Characterization of function of the GlgA2 glycogen/starch synthase in *Cyanobacterium* sp. Clg1 supports an ancient role in the synthesis of long chain glycogen and starch.

Derifa Kadouche[a2], Mathieu Ducatez[a2], Ugo Cenci[a], Catherine Tirtiaux[a], Eiji Suzuki[b], Yasunori Nakamura[b], Jean-Luc Putaux[c], Amandine Durand Terrasson[c], Sandra Diaz-Troya[d], Francisco Javier Florencio[d], Maria Cecilia Arias[a], Alexander Striebeck[e], Monica Palcic[f], Steven G. Ball[a,1] and Christophe Colleoni[a,1]

[a] Univ. Lille, CNRS, INRA, UMR 8576 - UGSF - Unité de Glycobiologie Structurale et Fonctionnelle, F-59000 Lille, France

[b] Department of Biological Production, Akita Prefectural University, 241-438, Shimoshinjo Nakano aza Kaidohata Nishi, Akita, 010-0195 Japan

[c] Centre de Recherches sur Les Macromolécules Végétales (CERMAV-CNRS), Université Grenoble Alpes, BP 53, F-38041 Grenoble Cedex 9, France

[d] Instituto de Bioquímica Vegetal y Fotosíntesis cicCartuja, CSIC-Universidad de Sevilla, Avda. Américo Vespucio, 49 41092 Sevilla, Spain

[e] Raw Materials Group, Carlsberg Laboratory, Gamle Carlsberg Vej 10, 1799 Copenhagen V, Denmark

[f] The Department of Biochemistry and Microbiology, University of Victoria, Victoria BC, Canada V8W 3P6

Running title: Evolution of starch metabolism.

[1] Co-corresponding authors. Christophe Colleoni and Steven G. Ball and, UMR8576 CNRS-USTL, Bâtiment C9, Université des Sciences et Technologies de Lille, Cité Scientifique, 59655 Villeneuve d'Ascq Cedex, France. Phone: 33 3 20436543, Fax : 33 3 20436555, e-mail : steven.ball@univ-lille1.fr ; christophe.colleoni@univ-lille1.fr

[2]These authors contributed equally to this work

Short title: starch synthesis in cyanobacteria

One sentence summary: In starch accumulating cyanobacteria, the GlgA2/SSIII/SSIV enzyme is mandatory to obtain polysaccharide aggregation into amylopectin.

The author(s) responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors is: Christophe Colleoni (christophe.colleoni@univ-lille1.fr).

At variance with the starch accumulating plants and most of the glycogen-accumulating cyanobacteria, Cyanobacterium sp. CLg1 synthesizes both glycogen and starch. We now report the selection of a starchless mutant of this cyanobacterium that retains wild-type amounts of glycogen. Unlike other mutants of this type found in plants and cyanobacteria, the mutant proved to be selectively defective for one of the two types of glycogen/starch synthase: GlgA2. This enzyme is phylogenetically related to the previously reported SSIII/SSIV starch synthase that is thought to be involved in starch granule-seeding in plants. This suggests that in addition to the selective polysaccharide debranching demonstrated to be responsible for starch rather than glycogen synthesis, the nature and properties of the elongation enzyme defines a novel determinant of starch versus glycogen accumulation. We show that the phylogenies of GlgA2 and of 16S rRNA display significant congruence. This suggests that this enzyme evolved together with cyanobacteria when they diversified over 2 billion years ago. The reasons that may have favored the ancient selection of such an activity are discussed.

INTRODUCTION

Soluble glycogen/starch synthases of the GT5 (CAZy family 5 glycosyl- transferases) family transfer glucose from a nucleotide sugar to the non-reducing end of a growing α-1,4-linked glucan. Among the very large family of prokaryotic GT5 enzymes, the soluble starch synthases III/IV (SSIII/IV) found in the green plant or alga plastid and in the glaucophyte cytosol are united into a highly supported monophyletic group together with glycogen/starch synthases found in all chlamydiales intracellular pathogens, in a restricted number of proteobacteria and a large number of cyanobacteria (Ball et al., 2013). However, erosion of phylogenetic signal did not enable a clear determination of the root position of this SSIII/IV/GlgA2 sub-family within the large size GT5 GS phylogenetic tree (Ball et al, 2013). SSIII/IV in green plants and algae are known to be essential for starch synthesis and play both a role in building the large-size chains within amylopectin and a function in starch particle seeding and (or) polysaccharide synthesis priming (for reviews see D'Hulst *et al.*, 2015; Nakamura, 2015). Little is known about the function of the corresponding enzymes in bacteria. Cyanobacteria represent one of the most ancient groups of prokaryotes and the founders of oxygenic photosynthesis (Summons et al., 1999, Crowe et al., 2013). Like in plants, photosynthetic carbon is temporarily assimilated via the Calvin cycle in the form of homopolymers of D-glucose, such as glycogen or starch that both consist of glucan chains

made of glucose residues linked in α-1,4 and branched by α-1,6 linkages. In spite of sharing the same chemical linkages, both polymers widely differ in physicochemical properties. Glycogen particles are highly branched polysaccharides (8 to 10 % of α-1,6 branches) resulting in the storage of small hydrosoluble particles with 30-50 nm maximal diameter in the cytosol of numerous organisms (Archaea, Bacteria and eukaryotes). One third of a maximum total 55000 glucose residues within a single particle is readily accessible to glycogen catabolism in the outer chains without cleaving off α-1,6 branches (Melendez-Hevia et al., 1993). Glycogen is thus a homogeneous structure and a very dynamic form of glucose storage that combines low osmotic activity and accessibility to hydrosoluble enzymes. Starch granules are usually made up of two α-glucan polymers, namely amylopectin and amylose. The minor fraction, amylose, is composed of linear weakly-branched glucan chains (<1% of α-1,6 branches) while the major fraction, amylopectin, harbors an ordered branch pattern of α-1,6 linkages leading to the cluster organization responsible for starch crystallinity (Hizukuri, 1986, Laohaphatanaleart et al., 2010, Bertoft et al., 2010). The synthesis of starch granules was initially believed to be a hallmark of three sister lineages -plants/green algae, red algae and Glaucophytes (i.e Archaeplastida) stemming from primary plastid endosymbiosis and some of their secondary endosymbiosis derivatives (i.e. Alveolates and Cryptophytes) (Cenci et al., 2014; Ball et al., 2015). Several lines of evidence suggest that starch metabolism has evolved shortly after plastid endosymbiosis from a pre-existing cytosolic eukaryotic glycogen metabolism enzyme network. In line with this hypothesis, an overview of gene origin in Archaeplastida lineages points out that most of the starch metabolism enzymes display a common host phylogeny. Only 4 genes of the inferred ancestral Archaeplastida network display a clear-cut bacterial origin with two originating from cyanobacteria (GBSS (granule-bound starch synthase), and ADP-glucose pyrophosphorylase (GlgC)) and the remainder two from chlamydial intracellular pathogens (GlgA and GlgX respectively soluble glycogen/starch synthase and glycogen/starch debranching enzyme) (Ball et al., 2013). Interestingly, extant unicellular diazotrophic cyanobacteria were recently reported to synthesize starch-like polysaccharides with an enzyme network mostly unrelated to the one at work in Archaeplastida (Cenci et al., 2013). The presence of GBSS in chroococcales unicellular diazotrophic cyanobacteria may suggest that the plastid ancestor could have been an ancient starch accumulator related to such organisms (Cenci et al., 2013). Indeed GBSS is an enzyme responsible for amylose synthesis within starch and requires the binding to semi-

crystalline polysaccharides to be active. We thus proposed that an ancestor of this group of diazotrophic unicellular cyanobacteria may define the plastid donor (Deschamps et al., 2008).

Because of the fastidious growth of many chroococcales unicellular diazotrophic cyanobacteria and because this group, like many other cyanobacteria, has resisted all attempts at genetic transformation, we applied a classical genetic approach to the dissection of starch metabolism in *Cyanobacterium* sp. CLg1. This strain initially reported as diazotrophic by Falcon et al. (2004) was axenized by us, but has lost the ability to fix nitrogen under laboratory conditions. *Cyanobacterium* sp. CLg1, has been reported to accumulate both a major starch fraction and a minor yet significant glycogen pool (Falcon et al., 2004; Cenci et al., 2013). We now report the selection of a starchless mutant of Cyanobacterium that synthesizes wild-type amounts of glycogen. This mutant proved to be selectively defective for the GlgA2 glycogen/starch synthase. This suggests that starch and glycogen are synthesized by at least partly distinct pathways in *Cyanobacterium* sp. CLg1. To our knowledge this is the first report for a requirement other than those previously assigned to starch debranching enzymes for the selective accumulation of starch rather than glycogen in living cells. The evolutionary implications of this novel function are discussed in the light of the origin and possible role of the SSIII-IV-GlgA family of glucan elongation enzymes within cyanobacteria.

Results

Selection of 187G11 a starchless mutant of *Cyanobacterium sp.* CLg1

A collection of 2.10^4 mutants was generated after UV mutagenesis followed by a minimum of 4 rounds of subcloning as previously detailed in Cenci et al. (2013). Following the first round of screening, the selected mutants were further subcloned to check for complete segregation of the mutant phenotype. After 3 years of segregation and phenotype screening we selected 7 strains defining the class C mutants, which contained water-soluble polysaccharides in amounts close to those of the wild-type reference but with significantly lower amounts of starch. 6 of these 7 mutants were previously reported in Cenci et al. (2013) but failed to reveal the biochemical explanation for the mutant phenotype. The seventh strain (187G11) displayed a very severe phenotype defined by the absence of iodine stain displayed after spraying cell patches with iodine vapors. This was correlated to a complete disappearance of starch, which fell below detection level (<0.5% of wild-type level) (Fig. 1). However the mutant remained able to accumulate normal amount of water soluble polysaccharide (0.34 ± 0.04 mg of WSP.mg^{-1} of protein) in comparison to the wild-type strain (0.26 ± 0.04 mg of WSP.mg^{-1} of protein). This phenotype is more severe than that exhibited by class A mutants which overproduced glycogen and retained very low levels (2-5% of wild type) of starch with modified structure (Cenci et al., 2013). Nevertheless unlike glycogen-less mutants of Synechocystis PCC6803 the mutant grew under 12h light/12 h darkness growth conditions albeit with a twofold increase in generation time (from 60 to 120 h) (Grundel et al., 2012) (Supplemental Fig. S1) in liquid medium. On solid media we did not observe significant delays in the appearance of single colonies.

187G11 displays normal glycogen levels of slightly modified structure.

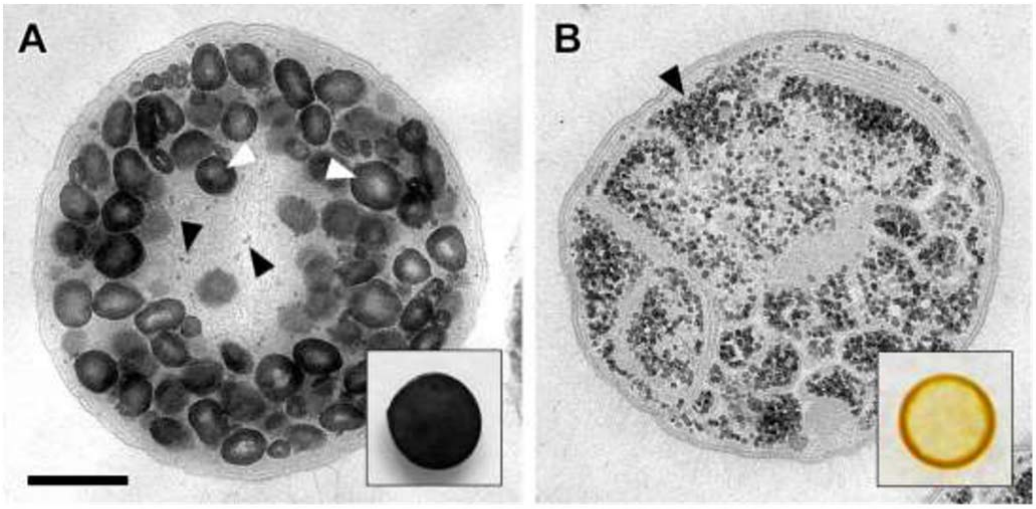

Figure 1. TEM images of ultrathin sections (70 nm) of wild type and 187G11 mutant strains. Polysaccharides in the wild-type CLg1 (A) and 187G11 mutant (B) are positively stained with PATAg. Both starch-like granules (white arrows) and glycogen particles (dots pointed by black arrows) were observed in the wild-type strain. Starch granules are absent in the 187G11 mutant and substituted by glycogen-like water-soluble polysaccharides (black arrow). The dark-blue iodine stain from a cell patch of the wild-type strain is shown in inset. The absence of starch granules in the 187G11 mutant yields a yellow-orange stain after spraying iodine vapors (inset in panel B). Bar: 500 nm

To characterize the water-soluble polysaccharide fraction accumulated in the 187G11 mutant, the latter was purified, sized by gel permeation chromatography and compared to wild type (Fig. 2A and 2B). Both mutant and wild type soluble polysaccharides are composed of high molecular weight polysaccharides (fractions 35 to 50) and short malto oligosaccharides (fractions 60 to 100). The formers were then examined by transmission electron microscopy (TEM) and further subjected to enzymatic debranching, followed by separation of chains by HPAEC-PAD (High Performance Anion Exchange Chromatography with Pulsed Amperometric Detection). Chain length distribution analysis (Fig. 2C and 2D) and TEM observation of negatively stained preparations (Fig. 2E and 2F) suggest that WSP of 187G11 is composed of highly branched glucan chains capable to exclude the uranyl acetate molecules in a fashion similar to wild-type soluble polysaccharides (Fig. 2E). Altogether, these results (Fig.1 and Fig. 2) suggest that 187G11 contains normal amounts of wild-type glycogen with a similar branched polysaccharide and with a chain-length distribution slightly enriched in small chains. Hence 187G11 synthesizes glycogen as efficiently as wild-type cells but selectively lacks starch.

187G11 is specifically defective for the major starch /glycogen synthase

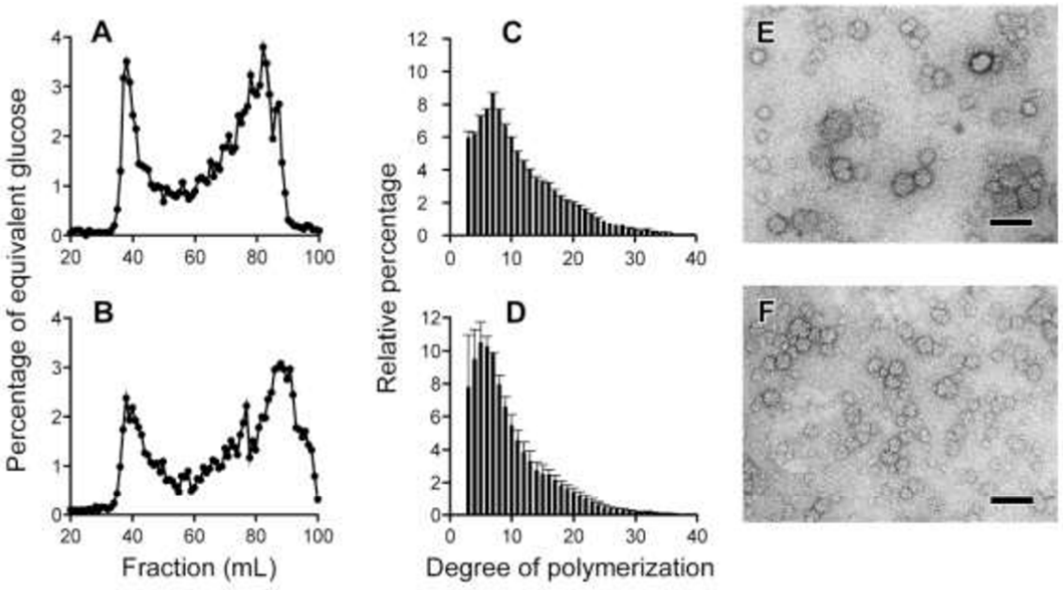

Figure 2. Structural analysis of water-soluble polysaccharide accumulated by the wild type and 187G11 mutant strains. Water-soluble polysaccharides purified from wild-type (A) and 187G11 mutant (B) strains were subjected to size exclusion chromatography analysis (TSKHW55 Toyopearl). The amount of total glucose was determined for each fraction by the phenol-sulfuric method (see methods). Results are expressed as weight percentages of equivalent glucose (black line). After complete digestion with commercial isoamylase, glucan chains were separated according to their degree of polymerization (DP) by HPAEC-PAD. The relative abundance for each DP (black bars) was determined for the wild type (C) and 187G11 mutant (D) from the mean of three independent extractions. TEM images of negatively stained preparations suggest that WSP of the wild type (E) and 187G11 mutant (F) are highly branched polysaccharides with a diameter below 50 nm similar to glycogen particles of rabbit liver (bars: 100 nm)

We undertook a large survey of starch metabolism enzymes through crude extract assays (ADP-glucose pyrophosphorylase and glycogen/starch synthase) and previously adapted zymogram procedures (phosphorylases, glycosyl hydrolases, and transferases including BE (branching enzyme), α-1,4 glucanotransferase, debranching enzymes, amylases, and glycogen/starch synthases) (Supplemental Fig. S2). We found a very large decrease in total glycogen primed glycogen/starch synthase activity (80% decrease with respect to wild-type (492 $nmol.min^{-1}.mg^{-1}$ of protein)) that correlated with the disappearance of the major glycogen/starch synthase. A second minor slow migrating glycogen/starch synthase was also witnessed selectively in the mutant 187G11 strain (Fig. 3). Unfortunately the unstable nature and low activity of this slow migrating band which we suspect to represent the GlgA2 mutant enzyme (see below) only allowed us to partially purify it from *Cyanobacterium* sp. CLg1 mutant crude extracts. We have previously published that mutants defective for a debranching enzyme (GlgX2) over-accumulated glycogen and witnessed a dramatic decrease in starch

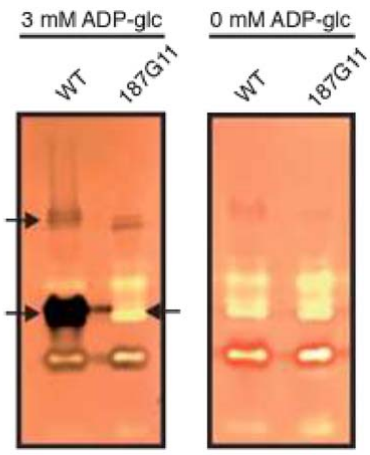

Figure 3. Zymogram analysis of glycogen/starch synthase activities from wild type and 187G11 mutants. Total protein of semi-purified crude extracts of both WT and 187G11 mutant strains were separated by native PAGE, then electro-transferred onto native PAGE containing 0.6% (p/v) of glycogen. The native gels were then incubated with or without 3 mM ADP-glucose. Glycogen/starch synthase activities are witnessed after iodine staining in the wild type's crude extract as two dark activity bands (black arrows). The fast migrating form disappears in the 187G11 mutant (black arrow). The total decrease in iodine staining was estimated through dilution to be between two to three orders of magnitude. This decrease is in line with that measured by quantitative radioactive assays in recombinant *E. coli* extracts.

amounts. To make sure that the phenotype displayed in 187G11 could not result from a combination of a direct effect on the glycogen/starch synthase and an indirect effect on the GlgX2 debranching enzyme, we semi-quantified GlgX2 by zymogram analysis through the procedures detailed in Cenci et al. (2013) and found the activity to be normal qualitatively and quantitatively (Supplemental Fig. S3).

Characterization of a glycogen/starch synthase mutation in the 187G11 genome

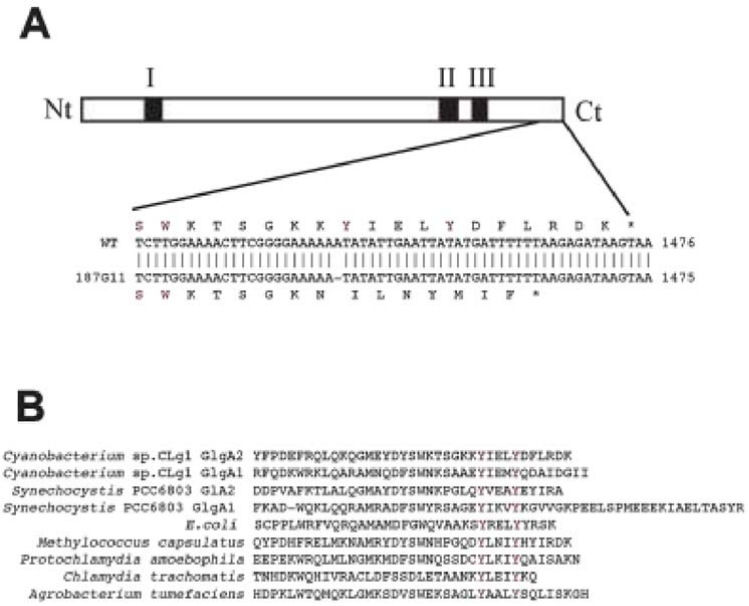

Figure 4. Molecular characterization of the starchless mutant 187G11. (A) One base pair deletion was identified in the glgA2 gene of 187G11. This point mutation results in a frameshift followed by appearance of a non-sense codon and the synthesis of truncated protein (GlgA2*) at the Carboxy terminus. Regions I, II and III (black boxes) previously characterized to be involved in the binding of ADP-glucose and in catalysis are conserved in GlgA2*. (B) Nevertheless, the YxxxY motif conserved throughout all bacterial GT5 glycosyl transferases has disappeared in GlgA2*.

We proceeded to sequence all of the genes previously found in the Cyanobacterium CLg1 genome related to either glycogen and (or) starch metabolism. These included all possible glycosyl hydrolases and glycosyl transferases found in the genome and known or suspected to be involved in glucan metabolism as listed in Table 1. In addition, we sequenced the unique gene encoding ADP-glucose pyrophosphorylase from the 187G11 strain. We found only one significant modification in the whole starch/glycogen metabolism network and no silent mutations. We thus found a one base pair deletion yielding a frameshift and a nonsense mutation toward the C-terminus of the GlgA2 glycogen synthase gene (Fig. 4). This mutation deletes a highly conserved region of the bacterial GT5 glycogen synthases and is therefore expected to impact the enzyme activity. We further investigated this impact on other GT5 glycogen synthases such as the *E. coli* enzyme by introducing mutations in one or both of the highly conserved tyrosine residues (Supplemental Fig. S4) of this region which resulted in a large decrease of the enzyme activity. Slow growth of the marine *Cyanobacterium* sp. CLg1 strain requires 4 years for a full cycle of mutant screening and purification which prevented us from selecting additional defective alleles. In addition chroococcales cyanobacteria are notorious for their resistance to genetic transformation precluding complementation of the

196 effect by genetic transformation. Hence not only did we sequence all genes of the
197 starch/glycogen metabolism network but in addition we assayed all possible enzymes of the
198 network to check that undetected mutations in regulatory genes would not modify the balance
199 of starch/glycogen metabolism enzymes. We found no evidence for any qualitative
200 modification in all assayable enzyme activities through crude extract assays and zymogram
201 procedures. However we did record a significant increase in starch (glycogen) phosphorylase
202 activity which was also noted in other mutants of *Cyanobacterium* sp. CLg1. Similar
203 increases have been noted in other cyanobacterial mutants by others (Fu and Xu, 2006; Cenci
204 et al., 2013). These increases may explain the slightly modified glycogen structures observed
205 in the mutants.
206
207 **Biochemical characterization of glycogen/starch synthases in wild-type and mutant**
208 **cyanobacteria**
209 In order to verify that the nonsense mutation detected in *glgA2* explains both the
210 disappearance of the major glycogen/starch synthase and the phenotype recorded in 187G11
211 we expressed wild-type GlgA1 and GlgA2 proteins as well as the mutant GlgA2* enzyme in
212 *E. coli* (Fig. 5A). In addition we checked for complementation of the *E. coli glgA* mutation by
213 our constructs. Interestingly the wild-type GlgA2 enzyme complemented the *E. coli* defect
214 only when *E. coli* was supplemented with maltose and not with mannitol, a property which
215 was shared by both GlgA1 and the mutant GlgA2* (Fig. 5B). All recombinant proteins cross-
216 reacted in purified extracts as expected with antibodies directed against the phylogenetically
217 related Synechocystis PCC 6803 GlgA1 and GlgA2 (named respectively GSII and GSI by
218 Yoo et al. 2014) (Supplemental Fig. S5). However we were unable to distinguish the activities
219 in crude extracts because of abundant cross reactions against other bacterial proteins. A
220 strongly decreased activity was scored for GlgA2* (0.303 $\mu mol.min^{-1}.mg^{-1}$) by comparison to
221 the wild-type enzyme (327 $\mu mol.min^{-1}.mg^{-1}$) while significant GlgA1 activities could be
222 reproducibly measured only by quantitative radioactive assays (4.1 $\mu mol.min^{-1}.mg^{-1}$) (see
223 methods). The absence of iodine stain in the strains expressing GlgA1 is suggestive of the
224 synthesis of very short glucan chains (iodine staining of glucans starts developing at 20°C for
225 chains longer than 12 Glc residues). We also expressed GlgA1 and GlgA2 from
226 Synechocystis PCC 6803. In a similar fashion, we found recombinant activity through iodine
227 staining with GlgA2 but not with GlgA1, which thus behaved like the *Cyanobacterium* sp.
228 CLg1 GlgA1.

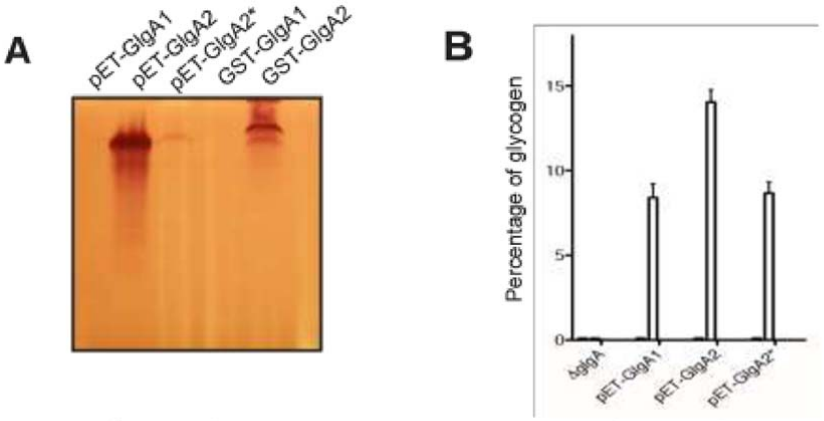

Figure 5: Complementation experiments and recombinant protein expressions of GlgA1 and GlgA2 of CLg1. (A) Recombinant protein expressions of GST tagged glycogen/starch synthase GlgA1and GlgA2 (pGEX-glgA1, pGEX-glgA2) and untagged proteins (pET-glgA1, pET-glgA2 and pET-glgA2*) were expressed in the ΔglgA mutant stain (JW3392-1) of *Escherichia coli*. Crude extracts were loaded on native PAGE containing glycogen. After migration, the gel was incubated overnight in the incubation buffer containing 3 mM ADP-glucose. Provided their glucan products are sufficiently long to form helices that stably trap iodine, glycogen/starch synthase activities may be revealed as black bands staining in an orange background after soaking the native gel in iodine solution. (B) Restoration of glycogen synthesis in the presence of mannitol or maltose as carbon source in the ΔglgA mutant expressing untagged protein GlgA1, GlgA2 and GlgA2* (pET-glgA1, pET-glgA2 and pET-glgA2*). On the histogram the mannitol response is displayed (left bar) on the left besides the maltose response (right bar). This response to mannitol was below detection level afforded by our sensitive amyloglucosidase glycogen amount assays The glycogen measured in the wild-type strain was used as reference (70 ± 20 µg of glycogen/mg of protein). The results were expressed as percentages of glycogen amounts accumulated by our wild-type reference.

GlgA2 is selectively bound to cyanobacterial starch and can prime polysaccharide synthesis in *Cyanobacterium* but not *E. coli.*

In a previous study, we found peptides from a 52 kDa starch bound protein that matched those from the *Cyanobacterium* sp. CLg1 GlgA2 (Fig. 6A) (Deschamps et al., 2008). We also found GlgA2 by similar means on the purified Cyanothece starch-like granules. We used the anti-GlgA1 and anti-GlgA2 antibodies raised against the Synechocystis enzymes that cross-reacted similarly with the corresponding GlgA1 and GlgA2 Cyanobacterium recombinant proteins to check for the presence of these proteins on the cyanobacterial starch granules. In both systems we found GlgA2 as a major starch bound protein with no GlgA1 detected (Fig. 6B).

The fact that the wild-type GlgA2 enzyme could complement a *glgA* defective *E. coli* mutant only in the presence of maltose suggested to us that this activity was dependent on the supply of MOS primers by the MalQ amylomaltase in *E coli*. This was confirmed by the absence of recombinant GlgA2 enzyme activity recorded on zymogram gels in the absence of glycogen

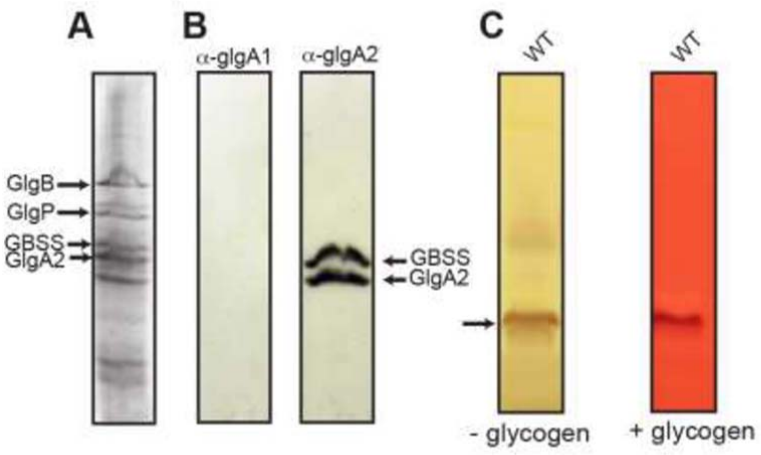

Figure 6. Western blot analysis of granule bound proteins and primer dependence of GlgA2 activity. (A) Proteins specifically bound to starch granules were analyzed onto SDS-PAGE. Major polypeptides were previously identified by mass spectrometry analysis (Deschamps et al., 2008): GlgB (85 kD branching enzyme); GlgP (72 kD glycogen phosphorylase); GBSS (57 kD granule bound starch synthase); GlgA2 (52 kD glycogen/starch synthase). (B) Western blot analysis was carried out on proteins attached to starch granules. Polypeptides were transferred onto PVDF membranes. Glycogen/starch synthases were immunodetected using antibodies raised against GlgA1 (α-GlgA1) and GlgA2 (α-GlgA2) of Synechocystis PCC6803. (C) Total proteins of the wild-type CLg1 (WT) strain were loaded and separated onto native-PAGE. After migration, the native PAGE gel was cut in two parts: one half of the gel was directly incubated in starch synthase buffer containing 3 mM of ADP-glucose (- glycogen) while the second half, was separately electro-transferred against another native PAGE containing glycogen (+ glycogen) during 2 hours. Starch synthase activity was revealed as black activity band with iodine solution after overnight of incubation (black arrow). Activities were too low to enable detection of GlgA2* in comparable experiments.

primer (Supplemental Fig. S6). This property was shared also by the GlgA2* mutant activity. However when GlgA2 was purified partially from Cyanobacterium extracts the wild-type protein was always able to prime glucan synthesis in the absence of added polysaccharide primer. We conclude that in Cyanobacterium the GlgA2 glycogen synthase is either modified or interacts with a Cyanobacterium specific factor or substrate absent from both glucose or maltose grown *E. coli* cells. We do not know if GlgA2* would behave similarly in Cyanobacterium extracts since we never obtained enough residual activity in the mutant to assay its primer dependence. We therefore believe that the mutant phenotype could be explained either by the spectacular decrease of enzyme activity on its own or by a combination of the latter and a possible inability to prime polysaccharide synthesis in vivo in *Cyanobacterium*.

Phylogenetic analysis of the SSIII-IV-GlgA2 glycogen/starch synthases

256 Several detailed and recent phylogenies of GT5 ADP-Glc requiring starch synthases have
257 appeared (Ball et al., 2013). However the databases have considerably increased in size since
258 our last analysis and now include a much larger diversity of cyanobacteria. For our
259 phylogenetic analysis we have selected only the SSIII-SSIV-GlgA2 monophyletic subgroup
260 of enzymes that was previously defined in these phylogenies with very high support and have
261 restricted our alignment to these sequences. The GT5 ADP-Glc dependent glycosyl-
262 transferases represent a distinctively prokaryotic group of enzymes with no representatives
263 within eukaryotes with the noticeable exception of Archaeplastida. It is thus reasonable to
264 assume that the green algae and plant sequences summarized in figure 7 were gained by LGT
265 from a prokaryotic source. Because the tree is unrooted and because the phylogeny of GT5
266 glycogen/starch synthases shows many signs of signal erosion we cannot exclude the unlikely
267 possibility that the sequence was donated to Chlamydiales by the Archaeplastida rather than

the reverse. Nevertheless the phylogeny represented in figure 7 and detailed in the supplemental Figure S7 demonstrates that despite the growing databases, the Chlamydiales remain the most plausible donors for the ancestral plant SSIII-IV starch synthase gene and that Cyanobacteria can be very confidently rejected as possible donors for the plant enzymes. The phylogeny now suggests in addition that GlgA2 may define a very ancient cyanobacterial enzyme as a substantial portion of the cyanobacterial diversity appears uninterrupted by foreign clades in a large size monophyletic group. Among the available cyanobacterial genome sequences, we estimate that 47 % of reported genomes lack GlgA2 while only 16 % lack GlgA1. Hence both sequences are largely distributed within cyanobacteria. Both GlgA1 and GlgA2 (2 cases) can function as sole glycogen/starch synthase. In addition both GlgA1 and GlgA2 absence is not confined to specific subgroups of cyanobacteria but is distributed throughout the cyanobacterial tree pointing to multiple gene loss events. A systematic search

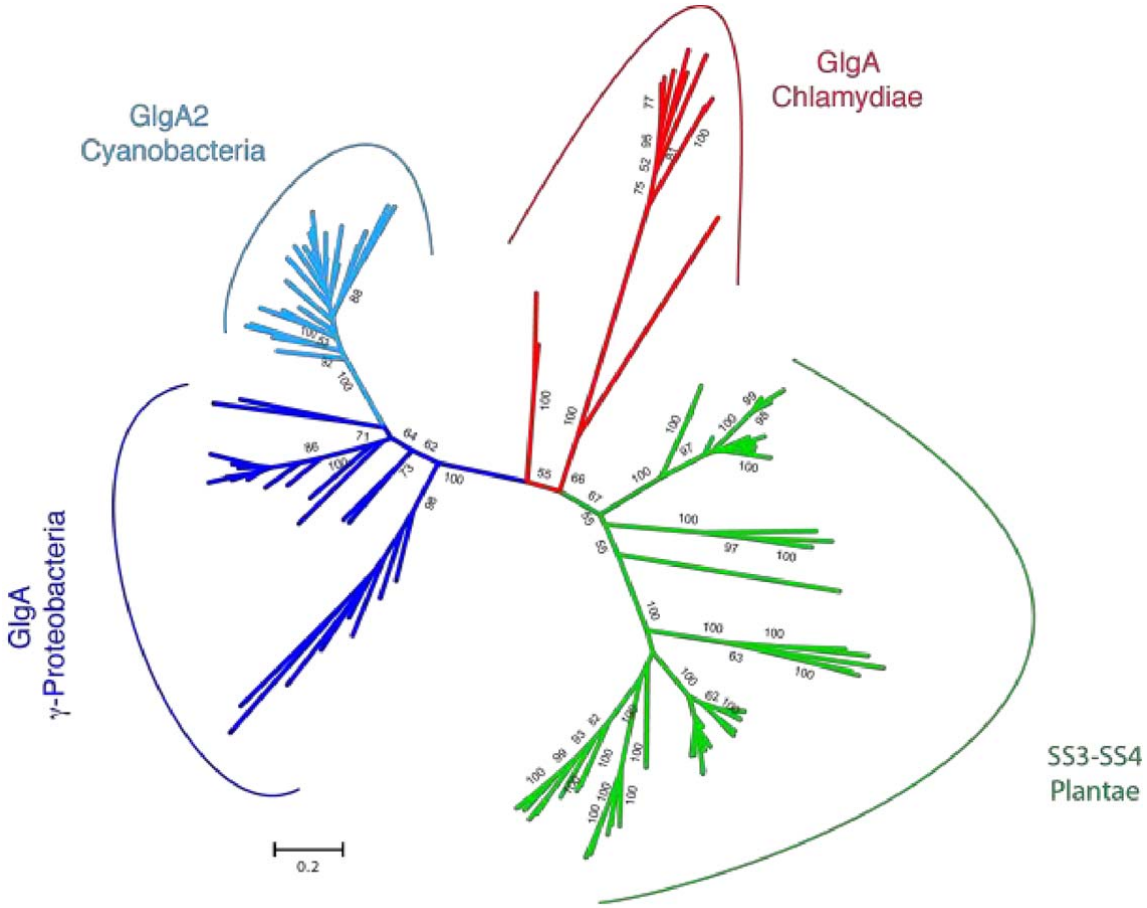

Figure 7. Unrooted phylogenetic tree of SSIII-IV-GlgA2 glycogen/starch synthases. Phylogenetic groups are color-coded according to their taxonomy: green correspond to green linage, red corresponds to Chlamydiales, blue corresponds to Proteobacteria and turquoise corresponds to Cyanobacteria. Posterior probability (PP) support values are indicatedat nodes supported by bootstrap values higher than 50 (1000 bootstrap replicates)

for GlgA2 in cyanobacteria has indeed yielded the finding of this enzyme in several of the most basal cyanobacterial clades (Colleoni and Suzuki, 2012) We further assayed the congruence of the GlgA2 phylogeny with the diversification of cyanobacteria (Supplemental Figures S8 and S9) as estimated through the 16S rRNA phylogeny and found a good level of congruence despite the intensive gene sharing and exchanges usually found in bacteria. A GlgA2-related sequence is also congruent with the diversification of Chlamydiales which contain no other glycogen/starch synthases (Ball *et al*., 2013). On the other hand the rather restricted distribution within mostly marine gamma proteobacteria (purple sulfur bacteria (chromatiaceae) and methanotrophs) may argue for a more recent LGTs within these groups.

291

Discussion

A novel determinant of starch versus glycogen structure; glucan product length and particle seeding

In this work, we bring suggestive evidence that cyanobacterial starch depends on a function provided by GlgA2 that cannot be supplied by GlgA1. We believe that this function is distinct from polysaccharide synthesis priming since normal glycogen synthesis priming occurs in the 187G11 mutant. In plants, mutants defective for both SSIII and SSIV which are phylogenetically-related to GlgA2, are also starchless but do not produce any glycogen (Szydlowski et al., 2009). In that case, however, the missing function is thought to consist of polysaccharide synthesis priming. Indeed, transgenic expression of the self-priming Agrobacterium GlgA glycogen synthase in the Arabidopsis SSIII-IV double mutants restores starch synthesis (Crumpton-Taylor et al., 2013). Expression of both GlgA1 and GlgA2 was successfully achieved in *E. coli*. However successful complementation of the *E. coli glgA* mutation could only be achieved in the presence of maltose and no complementation was observed on mannitol or glucose-grown *E. coli* cultures. Maltose is known to induce the maltose operon *per se* the MalQ amylomaltase which elongates malto-oligosaccharides by a series of transfer reactions at the expense of glucose formation. The synthesis of long glucans leading to glycogen production by action of glycogen branching enzyme is prevented by the presence of the MalP and MalZ gene products yielding glucose and glucose-1-P thereby feeding bacterial metabolism and recessing the long chains to maltotetraose yielding a MOS pool consisting of small-size glucans. Hence *E. coli* cells growing on maltose contain a significant pool of small-size MOS. Selective complementation of *glgA* in maltose grown *E.coli* cells suggest that *in vivo,* these bacteria are using the MOS pool to elongate glucans for glycogen synthesis. In the presence of mannitol or glucose the absence of a sizeable MOS pool prevents glycogen synthesis priming. This suggests that neither recombinant GlgA2 nor recombinant GlgA1 or GlgA2* are able to prime polysaccharide synthesis in the absence of MOS. This was confirmed for GlgA2 through zymogram analysis. However when GlgA2 was purified from *Cyanobacterium* sp Clg1, GlgA2 was systematically able to prime polysaccharide in the same zymogram analysis. We therefore conclude that the ability to prime does not define an intrinsic property of the cyanobacterial glycogen synthases and that this ability is dependent on either enzyme modification or supply of specific primers by other cyanobacterial factors. The GlgA2* mutant activity may or may not lack this essential property but nevertheless the very substantial decrease of its specific activity (at minima two

orders of magnitude (Figure 3)) precludes its normal function for cyanobacterial starch synthesis.

Our work suggests that *Cyanobacterium* sp. CLg1 displays two separate pathways for polysaccharide synthesis affording the possibility to regulate both of these pathways at least partly independently. The first pathway leads to the production of short-chain glycogen and the second yields starch. The starch specific pathway consists at minima of GlgA2 (this work) and of GlgX2 (Cenci et al., 2013). As we previously proposed, glycogen, because of the instant accessibility of the glucose stores of its outer chains, defines an optimal structure to ensure a fast adaptation of the carbon sink strength to optimize photosynthetic activity. On the other hand starch offers the opportunity to trap five-fold more (Cenci et al., 2013) carbon into a slow turnover storage polysaccharide form to ensure high respiration rates during the dark phase. These high respiration rates have been proposed by others to be needed in diazotrophic single cell cyanobacteria not only to supply nitrogenase with the required high levels of ATP and reducing power but also to further lower locally the oxygen levels through its respiratory consumption (Schneegurt et al., 1994). By reaching anoxia this would induce Nitrogenase synthesis and activity. We therefore predict that the absence of starch would abolish diazotrophy. Unfortunately the loss of diazotrophy of our axenic Cyanobacterium strain does not allow us to test this in a straightforward fashion. However in this respect, we wish to stress that the most abundant class of mutants of the green alga *Chlamydomonas reinhardtii* that are defective for hydrogen production under anoxic conditions by the oxygen-sensitive hydrogenase are those that we reported to substitute starch by glycogen synthesis (Posewitz et al., 2004).

All 6 cyanobacteria that have been proven to accumulate starch contain GlgA2. However many glycogen accumulating cyanobacteria also contain both GlgA1 and GlgA2 (Colleoni & Suzuki, 2012). We would like to propose that GlgA2 has evolved mainly to allow nitrogen fixation through the synthesis and mobilization of starch. Frequent loss of both diazotrophy and starch in cyanobacteria may not necessarily have been accompanied by that of GlgA2. In some cases GlgA2 may indeed have been lost as in Prochlorococcus and many related Synechococcus strains (Colleoni & Suzuki, 2012). In other cases, the single loss of the GlgX2 debranching enzyme would have converted the synthesis of high levels of starch into smaller levels of "phytoglycogen" (a polymer resembling glycogen but with slightly longer chains that results from impaired amylopectin crystallization) as evidenced in the GlgX2 mutants (Cenci et al., 2013). The pool of phytoglycogen induced by the loss of GlgX2

function would lead to the production of increased glycogen amounts made of slightly longer chains that escapes the hypothesized tight regulation of GlgA1 by photosynthesis. Hence maintenance of both long-chain and short-chain glycogen may have been desirable in some glycogen accumulating cyanobacteria such as *Synechocystis* sp. PCC6803.

Our work emphasizes that the intrinsic properties of the glycogen/starch synthase possibly defines a novel determinant of starch versus glycogen synthesis. We believe that the specific properties concerned consist in the synthesis of chains with a length (DP > 12) compatible for their selective debranching by GlgX2. Indeed GlgX2 was demonstrated to display little or no activity toward glycogen chains and to require the longer chains present in amylopectin-like molecules (Cenci et al., 2013). Our results concerning the biochemical properties of GlgA2 are in agreement with these speculations. The absence of iodine stained polysaccharide product in zymogram analysis of recombinant GlgA1 despite the presence of significant activity measured in our radioactive ADP-glucose incorporation assays strongly suggests the presence of a distributive mode of action for GlgA1. Indeed an hypothetical distributive mode of action had been solely deduced previously by others from the detailed glycogen structures produced in the single GlgA1 and GlgA2 mutants of Synechocystis (Yoo et al., 2014). That GlgA1 is responsible for short chain glycogen synthesis is indeed suggested by the measure of significant (20%) residual crude extract glycogen synthase activity found in the 187G11 mutant which is also in agreement with its function in the synthesis of the remaining short-chain glycogen pool. The small increase of short chains in the glycogen structure of the GlgA2 mutant of Cyanobacterium when compared to the wild-type can be explained either by the mutation of GlgA2 contributes in a minor fashion to glycogen synthesis or by the observed induction of phosphorylase activity in the mutant. We presently cannot distinguish between these two possibilities.

In addition to long chain synthesis, we believe that an additional function for starch versus glycogen synthesis carried by GlgA2 may consist of starch granule seeding. Indeed, a need exists at the core of the granule for the organization of the 3D crystalline growth of the individual granules. There is also a need to control starch granule size and hence starch granule seeding independently from glycogen to avoid physically blocking bacterial division.

The cyanobacterial origin of the SSIII-IV-GlgA2 subfamily of GT5 glycogen/starch synthases.

391 The phylogeny published in this and our previous work shows that two prokaryotic groups
392 show a significant level of congruence between their diversification and the GlgA2
393 phylogeny. These are the Chlamydiales and the cyanobacteria. The few bacteria mainly basal
394 gamma proteobacteria members presenting a GlgA2-SSIII-IV group can be easily explained
395 through lateral gene transfers (LGTs) from cyanobacteria in a common marine environment.
396 This work extends the distribution of cyanobacteria to the point where we can show that the
397 phylogeny of GlgA2 shows an appreciable level of congruence with this group of bacteria
398 especially when considering the high level of gene exchanges considered to occur in bacteria.
399 However some cyanobacteria lack either GlgA1, GlgA2 or both through selective gene losses,
400 but nevertheless, both enzymes are largely distributed within this group. Since diversification
401 of cyanobacteria was initiated between 2 to 3 billion years (Sanchez-Baracaldo et al., 2014),
402 the node uniting all cyanobacteria in Figure 7 is vastly more ancient than that uniting the
403 green algae and land-plants. This conclusion invalidates that the root of the GlgA2-SSIII-IV
404 group could lie within the Archaeplastida since the latter diversified after plastid
405 endosymbiosis (dated between 0.9 to 1.6 billion years). A GlgA2-like gene is on the other
406 hand universally distributed in all Chlamydiales where it defines the sole starch/glycogen
407 synthase present. Chlamydiales are considered to be members of the bacterial PVC clade
408 (consting of Planctomycetes, Verrucomicrobia, Chlamydiales). However no other glycogen
409 accumulating PVC members are reported to contain enzymes of similar phylogenetic origin
410 suggesting that the last common Chlamydiales ancestor may have received the gene by LGT
411 from other bacteria. Hence the GlgA2 type of enzyme displays a very ancient origin in both
412 Chlamydiales and cyanobacteria. GT5 (Glycosyl Transferase Cazy family 5) glycogen/starch
413 synthases that use ADP-Glc as substrate are very largely distributed in Bacteria and Archea.
414 This type of enzyme is only very distantly related to the UDP-Glc requiring GT5 or GT3
415 glycogen synthases distributed in glycogen storing eukaryotes. It is thus reasonable to
416 conclude that the green algae land plants and glaucophyte SSIII-SSIV glycogen/starch
417 synthases must have received the ancestor of SSIII-IV from a bacterial source, the most likely
418 being an ancient Chlamydiale. The unrooted tree presented in this work does not clarify the
419 origin of the SSIII-SSIV-GlgA subgroup of glycogen/starch synthases and a chlamydial
420 proteobacterial or cyanobacterial origin remain possible. Nevertheless if we now exclude the
421 Archaeplastida as the source for this enzyme, all these hypotheses agree with the presence of
422 a chlamydial LGT to the Archaeplastida. We believe that among the 3 possible origins a
423 cyanobacterial source defines the most probable scenario. Indeed while both Chlamydiales

and cyanobacteria display some level of congruence between their diversification and the phylogeny of their SSIII-IV-like glycogen synthase, the specialized function of GlgA2 in cyanobacteria evidenced in this work points to a possible link between the latter and diazotrophy in single cell cyanobacteria. This suggests a cyanobacterial origin since the conflict between oxygenic photosynthesis and diazotrophy probably predates the evolution of Chlamydiales from a PVC ancestor. This is further suggested by a suspected more ancient diversification of cyanobacteria and the presence of GlgA2 in some of the most basal clades. Also supporting such an origin is the unusual abundance of glycosyl hydrolases and glycosyltransferases related to storage polysaccharide metabolism in several cyanobacterial lineages including *Cyanobacterium* sp CLg1 which often display over twofold more enzymes than those found in most other bacteria and archaea (Colleoni & Suzuki, 2012). This higher complexity may reflect a very ancient subfunctionalization of storage polysaccharide metabolism into two types of distinct pathways: one producing glycogen controlled by photosynthesis and the other producing starch controlled independently possibly by Nitrogen metabolism. Gene losses, acquisition of multicellularity and adaptation and diversification of cyanobacteria may have yielded more simple networks from a subset of this very ancient and complex pathway.

Materials and Methods

Strains and culture conditions

187G11 mutant of *Cyanobacterium* sp. CLg1 was obtained through UV mutagenesis campaign and grown in Artificial Sea Water medium (Rippka et al., 1979) in the absence (AS0 medium) or in the presence of nitrogen source provided by 0.88 mM of sodium nitrate (ASNIII medium) as described in Cenci et al., 2013.

Transmission electron microscopy observation

Cyanobacteria were cultivated in 50 mL nitrogen-deprived medium (ASO) medium and harvested by centrifugation (5 min at 4000 g at 4 °C) after 2 weeks. The cells were fixed with glutaraldehyde, post-fixed with osmium tetroxide and embedded in Epon resin. 70 nm-thin sections were cut with a diamond knife in a Leica UC6 microtome and post-stained with periodic acid thiosemicarbazide silver proteinate (PATAg) (Gallant & Guilbot, 1969). Drops

of dilute suspensions of water-soluble polysaccharide fractions were deposited on glow-discharged carbon-coated copper grids and allowed to dry after negative staining with 2% uranyl acetate. All specimens were observed with a Philips CM200 transmission electron microscope (TEM) operating at 80 kV. Images were recorded on Kodak SO163 films.

Purification and structural analysis of water-soluble polysaccharide (WSP)

WSP were purified from wild type and mutant strains cultivated in 300 mL of ASO liquid medium during 12 days and harvested at the middle of the day by centrifugation at 3600 g during 15 min at 4 °C. The cell suspension (10 mL) was disrupted through a French Press. Starch-pellets were separated from WSP by spinning the lysate at 16000 g for 15 min at 4 °C. WSP in the supernatant and starch-pellet were quantified by amyloglucosidase assay following the instructions of R-Biopharm. Results are expressed in mg of polysaccharide/milligram of total protein. The total protein concentration was determined in the supernatant using the Bradford method (Bio-Rad). WSP were sized on exclusion chromatography (Toyopearl TSK HW 55) pre-equilibrated at 1 $mL.min^{-1}$ in 10% DMSO (D = 1.8 cm L = 60 cm). Polysaccharides were quantified in each fraction (1 mL) by the phenol-sulfuric acid method (Fox & Robyt, 1991). Polysaccharides contained in fractions 35 to 45 were pooled and further incubated *Pseudomonas* sp. isoamylase (1U) and pullulanase (1U) (Megazyme), in Sodium Acetate 55 mM pH 3.5. The linear glucan chains were separated according to their degree of polymerization by HPAEC-PAD as described previously (Colleoni et al., 1999).

Zymogram analysis

Cells were grown for 10 days in 3 liters liquid ASNIII medium and harvested at the middle of the day by centrifugation (3000 g at 4°C during 15 min.). The cell pellets were washed three times with 20 mL of cold Tris-acetate buffer (25 mM Tris-acetate, pH 7.5, 10 mM DTT) before disrupting by French press at 1250 psi. The lysate was centrifuged at 16000g during 15 minutes at 4°C. The supernatant (20 mL) was loaded on preparative anion-exchange chromatography (AEC) column (HitrapQ sepharose FF, 5ml column volume, GE Healthcare) pre-equilibrated in buffer A (150 mM NaCl, 25 mM Tris-acetate, pH 7.5, 5 mM dithiothreitol, 10% glycerol). The proteins were eluted at 4 $ml.min^{-1}$ using buffer B (150 mM NaCl, 25 mM Tris-acetate, pH 7.5, 5 mM dithiothreitol, 10% glycerol, 1 M NaCl) in 25 mL. Eluted proteins

were desalted and concentrated to 1 milliliter using ultrafiltration system (Millipore). The semi-purified crude extracts were separated by non-denaturing polyacrylamide gel electrophoresis (PAGE) containing 0.6% of rabbit glycogen (Sigma-Aldrich). After electrophoresis, gels were incubated overnight at room temperature in starch synthase buffer (70 mM Gly-gly pH 7.5, 135 mM $(NH_4)_2SO_4$, 280 mM NaF, 330 mM trisodium citrate, 290 mM sodium acetate, 3 mM ADP-glucose, 67 mM β-mercapto-ethanol). Starch synthase activities were then visualized as dark activity bands after soaking native PAGE in iodine solution (0.5 g I_2, 10 g KI).

Gene cloning and sequencing

Starch metabolism genes: glgC (KR020055), glgA1 (AHB52787), glgA2 (AHB52788), gbss (AHB52786), glgB1 (AFP43334), glgB2 (AFP43335), glgB (AFP43336), glgB4 (AHB52790), glgX1 (AGI19288), glgX2 (AGI19289), apu13 (AHB52783), apu57 (AHB52784), amg (AHB52785), glgP (AHB52789), malQ (AHB52791) were amplified from genomic DNA of mutant strain using primers designed in the untranslated region as described in Cenci et al., 2013. Starch/glycogen synthase genes (*glgA1*, *glgA2* and *glgA2**) were amplified from genomic DNA of wild type (*glgA1* and *glgA2*) and 187G11 mutant strains (*glgA2**). Primers include restriction sites in order to clone the *glgA* genes either in pGex (GE Healthcare) or pET15 (Novagen) expression vectors (underlined letters) :

BamHI-*glgA*1-pGexF GGATCCATTCCCTCTGAGTCTGTGTGGCAG GCAA,
NcoI-*glgA*1-pET15F CCATGGGCAAAATATTATTTGTGGCGGCAGAAGCATC,
XhoI-*glgA*1R CTCGAGTTAAATAATTCCATCGATCGCATCTTGATAC,
EcoRI-*glgA*2-pGexF GAATTCTATATAGTTCAAATTGCCTCCGAATGTCCT,
NcoI-*glgA*2/*glgA*2*-pET15F CTCGAGTATATAGTTCAAATTGCCTCCGAATGTCCT,
XhoI-*glgA*2/*glgA2**R CTCGAGTTACTTATCTCTTAAAAAATCATATAATTCA,
The PCR experiments were conducted at 95 °C for 5 min; 30 cycles of denaturation at 98 °C for 30 s, annealing 30 s at 59.6 °C for *glgA*1, *glgA*2 and *glgA**, and extension 1 min 30 s at 72 °C, and a final elongation step at 72 °C for 5 min. The PCR products were cloned into pCR-BluntII-TOPO vector (Invitrogen) and transferred into the chemical competent *E. coli* TOP10 Mach1™-TR, and plated on LB agar with kanamycin and X-gal. Purified plasmids were sequenced by GATC Biotech Company according to Sanger methods. Each gene was sequenced on both strands using additional primers when required. The presence of mutation was identified by alignment with the wild type gene using the BLASTn program. The

519 insertion of the cloning product in pGEX or pET15 was done by the T4 ligase (Thermo
520 Scientific) and transferred into the chemical competent *E. coli* TOP10 Mach1TM-TR, and
521 plated on LB agar with Ampicillin. Purified plasmids were also sequenced by GATC Biotech
522 Company.
523
524 **Protein expressions in *Escherichia coli* JW3392-1 Δ*glgA***

525 *Escherichia coli* wild type strain (BW25113) and the derivative single knockout *ΔglgA*
526 mutant (JW3392-1) of Keio collection (Baba et al., 2006) were provided by E.coli Stock
527 center (http://cgsc.biology.yale.edu). Δ*glgA* mutant was lysogenized with λDE3 phage in
528 order to insert the inducible T7 RNA polymerase gene (λDE3 lysogenization kit, Novagen).
529 Δ*glgA* DE3 mutant was then transformed with pET-15-*glgA*1, pET-15-*glgA*2, pET-15-
530 *glgA*2* pGEX-*glgA*1, pGEX-*glgA*2. pET and pGEX expression vectors allow the synthesis of
531 recombinant protein without and with N-terminus GST tag, respectively. Transformed *E.coli*
532 strains were grown in 200 mL autoinductible medium (Formedium^{TM}) in the presence of
533 ampicillin (100 µg/mL) at 30°C during 36 hours. The cells were harvested by centrifugation
534 16000g for 10 min at 4 °C and the pellets were resuspended in 5 mL of cold buffer
535 (Tris/acetate 25 mM pH 7.5 10 mM DTT) before lysing the cells by sonication. Crude extacts
536 were fractionnated and stored at -80°C for further analyses. Starch synthase activities in the
537 *E.coli* crude extract were determined either by ^{14}C-glucose incorporation assay (described
538 below) or by zymogram analysis. The complementation experiment was carried out in 250
539 mL M9 liquid medium (38 mM Na_2HPO_4, 22mM KH_2PO_4, 8.5 mM NaCl, 18 mM NH_4Cl
540 chloride, 0.1 mM $CaCl_2$, 2 mM $MgSO_4$, 0.4% casaminoacids.) supplemented with either 2%
541 glucose, mannitol or maltose. After 12 h incubation at 37 °C, the cells are harvested by
542 centrifugation (15 min at 4000 g) and the pellets are washed and resuspended in cold buffer
543 (Tris/acetate 25 mM pH 7.5 10 mM DTT). After sonication (three times for 30 seconds) and
544 centrifugation, the amount of glycogen and protein were measured in the supernatants using
545 amyloglucosidase assay (R-biopharm) and bradford method (Biorad), respectively. The
546 results are expressed as mg WSP.mg^{-1} total protein.

547 **Western blot analysis**

548 Proteins bound to starch granules of CLg1 were extracted by denaturing 1 mg of purified
549 starch granules in 50 µL of SDS/β-mercaptoethanol buffer for 10 min at 95 °C. After

centrifugation at 10000 g during 10 min, proteins specifically attached to starch granules are found in the supernatant. Granule-bound proteins were loaded onto 9 % SDS-PAGE. Western blot was then carried out as described presiously (Ral et al., 2006). Polyclonal primary antibodies raised against GlgA1 and GlgA2 of synechocystis PCC6803 and secondary antibody were diluted at 1:1000 and 1:20000 in blocking buffer, respectively. The immunocomplexes were detected by chimioluminescence following the instructions of ECL^{TM} prime western blotting reagent kit (GE healthcare).

Starch/glycogen synthase assay

The starch synthase activities were measured by following the incorporation of ^{14}C-glucose onto glycogen particles. The reaction was carried out at the initial velocity by incubating 40 µL of enzyme preparation and 60 µL of incubation buffer (50 mM HEPES-NaOH pH 7, 10 mg.mL^{-1} glycogen, 100 mM $(NH_4)_2SO_4$, 10 mM Dithiothreitol , 0.5 mg.mL^{-1} bovine serum albumin, 3 mM of ADP-Glucose and 2 µM ADP-^{14}C-[U]-Glc) for 15 minutes at 35°C. The reaction is stopped by precipating labeled glycogen with 1 mL methanol-KCl (75%v/v; 1%wt/v). The samples are stored at -20 °C during 10 min and then centrifuged 5 min at 3000 g at 4 °C. After centrifugation, the glycogen-pellets are resuspended with 200 µL of distilled water. This step was repeated twice before mixing the sample with 2.5 mL of scintillation liquid. The radioactivity incorporated onto glycogen was determined by liquid scintillation counting.

Phylogenetic tree

Homologs of GlgA were identified in GenBank or other sources using BLASTp and aligned with MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle/). The alignment was manually refined using SeAl (http://tree.bio.ed.ac.uk/software/seal/) and blocks of missing data in some taxa or regions of low identity were manually removed (final alignment of 595 amino acids available from SGB). This reduced alignment was analyzed under maximum likelihood (ML). The best-fitting amino acid substitution model was selected according to the Akaike informational criterion with ProtTest using the default values (Abascal et al., 2005). The LG (Le et al., 2008) model with heterogeneous gamma rate distribution across sites (+ G) was selected by ProtTest for this protein data set. The LG model parameter values were used under RAxML v.7.2.8 (Stamatakis, 2006) for the ML tree searches. The stability of monophyletic groups was assessed using RAxML with 1000 bootstrap replicates.

Supplemental data

Supplemental Figure S1: Comparison of growth rates of 187G11 mutant and wild-type strains in continuous and day/night cycles.

Supplemental Figure S2: Analysis of starch metabolizing enzymes in the crude extract of 187G11.

Supplemental Figure S3: Semi-quantitative assay of GlgX2 activity in the 187G11 extract.

Supplemental Figure S4: Highly conserved tyrosine residues in starch/glycogen synthase

Supplemental Figure S5: GST-GlgA1 and GST-A2 recombinant proteins.

Supplemental Figure S6: GlgA2/SSIII self-priming activities.

Supplemental Figure S7: Detailed phylogenetic tree of glycogen/starch synthases belonging to SSIII/SSIV/GlgA2 family

Supplemental Figure S8: Maximum likelihood phylogenies of GlgA2 and 16S RNA of cyanobabacteria.

Supplemental figure S9: Congruence of maximum likelihood phylogenies of GlgA2 and 16S RNA of cyanobacteria.

Acknowledgments

SB and CC were supported by CNRS, Université des Sciences et Technologies de Lille, Région Nord-Pas-de-Calais, and Agence Nationale de la Recherche (grant # ANR-BLAN07-3_186613).

Authors contributions: C.C. and S. G. B. designed research and wrote the paper; K.D. and M.D. have performed the molecular characterization of 187G11 mutant strain; U.C and C.T. carried out the mutagenesis campaign and the screening process; E.S and Y.N. sequenced the CLg1 genome; TEM observations were performed by J-L.P and A.D.T.; A.S. and S.G. characterized the water soluble polysaccharide; S.D-T and F.J.F. produced the antibodies raised against glycogen/starch synthases. M-C.A. performed phylogenetic analysis; MD, KD, AS and MP have expressed GlgA1 and GlgA2 recombinant proteins.

Activity	Gene	Cazy Classification	Accession N°	sequencing
ADP-glucose pyrophosphorylase	*glgC*	-	KR020055	+
glycogen/starch synthase	*glgA1*	GT5	AHB52787	+
glycogen/starch synthase	*glgA2*	GT5	AHB52788	K480N
glycogen/starch synthase	*gbss*	GT5	AHB52786	+
Branching enzyme	*glgB1*	GH13	AFP43334	+
Branching enzyme	*glgB2*	GH13	AFP43335	+
Branching enzyme	*glgB3*	GH13	AFP43336	+
Putative branching enzyme	*glgB4*	GH57	AHB52790	+
Debranching enzyme	*glgX1*	GH13	AGI19288	+
Debranching enzyme	*glgX2*	GH13	AGI19289	+
Debranching enzyme (amylopullulanase-GH13)	*apu13*	GH13	AHB52783	+
Putative debranching enzyme (amylopullulanase-GH57)	*apu57*	GH57	AHB52784	+
Debranching enzyme (amylo-1,6 glucosidase)	*amg*	GH133	AHB52785	+
Phosphorylase	*glgP*	GT35	AHB52789	+
α-1,4 glucanotransferase	*malQ*	GH77	AHB52791	+

Table 1: Summary of starch metabolism genes sequenced in the 187G11 mutant. Each gene was amplified using primers designed in the untranslated region. PCR products were cloned and sequenced on both strands using additional primers when required. GH and GT stand for Glycosyl Hydrolase and Glycosyl Transferase, respectively.

625

Parsed Citations

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21: 2104-2105
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y, Baba M, Datsenko KA, Tomita M, Wanner BL, Mori H (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2: 2006 0008
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ball SG, Colleoni C, Arias MC (2015) The transition from glycogen to starch metabolism in cyanobacteria and eukaryotes. In: Nakamura Y, ed. Starch: metabolism and structure. Springer, 93-158.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ball SG, Subtil A, Bhattacharya D, Moustafa A, Weber AP, Gehre L, Colleoni C, Arias MC, Cenci U, Dauvillee D (2013) Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis? Plant Cell 25: 7-21
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Bertoft E, Laohaphatanalert K, Piyachomkwan K, Sriroth K (2010) The fine structure of cassava starch amylopectin. Part 2: building block structure of clusters. Int J Biol Macromol 47: 325-335
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cenci U, Chabi M, Ducatez M, Tirtiaux C, Nirmal-Raj J, Utsumi Y, Kobayashi D, Sasaki S, Suzuki E, Nakamura Y, Putaux JL, Roussel X, Durand-Terrasson A, Bhattacharya D, Vercoutter-Edouart AS, Maes E, Arias MC, Palcic M, Sim L, Ball SG, Colleoni C (2013) Convergent evolution of polysaccharide debranching defines a common mechanism for starch accumulation in cyanobacteria and plants. Plant Cell 25: 3961-3975
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Cenci U, Nitschke F, Steup M, Minassian BA, Colleoni C, Ball SG (2014) Transition from glycogen to starch metabolism in Archaeplastida. Trends Plant Sci 19: 18-28
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Colleoni C, Dauvillée D, Mouille G, Morell M, Samuel M, Slomiany MC, Liénard L, Wattebled F, d'Hulst C, Ball S (1999) Biochemical characterization of the chlamydomonas reinhardtii alpha-1,4 glucanotransferase supports a direct function in amylopectin biosynthesis. Plant Physiol 120: 1005-1014
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Colleoni C, Suzuki E (2012) Storage polysaccharide metabolism in Cyanobacteria.In: Tetlow I, ed. Essential reviews in experimental biology: Starch: Origins, Structure and metabolism. Society Experimental Biology, 217-254.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Crowe SA, Dossing LN, Beukes NJ, Bau M, Kruger SJ, Frei R, Canfield DE (2013) Atmospheric oxygenation three billion years ago. Nature 501: 535-538
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Crumpton-Taylor M, Pike M, Lu KJ, Hylton CM, Feil R, Eicke S, Lunn JE, Zeeman SC, Smith AM (2013) Starch synthase 4 is essential for coordination of starch granule formation with chloroplast division during Arabidopsis leaf expansion. New Phytol 200: 1064-1075
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Deschamps P, Colleoni C, Nakamura Y, Suzuki E, Putaux JL, Buleon A, Haebel S, Ritte G, Steup M, Falcon LI, Moreira D, Loffelhardt W, Raj JN, Plancke C, d'Hulst C, Dauvillee D, Ball S (2008) Metabolic symbiosis and the birth of the plant kingdom. Mol Biol Evol 25: 536-548
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

D'Hulst C, Wattebled F, Szydlowski N (2015) Starch biosynthesis in leaves and its regulation. In: Nakamura Y, ed. Starch:

metabolism and structure. Springer, 161-210.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Falcon LI, Lindvall S, Bauer K, Bergman B, Carpenter EJ (2004) Ultrastructure of unicellular N2 fixing cyanobacteria from the tropical north atlantic and subtropical north pacific oceans. J phycol 40: 1074-1078
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Fox JD, Robyt JF (1991) Miniaturization of three carbohydrate analyses using a microsample plate reader. Anal Biochem 195: 93-96
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Fu J, Xu X (2006) The functional divergence of two glgP homologues in Synechocystis sp. PCC 6803. FEMS microbiology letters 260: 201-209
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gallant D, Guilbot A (1969) Etude de l'ultrastructure du grain d'amidon à l'aide de nouvelles méthodes de préparation en microscopie électronique. Die Stärke 6: 156-163
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Grundel M, Scheunemann R, Lockau W, Zilliges Y (2012) Impaired glycogen synthesis causes metabolic overflow reactions and affects stress responses in the Cyanobacterium Synechocystis sp. PCC 6803. Microbiology 158: 3032-3043
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Gupta RS (2009) Protein signatures (molecular synapomorphies) that are distinctive characteristics of the major cyanobacterial clades. Int J Syst Evol Microbiol 59: 2510-2526
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Hizukuri S (1986) Polymodal distribution of the chain lengths of amylopectin and the crystalline structure of starch granules. Carbohydr Res 147: 342-347
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Laohaphatanaleart K, Piyachomkwan K, Sriroth K, Bertoft E (2010) The fine structure of cassava starch amylopectin. Part 1: Organization of clusters. Int J Biol Macromol 47: 317-324
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Le SQ, Lartillot N, Gascuel O (2008) Phylogenetic mixture models for proteins. Philos Trans R Soc Lond B Biol Sci 363: 3965-3976
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Melendez-Hevia E, Waddell TG, Shelton ED (1993) Optimization of molecular design in the evolution of metabolism: the glycogen molecule. Biochem J 295 (Pt 2): 477-483
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Nakamura Y (2015) Biosynthesis of reserve starch. In: Nakamura Y, ed. Starch: metabolism and structure. Springer, 161-210.
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Posewitz MC, Smolinski SL, Kanakagiri S, Melis A, Seibert M, Ghirardi ML (2004) Hydrogen photoproduction is attenuated by disruption of an isoamylase gene in Chlamydomonas reinhardtii. Plant Cell 16: 2151-2163
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Ral JP, Colleoni C, Wattebled F, Dauvillee D, Nempont C, Deschamps P, Li Z, Morell MK, Chibbar R, Purton S, d'Hulst C, Ball SG (2006) Circadian clock regulation of starch metabolism establishes GBSSI as a major contributor to amylopectin synthesis in Chlamydomonas reinhardtii. Plant Physiol 142: 305-317
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY (1979) Generic Assignments, strain histories and properties of pure cultures of cyanobacteria. Journal of General Microbiology 111: 1-61
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Sanchez-Baracaldo P, Ridgwell A, Raven JA (2014) A neoproterozoic transition in the marine nitrogen cycle. Curr Biol 24: 652-657
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

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
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Stamatakis A (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688-2690
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Summons RE, Jahnke LL, Hope JM, Logan GA (1999) 2-Methylhopanoids as biomarkers for cyanobacterial oxygenic photosynthesis. Nature 400: 554-557
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Szydlowski N, Ragel P, Raynaud S, Lucas MM, Roldan I, Montero M, Munoz FJ, Ovecka M, Bahaji A, Planchot V, Pozueta-Romero J, D'Hulst C, Merida A (2009) Starch granule initiation in Arabidopsis requires the presence of either class IV or class III starch synthases. Plant Cell 21: 2443-2457
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title

Yoo SH, Lee BH, Moon Y, Spalding MH, Jane JL (2014) Glycogen Synthase Isoforms in Synechocystis sp. PCC6803: Identification of Different Roles to Produce Glycogen by Targeted Mutagenesis. PLoS One 9: e91524
Pubmed: Author and Title
CrossRef: Author and Title
Google Scholar: Author Only Title Only Author and Title