 48:723–739.
- James MG, Robertson DS, Myers AM. 1995. Characterization of the maize gene *sugary 1*, a determinant of starch composition in kernels. *Plant Cell*. 7:417–429.
- Loftus B, Anderson I, Davies R, et al. (51 co-authors). 2005. The genome of the protist parasite *Entamoeba histolytica*. *Nature*. 433:865–868.
- Miao X, Wu Q, Wu G, Zhao N. 2003. Changes in photosynthesis and pigmentation in an agp deletion mutant of the cyanobacterium *Synechocystis* sp. *Biotechnol Lett*. 25:391–396.
- Mouille G, Maddelein ML, Libessart N, Talaga P, Decq A, Delrue B, Ball S. 1996. Preamylopectin processing: a mandatory step for starch biosynthesis in plants. *Plant Cell*. 8:1353–1366.
- Myers AM, Morell MK, James MG, Ball SG. 2000. Recent progress toward understanding biosynthesis of the amylopectin crystal. *Plant Physiol*. 122:989–998.
- Nakamura Y, Kubo A, Shimamune T, Matsuda T, Harada K, Satoh H. 1997. Correlation between activities of starch debranching enzymes and α -polyglucan structure in endosperms of *sugary-1* mutants of rice. *Plant J*. 12:143–153.
- Nakamura Y, Takahashi J, Sakurai A, et al. (13 co-authors). 2005. Some cyanobacteria synthesize semi-amylopectin type polyglucans instead of glycogen. *Plant Cell Physiol*. 46:539–545.
- Page RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci*. 12(4):357–358.
- Patron NJ, Keeling PK. 2005. Common evolutionary origin of starch biosynthetic enzymes in green and red algae. *J Phycol*. 41:1131–1141.
- Plancke C, Colleoni C, Deschamps P, et al. (12 co-authors). Forthcoming 2008. The pathway of cytosolic starch synthesis in the model glaucophytes *Cyanophora paradoxa*. *Eukaryot. Cell*.
- Preiss J. 1984. Bacterial glycogen synthesis and its regulation. *Annu Rev Microbiol*. 38:419–458.
- Ral JP, Derelle E, Ferraz C, et al. (13 co-authors). 2004. Starch division and partitioning a mechanism for granule propagation and maintenance in the picophytoplanktonic green alga *Ostreococcus tauri*. *Plant Physiol*. 136:3333–3340.
- Reddy KJ, Benjamin Haskell J, Sherman DM, Sherman LA. 1993. Unicellular, aerobic nitrogen-fixing cyanobacteria of the Genus *Cyanothece*. *J Bacteriol*. 175:1284–1292.
- Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol*. 111:1–61.
- Roach PJ. 2002. Glycogen and its metabolism. *Curr Mol Med*. 2:101–120.
- Rodríguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Löffelhardt W, Bohnert HJ, Philippe H, Lang BF. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr Biol*. 15:1325–1330.
- Scheidig A, Fröhlich A, Schulze S, Lloyd JR, Kossmann J. 2002. Downregulation of a chloroplast-targeted beta-amylase leads to a starch-excess phenotype in leaves. *Plant J*. 30:581–591.
- Schneegurt MA, Sherman DM, Nayar S, Sherman LA. 1994. Oscillating behavior of carbohydrate granule formation and dinitrogen fixation in the cyanobacterium *Cyanothece* sp. strain ATCC 51142. *J Bacteriol*. 176:1586–1597.
- Schneegurt MA, Sherman DM, Sherman LA. 1997. Composition of carbohydrate granules of the cyanobacterium, *Cyanothece* sp. Strain ATCC51142. *Arch Microbiol*. 167:89–98.
- Song J, Xu Q, Olsen R, Loomis WF, Shaulsky G, Kuspa A, Sucgang R. 2005. Comparing the *Dictyostelium* and *Entamoeba* genomes reveals an ancient split in the Conosa lineage. *PLoS Comput Biol*. 1:579–584.
- Suzuki E, Umeda K, Nihei S, Moriya K, Ohkawa H, Fujiwara S, Tsuzuki M, Nakamura Y. 2007. Role of GlgX protein in glycogen metabolism of the cyanobacterium, *Synechococcus elongatus* PCC 7942. *Biochim. Biophys Acta*. 1770:763–773.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 22:4673–4680.
- van de Wal M, D'Hulst C, Vincken JP, Buléon A, Visser R, Ball S. 1998. Amylose is synthesized in vitro by extension of and cleavage from amylopectin. *J Biol Chem*. 273:22232–22240.
- Wattebled F, Dong Y, Dumez S, et al. (11 co-authors). 2005. Mutants of *Arabidopsis* lacking a chloroplastic isoamylase accumulate phytylglycogen and an abnormal form of amylopectin. *Plant Physiol*. 138:184–195.
- Weber APM, Linka M, Bhattacharya D. 2006. Single, ancient origin of a plastid metabolite translocator family in Plantae from an endomembrane-derived ancestor. *Eukaryot Cell*. 5:609–612.
- Zabawinski C, Van Den Koornhuysen N, D'Hulst C, Schlichting R, Giersch C, Delrue B, Lacroix JM, Preiss J, Ball S. 2001. Starchless mutants of *Chlamydomonas reinhardtii* lack the small subunit of an heterotetrameric ADP-glucose pyrophosphorylase. *J Bacteriol*. 183:1069–1077.
- Zeeman SC, Smith SM, Smith AM. 2007. The diurnal metabolism of leaf starch. *Biochem J*. 401:13–28.
- Zeeman SC, Umemoto T, Lue W-L, Au-Yeung P, Martin C, Smith AM, Chen JA. 1998. A mutant of *Arabidopsis* lacking a chloroplastic isoamylase accumulates both starch and phytylglycogen. *Plant Cell*. 10:1699–1712.

William Martin, Associate Editor

Accepted December 13, 2007