Photosynth Res
DOI 10.1007/s11120-010-9615-z

REGULAR PAPER

Evolution of enzymes involved in the photorespiratory 2-phosphoglycolate cycle from cyanobacteria via algae toward plants

Ramona Kern · Hermann Bauwe · Martin Hagemann

Received: 15 September 2010 / Accepted: 22 December 2010

Abstract The photorespiratory pathway was shown to be essential for organisms performing oxygenic photosynthesis, cyanobacteria, algae, and plants, in the present day O_2-containing atmosphere. The identification of a plant-like 2-phosphoglycolate cycle in cyanobacteria indicated that not only genes of oxygenic photosynthesis but also genes encoding photorespiratory enzymes were endosymbiotically conveyed from ancient cyanobacteria to eukaryotic oxygenic phototrophs. Here, we investigated the origin of the photorespiratory pathway in photosynthetic eukaryotes by phylogenetic analysis. We found that a mixture of photorespiratory enzymes of either cyanobacterial or α-proteobacterial origin is present in algae and higher plants. Three enzymes in eukaryotic phototrophs clustered closely with cyanobacterial homologs: glycolate oxidase, glycerate kinase, and hydroxypyruvate reductase. On the other hand, the mitochondrial enzymes of the photorespiratory cycle in algae and plants, glycine decarboxylase subunits and serine hydroxymethyltransferase, evolved from proteobacteria. Other than most genes for proteins of the photosynthetic machinery, nearly all enzymes involved in the 2-phosphogylcolate metabolism coexist in the genomes of cyanobacteria and heterotrophic bacteria.

Keywords Cyanobacteria · Eukaryotic algae · Evolution · Photorespiration · Phylogeny · Plant

Abbreviations

2PG	2-Phosphoglycolate
3PGA	3-Phosphoglycerate
BS	Bootstrap value
CCM	Inorganic carbon concentrating mechanism
GDC	Glycine decarboxylase
GGT	Glutamate:glyoxylate aminotransferase
GK	Bacterial-type glycerate kinase
GLYK	Plant-type glycerate kinase
GOX	Glycolate oxidase
HPR	Hydroxypyruvate reductase
PDC	Pyruvate dehydrogenase
PGP	2-Phosphoglycolate phosphatase
Rubisco	Ribulose 1,5-bisphosphate carboxylase/oxygenase
SGT	Serine:glyoxylate aminotransferase
SHMT	Serine hydroxymethyltransferase

Introduction

The present day O_2-containing atmosphere was mainly generated via the biological activity of oxygenic photosynthesis. Organisms performing this type of photosynthesis use light as energy source to split water for the gain of energy-rich reductants for CO_2 assimilation. Oxygenic photosynthesis evolved about 2.5 billion years ago among cyanobacteria. During their long evolution, cyanobacteria developed a remarkable diversity regarding morphology and physiology that is reflected in a broad range of genome sizes, ranging from 1.6 to 9.7 Mbp (http://genome.

Electronic supplementary material The online version of this article (doi:10.1007/s11120-010-9615-z) contains supplementary material, which is available to authorized users.

R. Kern · H. Bauwe · M. Hagemann (✉)
Universität Rostock, Institut für Biowissenschaften, Abteilung Pflanzenphysiologie, Albert-Einstein-Straße 3, 18059 Rostock, Germany
e-mail: martin.hagemann@uni-rostock.de

Published online: 11 January 2011

kazusa.or.jp/cyanobase/). The analysis of the about 50 presently available genome sequences allowed the distinction of four main cyanobacterial clades (Gupta and Mathews 2010). Among them, the two subtypes of α- and β-cyanobacteria can be distinguished, which were initially identified by their distinct CO_2 assimilation proteins (Badger et al. 2002). Beside genome duplication, extensive lateral gene transfer was involved in the evolution of the different cyanobacterial genome organizations. According to the endosymbiotic theory (Mereschkowsky 1905), an ancient cyanobacterium was stably engulfed as ancestor for the chloroplasts in all eukaryotic photosynthetic organisms. The primary appearance of plastids occurred about 1.5 billion years ago and gave rise to the development of the glaucophyte, red and green algae and later on higher plants (Yoon et al. 2004). Recent genome comparisons indicate that the ancient cyanobacterium, which provided the starting point for the evolution of chloroplasts, contained a relatively large genome comparable to filamentous, N_2-fixing strains such as *Nostoc* (*Anabaena*) sp. PCC 7120 (Deusch et al. 2008). Subsequently, green and mainly red algal secondary endosymbiosis occurred resulting in the diversity of extant algae groups (for reviews, see Keeling 2010; Parker et al. 2008).

Obviously, the invention of oxygenic photosynthesis allowed for the universal distribution of cyanobacteria in all light-exposed habitats; however, the generation of O_2 and later the O_2-rich atmosphere introduced problems for the coupled process of CO_2 fixation. All oxygenic phototrophs use ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) as the main carboxylating enzyme. Rubisco is a very ancient enzyme, which evolved from enolases involved in the methionine salvage cycle of bacteria (Ashida et al. 2003). While in the ancient, CO_2-rich and O_2-free atmosphere, the rather low affinity and specificity of Rubisco for CO_2 was sufficient, in the present day atmosphere with low CO_2/O_2 ratio, the carboxylating activity of Rubisco is often limiting the photosynthetic rate in oxygenic phototrophs. In order to adapt to the increasingly unfavorable CO_2 conditions, many organisms evolved an efficient inorganic carbon concentrating mechanism (CCM). Among cyanobacteria, the CCM employs high-affinity uptake systems for CO_2 and bicarbonate, while Rubisco is concentrated inside carboxysomes where carbonic anhydrase converts the intracellular bicarbonate back into CO_2 (Badger et al. 2006; Kaplan and Reinhold 1999). A different mode of CO_2 enrichment evolved in C4 plants, which use phosphoenolpyruvate-carboxylase as primary carboxylating enzyme in the leaf mesophyll. The produced C_4 acids are subsequently transported and decarboxylated within the Rubisco-containing cells of the bundle sheath. Many eukaryotic algae show also clear CCM activity, but the underlying mechanisms are less well investigated (Giordano et al. 2005; Raven et al. 2008).

In addition to these effects caused by limiting CO_2, elevated O_2 levels resulted in the production of 2-phosphoglycolate (2PG) by the oxygenase activity of Rubisco. In plants, this toxic compound is converted into the Calvin–Benson cycle intermediate 3-phosphoglycerate (3PGA) via the photorespiratory 2PG cycle, which requires a number of enzymes localized in different subcellular compartments, the chloroplast, the peroxisome, the mitochondrion, and the cytosol (Fig. 1; Bauwe et al. 2010; Tolbert 1997). The cycle starts in the chloroplast, in which 2PG is hydrolyzed to glycolate. In the peroxisome, glycolate is then oxidized by glycolate oxidase (GOX) to glyoxylate, which is subsequently aminated to glycine. Next, two molecules of glycine are converted into serine, CO_2, and NH_4^+ by the concerted action of glycine decarboxylase (GDC), composed of the four subunits P-, H-, T-, and L-protein, and serine hydroxymethyltransferase (SHMT) in the mitochondrion. Again in the peroxisome, serine is deaminated to hydroxypyruvate, which is subsequently reduced to glycerate by hydroxypyruvate reductase (HPR1). Finally, glycerate is phosphorylated in the chloroplast to 3PGA by glycerate 3-kinase (GLYK). Plant mutants with defects in this pathway show a typical "photorespiratory phenotype," i.e., they can only grow at enhanced CO_2 concentration suppressing Rubisco oxygenase reaction and bleach under ambient conditions (Bauwe et al. 2010; Somerville 2001). It is hence well accepted that the operation of the photorespiratory cycle is essential for the survival of plants in air. While it was initially believed that, due to the evolution of efficient CCMs, cyanobacteria and C_4 plants do not require a photorespiratory 2PG cycle, recent results clearly demonstrated that it is essential for these organisms as well (Eisenhut et al. 2008; Zelitch et al. 2009).

In evolutionary terms, the occurrence of a plant-like photorespiratory 2PG cycle among cyanobacteria gave rise to the hypothesis that not only oxygenic photosynthesis but also the photorespiratory 2PG cycle was endosymbiotically conveyed from ancient cyanobacteria into eukaryotic oxygenic phototrophs (Eisenhut et al. 2008). In this article, we report on a systematic phylogenetic analysis of the enzymes composing the plant-like photorespiratory 2PG cycle. To bridge the evolutionary gap between cyanobacteria and streptophytes, we included putative photorespiratory enzymes from different eukaryotic algae into the study. Our analysis of 11 photorespiratory proteins in cyanobacteria, algae and plants revealed their almost even divergence from cyanobacteria (ancestor of chloroplasts) and proteobacteria (ancestor of mitochondria).

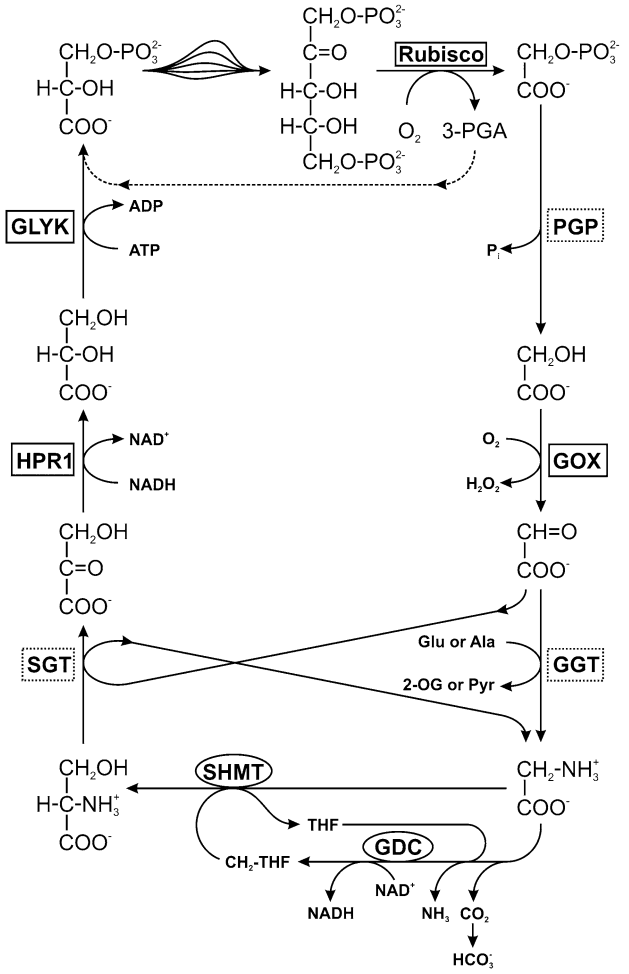

Fig. 1 Schematic display of the 2-phosphoglycolate metabolism in cyanobacteria, algae, and plants summarizing major results of our analyses. *Encircled enzyme* names represent plant and algal proteins most closely related to their α-proteobacterial homologs, while *framed enzyme* names are most closely related to their cyanobacterial counterpart. The *dashed frames* indicate plant and algal proteins for which the phylogeny could not resolved

Materials and methods

Sequence information

The sequences of the proteins involved in the 2PG-metabolism in *Arabidopsis thaliana* (Bauwe et al. 2010; Foyer et al. 2009) and *Nostoc* (*Anabaena*) sp. PCC 7120 (Eisenhut et al. 2008) were extracted from the databases TAIR (http://www.arabidopsis.org/) and CyanoBase (http://genome.kazusa.or.jp/cyanobase/), respectively, and were subsequently used to identify similar proteins in the completely sequenced genomes of 11 streptophytes or eukaryotic algae as well as 12 cyanobacteria using the BlastP algorithm (Altschul et al. 1990). Furthermore, the incomplete genomes of *Nodularia spumigena* CCY9414 and *Emiliania huxleyi* were included in this study. Homologous proteins of *Escherichia coli* K12 and *Rhodopseudomonas palustris* HaA2 and CGA009 were used as representatives of γ- and α-proteobacteria, respectively. These sequences were retrieved from the following databases: (1) GenBank (http://www.ncbi.nlm.nih.gov/genbank/) for *Populus trichocarpa*, *Zea mays*, *Physcomitrella patens*, *Oryza sativa*, *Sorghum bicolor*, *Ostreococcus tauri*, *Chlamydomonas reinhardtii*, *Micromonas* sp. RCC299, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, *Nodularia spumigena* CCY9414, *Escherichia coli* K12, *Rhodpseudomonas palustris* HaA2, and CGA009; (2) JGI (http://www.jgi.doe.gov/) for *Sorghum bicolor* and *Emiliania huxleyi*; (3) *Cyanidioschyzon merolae* genome browser (http://merolae.biol.s.u-tokyo.ac.jp/) (Matsuzaki et al. 2004); (4) CyanoBase for all cyanobacteria except *N. spumigena*. The accession numbers or ORF names and the E-values of similar proteins are listed in the Supplementary Tables S1 and S2. Additional, 16S rRNA gene sequences from cyanobacteria or chloroplasts of eukaryotic phototrophs were retrieved from GenBank or CyanoBase.

Phylogenetic analysis

For phylogenetic analyses, homologous nucleotide and amino acid sequences were aligned using the ClustalW algorithm (Thompson et al. 1994) integrated in the BioEdit Sequence alignment editor (Hall 1999). The alignment was checked manually, and clearly non-conserved regions of the alignment were removed using GBlocks (Castresana 2000). Maximum likelihood tree of 16S rRNA and photorespiratory proteins were inferred using PhyML (Guindon and Gascuel 2003) performed through the phylogeny.fr web service (Dereeper et al. 2008). For the 16S rRNA tree, the Generalised Time Reversible (GTR) substitution model was selected assuming an estimated proportion of invariant sites and four gamma-distributed rate categories to account for rate heterogeneity across sites. Reliability for internal branch was assessed using the bootstrapping method (400 bootstrap replicates). The Le and Gascuel evolutionary model (Le and Gascuel 2008) was selected for the protein phylogenies assuming an estimated proportion of invariant sites and a gamma correction (four categories). Bootstrap values (BS) were inferred from 400 replicates. Graphical representation and edition of the phylogenetic tree were performed with TreeDyn (v198.3) (Chevenet et al. 2006).

Results and discussion

Organism phylogeny based on 16S rRNA sequences

A phylogenetic tree of oxygenic photosynthetic organisms was derived from 16S rRNA sequences of cyanobacteria and from chloroplasts of eukaryotic phototrophs. The

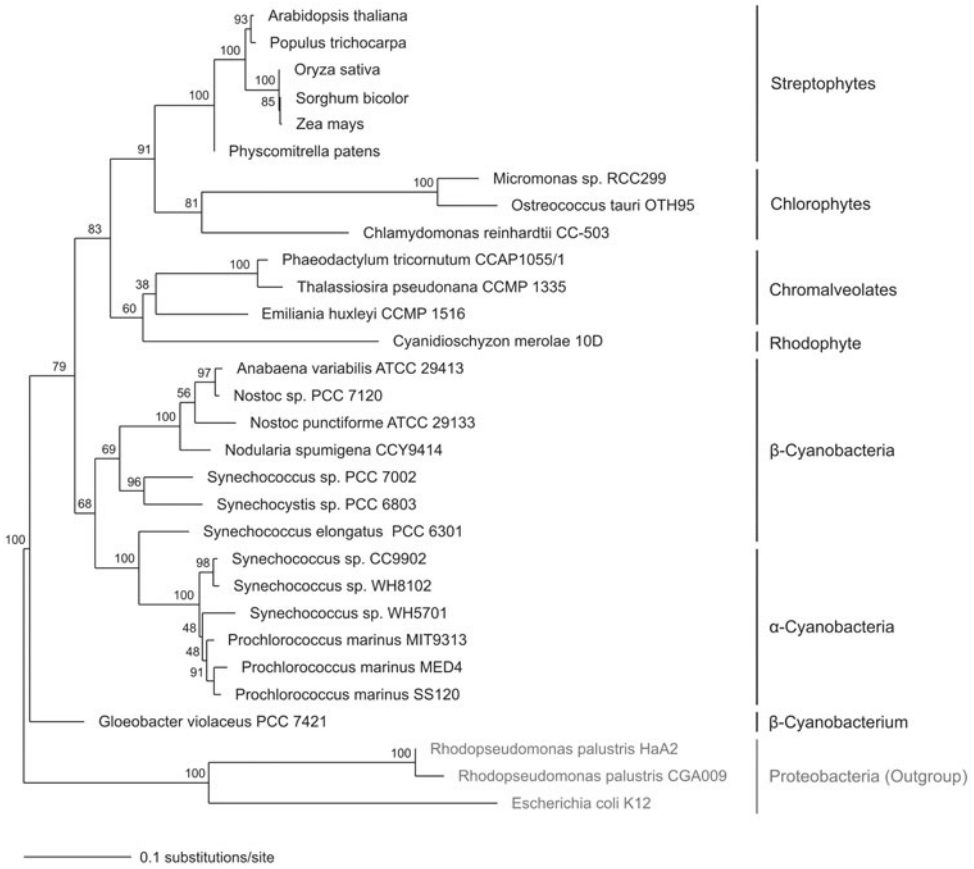

Fig. 2 Maximum likelihood tree of 16S rRNA genes based on 30 organisms. *Numbers* at the node indicated bootstrap values (%) for 400 replicates. The distance scale (substitutions per site) is shown in the bottom left-hand corner

phylogeny presented here (Fig. 2) corresponds well with current models of the evolution of photosynthetic organisms (e.g., Le Corguille et al. 2009; Nelissen et al. 1995; Worden et al. 2009). Two major clades were observed comprising cyanobacteria and all eukaryotic phototrophs as sister groups. Eukaryotic organisms are clustered into one group of red algae and Chromalveolates, which arose from a secondary endosymbiotic event engulfing a red alga. Green algae and land plants (Streptophyta) comprise the second group of eukaryotic phototrophs. *Gloeobacter violaceus*, a primitive cyanobacterium without thylakoid membranes, is found at the base of all oxygenic phototrophs when *Rhodopseudomas* and *E. coli* 16S rRNA sequences are used as outgroup. The large cluster of cyanobacteria is mainly separated in α- and β-cyanobacteria with the exception of *Synechococcus elongatus* PCC 7942 as observed before by Gupta and Mathews (2010). It shall be mentioned that an identical clustering was obtained when combined 16S, 23S, and ITS rRNA sequences were used for the phylogenetic analysis.

This phylogenetic tree served as starting point to evaluate the phylogenetic trees of proteins involved in photorespiratory 2PG metabolism for the presence or absence of a cyanobacterial origin. If the proteins in the eukaryotic clades cluster as sister group to the cyanobacterial proteins, a cyanobacterial origin is assumed.

Plant photorespiratory proteins with a cyanobacterial origin

Glycolate oxidase

It was assumed that all cyanobacteria perform the glycolate-into-glyoxylate conversion by glycolate dehydrogenases, while the plants' photorespiratory cycle employs the peroxisomal enzyme GOX for this purpose (Eisenhut et al. 2008). Glycolate dehydrogenases, which are evolutionary related to the cyanobacterial enzymes, are found in plant mitochondria but their function is not directly related to the photorespiratory 2PG cycle (Bari et al. 2004; Engqvist et al. 2009). However, BlastP searches using the *Arabidopsis* GOX1 revealed GOX-like proteins among cyanobacteria which form the sister group to GOX proteins from phototrophic eukaryotes (Fig. 3). This relationship between the cyanobacterial and the phototrophic eukaryotic clade strongly supports a cyanobacterial origin of GOX proteins (BS: 98%). Surprisingly, the cyanobacterial clade exclusively contains members of the β-cyanobacteria capable to

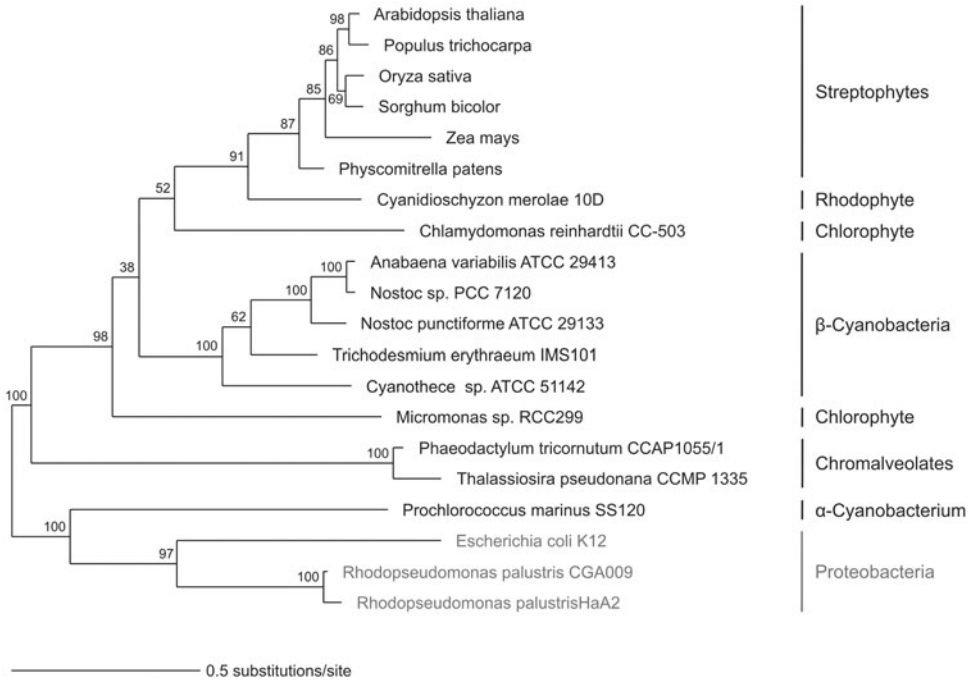

Fig. 3 Maximum likelihood tree of glycolate oxidases (GOX) based on 20 sequences. *Numbers* at the node indicated bootstrap values (%) for 400 replicates. The distance scale (substitutions per site) is shown in the *bottom left-hand corner*. To expand the dataset, putative glycolate oxidases of *Trichodesmium erythraeum* IMS101 (Tery_2398) and *Cyanothece* sp. ATCC 51142 (cce_1717) were included. Further information (protein accession numbers and sequence similarities) is given in Supplemental Tables S1 and S2

fix N_2. It was reported that Chlorophytes including *Chlamydomonas reinhardtii* and *Micromonas* sp., just as all cyanobacteria, possess a NAD^+-dependent glycolate dehydrogenase for the photorespiratory glycolate-into-glyoxylate conversion (Atteia et al. 2009; Beezley et al. 1976; Nelson and Tolbert 1970). Different from these reports, our BlastP searches revealed the presence of putative GOX proteins in the Chlorophyte *Chlamydomonas* (Supplementary Table S2, E-value: $8E^{-99}$) and the Prasinophyte *Micromonas* (E-value: $4E^{-74}$). These chlorophytic putative GOX proteins as well as a related red algal protein cluster in-between the plant and the cyanobacterial proteins as it should be expected for a protein of cyanobacterial origin.

Two less similar proteins from the two diatoms *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* build the basal branch of the monophyletic group of GOX-like proteins from N_2-fixing β-cyanobacteria and phototrophic eukaryotes. Similar to green algae, previous studies indicated that glycolate oxidation in diatoms is catalyzed by a glycolate dehydrogenase localized in peroxisomes (Winkler and Stabenau 1995). However, our and previously reported BlastP analyses identified the two GOX-like proteins, one with a peroxisomal targeting PTS1 and the other (XP_002178591 of *Phaeodactylum*) possibly targeted to mitochondria (Kroth et al. 2008). A similar GOX-like protein was detected in the genome of the picoplanktonic cyanobacterium *Prochlorococcus marinus* SS120. Similar GOX-like proteins also exist in the genomes of further *Prochlorococcus* strains (e.g., MIT9211, CCMP1375, NATL2A, NATL1A), which build a monophyletic group with the proteobacterial GOX-like proteins from *E. coli* and *Rhodopseudomonas* (Fig. 3 shows strain SS120 only). Interestingly, all *Prochlorococcus* strains harboring this proteobacterial-like putative GOX separate from other *Prochlorococcus* and *Synechococcus* strains in phylogenetic analyses of cyanobacteria reported by Gupta and Mathews (2010). It should be mentioned that only the biochemical function of the spinach GOX (Macheroux et al. 1992) and that of *Nostoc* sp. PCC 7120 (own unpublished results) has been verified, while the function of all putative algal and proteobacterial GOX-like enzymes still has to be studied.

Glycerate kinase

Plant glycerate 3-kinases (GLYK) involved in the photorespiratory 2PG cycle are structurally and phylogenetically distinct from glycerate 2-kinases (GK) in bacteria and animals (Bartsch et al. 2008; Boldt et al. 2005). In contrast to most other cyanobacteria, *Synechocystis* sp. PCC 6803 (Bartsch et al. 2008) as well as the thermophilic *Synechococcus* strains JA-3-3Ab and JA-2-3B′ possess a bacterial-type GK. The two *Synechococcus* strains and *Gloeobacter* collectively build the basal branch of the cyanobacterial species tree (Gupta and Mathews 2010). It seems as if the plant-like GLYK evolved among cyanobacteria but was lost from *Synechococcus* JA-3-3Ab and JA-2-3B′a as well as *Synechocystis* PCC 6803 and became replaced by bacterial GKs via horizontal gene transfer (HGT). Because of the different phylogeny, we excluded the bacterial-type GKs from our analyses of the plant GLYKs.

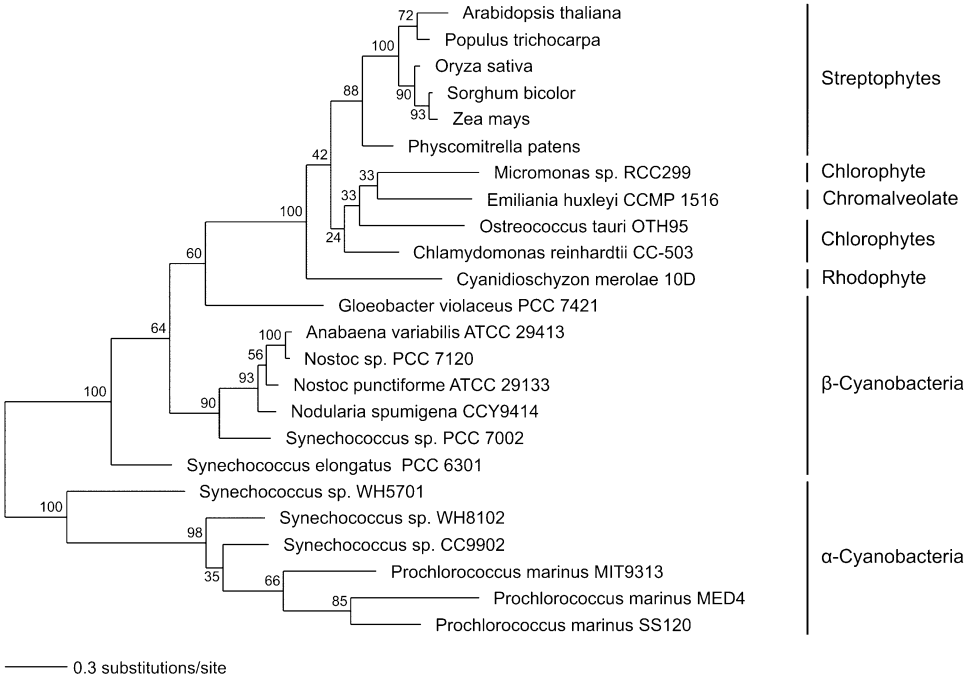

Fig. 4 Maximum likelihood tree of plant-type glycerate kinases (GLYK) based on 24 sequences. *Numbers* at the node indicated bootstrap values (%) for 400 replicates. The distance scale (substitutions per site) is shown in the *bottom left-hand corner*. Further information (protein accession numbers and sequence similarities) is given in Supplemental Tables S1 and S2

The phylogenetic tree of GLYK proteins strongly supports a cyanobacterial origin (Fig. 4). The *β*-cyanobacterial clade builds the sister group to the plant and algal clade supporting a N_2-fixing, heterocyst-forming plastid ancestor (Deusch et al. 2008). Unexpectedly, the GLYK from *Gloeobacter violaceus* PCC 7421 is positioned at the base of the plant and algal clade and not inside the *β*-cyanobacterial clade. A neighbor-joining tree shows this phylogeny, too, with a bootstrap support of 79% (data not shown).

Unfortunately, the relationship among the algal homologs could not be clearly resolved because of low BS among these branches (all <50%). Nevertheless, all algal GLYK proteins cluster in-between GLYKs from cyanobacteria and Streptophytes supporting the GLYK origin among oxygenic phototrophs in the cyanobacterial radiation. The close relationship among the algal GLYKs could result from a cryptic green endosymbiont (e.g., Prasinophyte) in the ancestor of the Chromalveolates before the secondary endosymbiosis with a red alga (Frommolt et al. 2008; Moustafa et al. 2009). In this study, we cannot distinguish whether the GLYKs in the Chromalveolates came from a green or a red algal endosymbiont, whereas Maruyama et al. (2009) postulated a red algal origin of GLYK for the Chromalveolate *Phytophthora*.

As reported before by Kroth et al. (2008), we were unable to identify any protein similar to GLYK as well as to bacterial-type GK in the genomes of the diatoms *Thalassiosira* and *Phaeodactylum*. These authors suggested that the glycine and serine produced in the upstream reactions of the photorespiratory 2PG cycle could be directed to other metabolic pathways. The presence of GLYK proteins in other Chromalveolates such as *Emiliania huxleyii* (Haptophyta) or *Ectocarpus siliculosus* (Phaeophyceae; Accession number CBN78958) point to a loss of this gene in diatoms. It should be noted that GLYK-encoding genes are frequently present in fungal genomes, where they probably originate from a HGT event (Maruyama et al. 2009).

Hydroxypyruvate reductase

Until now, no enzyme was experimentally verified to catalyze the hydroxypyruvate reduction in cyanobacteria. BlastP searches with *Arabidopsis* hydroxypyruvate reductase 1 (HPR1; Timm et al. 2008) identified several cyanobacterial proteins with considerably high similarities to HPR1 which are usually annotated as hydroxyacid reductase/dehydrogenase. Phylogenetic analyses including these enzymes from cyanobacteria showed that only a few putative HPRs from N_2-fixing *β*-cyanobacteria cluster in a monophyletic group with plant and algal HPRs. These proteins were used for reconstructing a maximum likelihood tree of HPR1 (Fig. 5). To expand the data set, additional putative HPRs from the *β*-cyanobacteria *Cyanothece* sp. PCC 8801 and ATCC 51142 (ORF names: PCC8801_3049 and cce_2560) were included. The HPR1 tree separated the cyanobacterial proteins into two groups. The first group including All8087 from *Nostoc* (*Anabaena*) sp. PCC 7120 forms a monophyletic group with proteobacterial representatives as sister to the plant and algal clade (only including *Chlamydomonas* and *Cyanidioschyzon*) as supported by a high BS (100%).

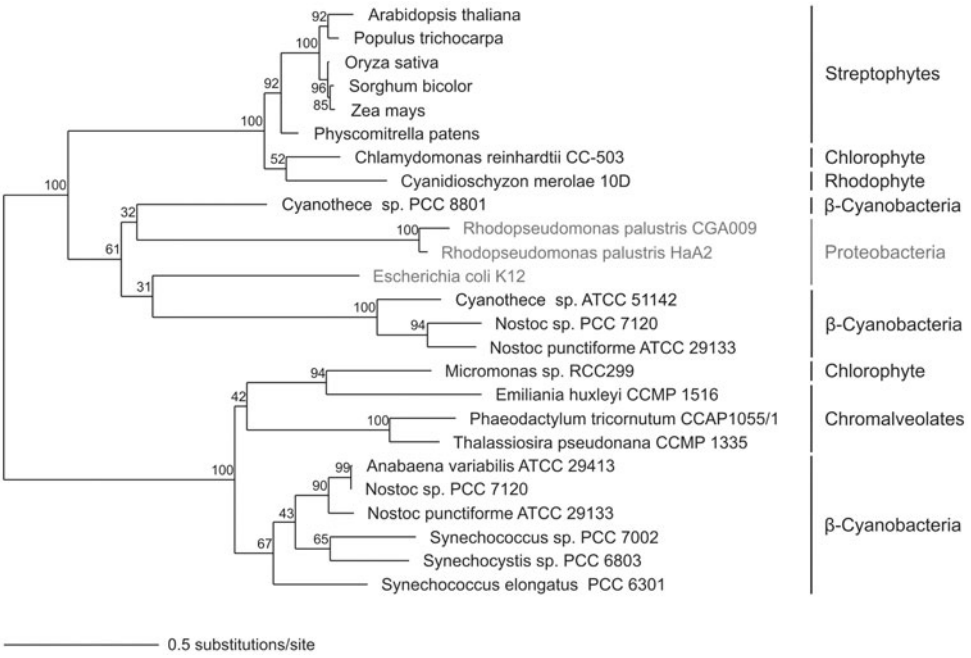

Fig. 5 Maximum likelihood tree of hydroxypyruvate reductases (HPR) based on 25 sequences including the plant photorespiratory HPR1. *Numbers* at the node indicated bootstrap values (%) for 400 replicates. The distance scale (substitutions per site) is shown in the *bottom left-hand corner*. To expand the dataset, putative HPRs of *Cyanothece* sp. PCC 8801 (PCC8801_3049) and *Cyanothece* sp. ATCC 51142 (cce_2560) were included. Further information (protein accession numbers and sequence similarities) is given in Supplemental Tables S1 and S2

Surprisingly, the algal HPRs from the Chromalveolates *Phaeodactylum* and *Emiliania* as well as the Prasinophyte *Micromonas* build a clade with the second group of cyanobacterial HPR-like proteins including the *Nostoc* Alr0058. This sister relation argues for a cyanobacterial origin of HPR in Prasinophytes. The positioning of the *Micromonas* enzyme in a monophyletic group with two diatoms and a Haptophyte as well as the lack of homologs in *Cyanidioschyzon* point to a "green" origin of this gene in Chromalveolates. Further investigations are needed to support this view.

Plant photorespiratory proteins with α-proteobacterial origin

Glycine decarboxylase: P-protein, T-protein, L-protein

In all phylogenies of the GDC proteins (Fig. 6 and Supplementary Figs. S1 and S2), the GDC proteins of α-proteobacteria form a monophyletic group with phototrophic eukaryotes (BS between 69 and 89%). Neighbor-joining phylogenies also show the same tree topology (data not shown) with higher bootstrap support (almost 100%). These phylogenies point to an α-proteobacterial origin of the GDC proteins in mitochondria of plants and algae. The P-, T-, and L-protein phylogenies were generally congruent and supported by high BS for main nodes. The main difference is in the positioning of the *E. coli* T-protein. In the P- and L-protein phylogeny, *E. coli* forms a monophyletic group with *Rhodopseudomonas* and phototrophic eukaryotes, whereas in the T-protein tree *E. coli* appeared at the base of cyanobacteria. Interestingly, the P-protein of *β*-cyanobacteria stays in a sister group relation with the monophyletic group of proteobacteria, plants, and algae (Fig. 6). Possibly, this is the result of an early divergence of α- and *β*-cyanobacteria (Gupta and Mathews 2010). The phylogenies of T- and L-protein could not clearly define the branch order of cyanobacteria, especially for the position of *Gloeobacter violaceous* PCC 7421 and *Synechococcus* sp. PCC 6301 because of low bootstrap support.

In plant chloroplasts, a second L-protein is part of the pyruvate dehydrogenase (PDC) multi-enzyme complex forming acetyl-CoA. In contrast to the mitochondrial L-protein, phylogenetic analysis supported a cyanobacterial origin of the chloroplastidal PDC L-protein (Lutziger and Oliver 2000 and not shown own data).

In contrast to the three large subunits of GDC, the phylogenetic relation of the GDC H-protein subunit could not be finally resolved. The H-protein is rather short, which made statistical save branching impossible. However, it can be assumed that the H-protein subunit is most probably also obtained from the α-proteobacterial endosymbiont for mitochondria as shown above for the other GDC subunits.

Serine hydroxymethyltransferase

The *Arabidopsis* genome encodes five (possibly even seven) SHMT isoforms, which are targeted to several subcellular compartments including mitochondria (Bauwe et al. 2010; McClung et al. 2000; Voll et al. 2006). In this

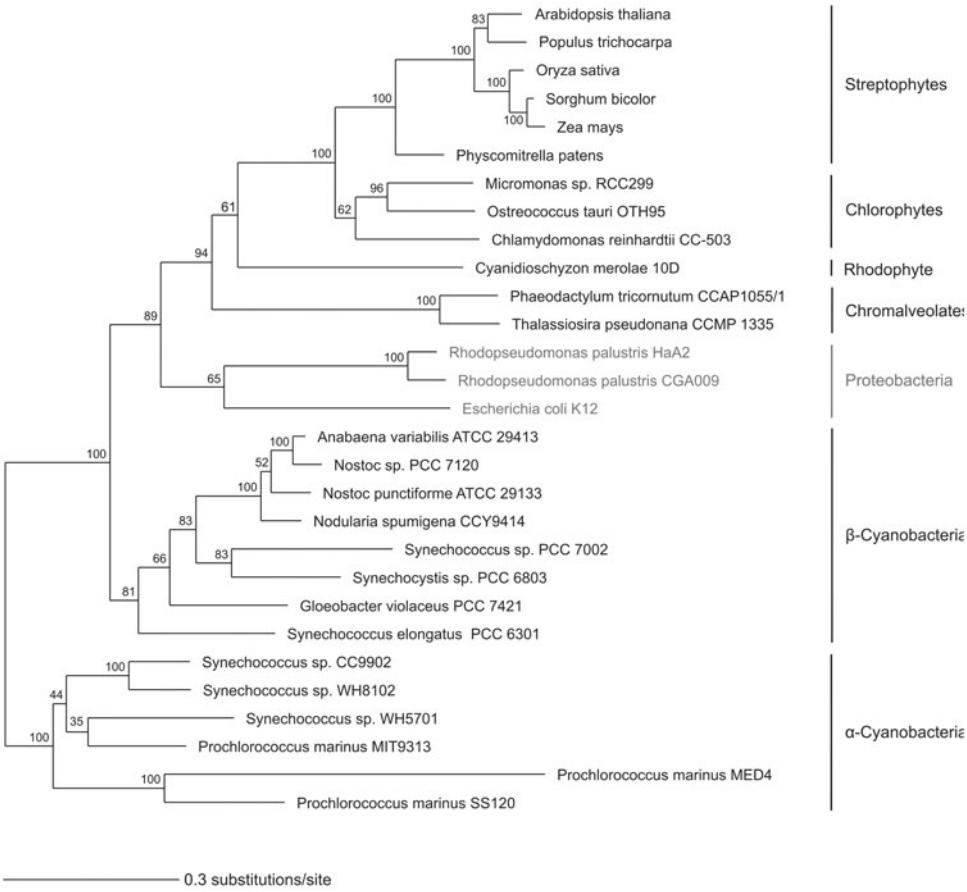

Fig. 6 Maximum likelihood tree of P-proteins of glycine decarboxylase (GDC) based on 29 sequences. *Numbers* at the node indicated bootstrap values (%) for 400 replicates. The distance scale (substitutions per site) is shown in the *bottom left-hand corner*. Further information (protein accession numbers and sequence similarities) is given in Supplemental Tables S1 and S2

study, we included only the two mitochondrial SHMT isoforms with SHMT1 as the major photorespiratory SHMT predominantly expressed in leaves (Voll et al. 2006). The SHMT tree (Fig. 7) positions both isoforms at the same node arguing for a gene duplication and further independent evolution resulting in two different SHMT's in the *Arabidopsis* genome (e.g., Blanc et al. 2003).

The plant and algal SHMT proteins form a monophyletic group with the *Emiliania* protein as the basal branch. The SHMTs of cyanobacteria together with proteobacteria build another monophyletic cluster with the *E. coli* homolog as the most basal node. Therefore, neither the α-proteobacterial nor the cyanobacterial clade stays in a clear sister group relationship with plants and algae. Clarification of the origin of plant and algal SHMTs will require more detailed phylogenetic analyses including SHMTs from fungi and animals. At present, since the photorespiratory SHMT reaction is localized in mitochondria (McClung et al. 2000; Voll et al. 2006) and proteomic studies identified the vast majority of mitochondrial proteins as acquisitions from the α-proteobacterial ancestor of this organelle (Atteia et al. 2009; Esser et al. 2004), an α-proteobacterial origin for the SHMT is most likely.

Plant photorespiratory proteins of unknown origin

Serine:glyoxylate aminotransferase and glutamate:glyoxylate aminotransferase

In plants, two peroxisomal glyoxylate aminotransferases are involved in the photorespiratory pathway (Igarashi et al. 2003; Liepman and Olsen 2001). Both enzymes use glyoxylate as the amino acceptor (Somerville and Ogren 1980). Although these aminotransferases show a broad substrate spectrum, in plants the serine:glyoxylate aminotransferase (SGT) and the glutamate:glyoxylate aminotransferase (GGT) are distinguished by their unique substrate specificities (Liepman and Olsen 2001). The SGT corresponds to At2g13360, while the GGT is encoded by At1g23310 in *Arabidopsis*. In many but not all cyanobacterial genomes, two different groups of putative SGTs are present. Interestingly, chromalveolate and prasinophyte algae seem to lack proteins homologous to SGT. Apparently, these algal groups are completely free of class IV aminotransferases and likely use other aminotransferase. Unfortunately, no functional analysis has been performed with the aminotransferases from cyanobacteria and algae. Moreover, the multi-specific nature of most aminotransferases makes a

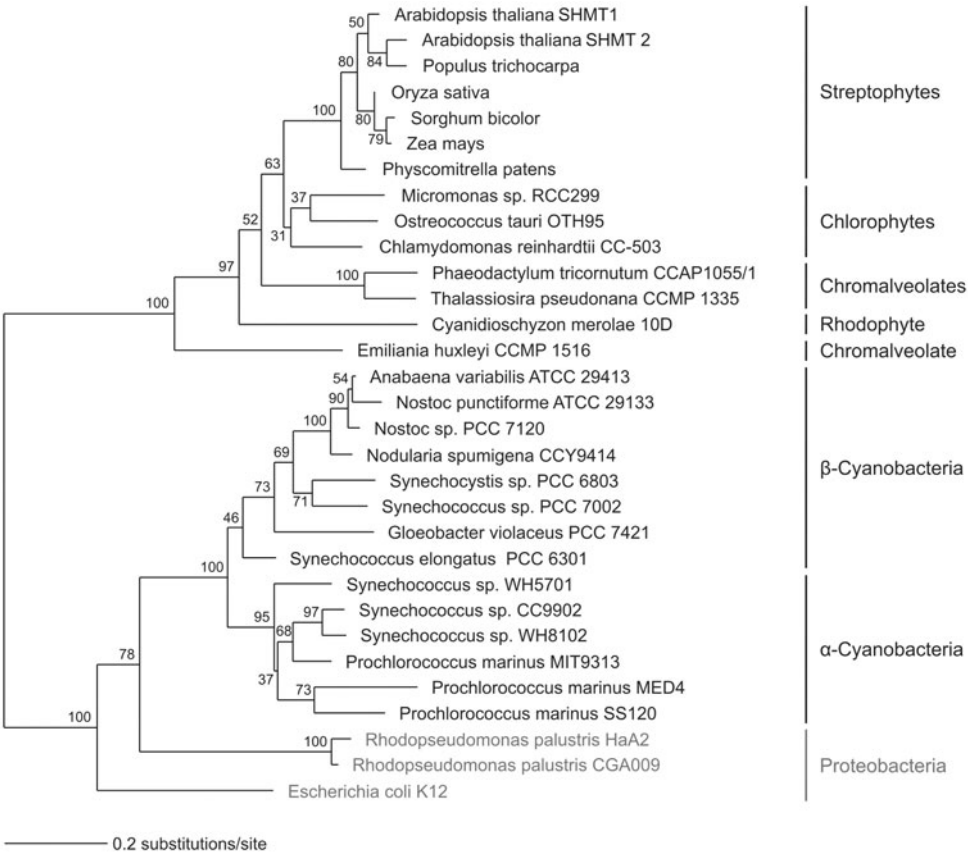

Fig. 7 Maximum likelihood tree of serine hydroxymethyltransferases (SHMT) based on 31 sequences. *Numbers* at the node indicated bootstrap values (%) for 400 replicates. The distance scale (substitutions per site) is shown in the *bottom left-hand corner*. Further information (protein accession numbers and sequence similarities) is given in Supplemental Tables S1 and S2

prediction of the main substrate of putative aminotransferases difficult (Liepman and Olsen 2004). Despite this unclear situation, we tried to also construct a maximum likelihood phylogeny of the two photorespiratory aminotransferases. Both trees seem to indicate a non-cyanobacterial origin (data not shown).

2-phosphoglycolate phosphatase

The 2PG phosphatase (PGP) catalyzes the initial dephosphorylation of 2PG produced in the oxygenase reaction of Rubisco. In *Arabidopsis*, the PGP1 (At5g36700) was identified as the chloroplastidal enzyme for 2PG dephosphorylation in the photorespiratory cycle (Schwarte and Bauwe 2007). Additionally, there exist over 10 proteins with similarities to the PGP1 in the *Arabidopsis* genome. Within cyanobacteria, the identified PGP-like proteins show much variability in their primary structure and only few conserved positions. Hence, the detection of homologous proteins among cyanobacteria through BlastP searches is difficult. It appears that the cyanobacterial PGPs are phylogenetically unrelated to the plant PGP1; however, functional data for putative cyanobacterial PGPs are not yet available and must be acquired before a sound hypothesis on the origin of photorespiratory PGPs in higher plants can be developed.

Conclusion

Cyanobacteria were the first lineage in which oxygenic photosynthesis evolved and were hence the first organisms in which Rubisco was directly exposed to oxygen (Mulkidjanian et al. 2006). Particularly in habitats with decreased O_2-diffusion (for example, stromatolites), already the ancient cyanobacteria might be characterized by an O_2-rich cytoplasm, which stimulated the oxygenase reaction of Rubisco leading to 2PG synthesis. Our findings suggest that the ancient cyanobacteria as well as other bacteria were already equipped with enzymes for the metabolism of glycolate and related small organic acids as a carbon source. We propose that this ancient catabolic glycolate conversion was also used as the starting point to evolve the photorespiratory 2PG cycle for the removal and recycling of this toxic intermediate originating from the Rubisco oxygenase activity in cyanobacterial cells. Only much later, when the atmospheric concentration of O_2 increased and that of CO_2 decreased, CCMs were invented

Table 1 Enzymes involved in 2-phosphoglycolate metabolism in *Arabidopsis thaliana*, their localization, and bacterial origin

Enzyme	*A. thaliana*	Localization in plants	Origin
Phosphoglycolate phosphatase	At5g36790	Chloroplast	?
Glycolate oxidase	At3g14420	Peroxisome	Cyanobacteria
Serine:glyoxylate aminotransferase	At2g13360	Peroxisome	?
Glutamate:glyoxylate aminotransferase	At1g23310	Peroxisome	?
Glycine decarboxylase			
T-protein	At1g11860	Mitochondrion	α-Proteobacteria
P-protein	At4g33010	Mitochondrion	α-Proteobacteria
H-protein	At2g35370	Mitochondrion	?
L-protein	At3g17240	Mitochondrion	α-Proteobacteria
Serine hydroxymethyltransferase	At4g37930	Mitochondrion	α-Proteobacteria
Hydroxypyruvate reductase 1	At1g68010	Peroxisome	Cyanobacteria
Glycerate kinase	At1g80380	Chloroplast	Cyanobacteria

to concentrate CO_2 in Rubisco-containing carboxysomes. Indeed, current studies suggest that CCMs evolved in Carboniferous period only about 360–290 million years ago (Raven et al. 2008).

Comparative analysis of cyanobacterial genomes including the picoplanktonic *Prochlorococcus* and *Synechococcus* strains point to the presence of the photorespiratory 2PG cycle in all cyanobacteria (Eisenhut et al. 2008; and data presented here). Our analysis of 11 photorespiratory proteins in cyanobacteria, algae, and plants revealed an almost even divergence (summarized in Table 1; Fig. 1) from cyanobacteria (ancestor of chloroplasts) and proteobacteria (ancestor of mitochondria). Interestingly, proteins in the chloroplast (except the unclear situation with PGP) and the cooperating enzymes in the peroxisome (GOX and HPR) likely originate from cyanobacteria, while the enzymatic steps for glycine-into-serine conversion in mitochondria and possibly also the associated aminotransferases in the peroxisome were inherited from the proteobacterial endosymbiont. Such a mixture of bacterial enzymes in the modern plant metabolism has been also found for other pathways (Schnarrenberger and Martin 2002). Moreover, our analyses also support the view that an N_2-fixing cyanobacterium was used as primary endosymbiont since only in these cyanobacteria plant-like GOX and HPR proteins still exist. Notably, in contrast to most genes of the photosynthetic machinery, nearly all enzymes involved in 2PG metabolism coexist in cyanobacterial and other eubacterial genomes like *Rhodopseudomonas palustris* (Mulkidjanian et al. 2006). These relationships point to the existence of an ancient glycolate metabolism in the common ancestor of phototrophic bacteria and chloroplasts. During the evolution of higher plants, these glycolate converting enzymes were recruited to generate the modern higher plant photorespiration 2PG cycle, which also included the compartimentation of single enzymatic steps.

Acknowledgments The work was supported by grants of the DFG (Deutsche Forschungsgemeinschaft) in the frame of Forschergruppe FOR 1186-Promics.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Ashida H, Saito Y, Kojima C, Kobayashi K, Ogasawara N, Yokota A (2003) A functional link between RuBisCO-like protein of *Bacillus* and photosynthetic RuBisCO. Science 302:286–290

Atteia A, Adrait A, Brugiere S, Tardif M, van Lis R, Deusch O, Dagan T, Kuhn L, Gontero B, Martin W, Garin J, Joyard J, Rolland N (2009) A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the α-proteobacterial mitochondrial ancestor. Mol Biol Evol 26:1533–1548

Badger MR, Hanson D, Price GD (2002) Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. Funct Plant Biol 29:161–173

Badger MR, Price GD, Long BM, Woodger FJ (2006) The environmental plasticity and ecological genomics of the cyanobacterial CO_2 concentrating mechanism. J Exp Bot 57:249–265

Bari R, Kebeish R, Kalamajka R, Rademacher T, Peterhansel C (2004) A glycolate dehydrogenase in the mitochondria of *Arabidopsis thaliana*. J Exp Bot 55:623–630

Bartsch O, Hagemann M, Bauwe H (2008) Only plant-type (GLYK) glycerate kinases produce D-glycerate 3-phosphate. FEBS Lett 582:3025–3028

Bauwe H, Hagemann M, Fernie AR (2010) Photorespiration: players, partners and origin. Trends Plant Sci 15:330–336

Beezley BB, Gruber PJ, Frederick SE (1976) Cytochemical localization of glycolate dehydrogenase in mitochondria of *Chlamydomonas*. Plant Physiol 58:315–319

Blanc G, Hokamp K, Wolfe KH (2003) A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. Genome Res 13:137–144

Boldt R, Edner C, Kolukisaoglu U, Hagemann M, Weckwerth W, Wienkoop S, Morgenthal K, Bauwe H (2005) D-Glycerate 3-kinase, the last unknown enzyme in the photorespiratory cycle in *Arabidopsis*, belongs to a novel kinase family. Plant Cell 17:2413–2420

Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 17:540–552

Chevenet F, Brun C, Banuls AL, Jacq B, Christen R (2006) TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinform 7:439–448

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36:465–469

Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, Martin W, Dagan T (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. Mol Biol Evol 25:748–761

Eisenhut M, Ruth W, Haimovich M, Bauwe H, Kaplan A, Hagemann M (2008) The photorespiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. Proc Natl Acad Sci USA 105:17199–17204

Engqvist M, Drincovich MF, Flügge U-I, Maurino VG (2009) Two d-2-hydroxy-acid dehydrogenases in *Arabidopsis thaliana* with catalytic capacities to participate in the last reactions of the methylglyoxal and β-oxidation pathways. J Biol Chem 284:25026–25037

Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastiani F, Gelius-Dietrich G, Henze K, Kretschmann E, Richly E, Leister D, Bryant D, Steel MA, Lockhart PJ, Penny D, Martin W (2004) A genome phylogeny for mitochondria among α-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. Mol Biol Evol 21:1643–1660

Foyer CH, Bloom AJ, Queval G, Noctor G (2009) Photorespiratory metabolism: Genes, mutants, energetics, and redox signaling. Annu Rev Plant Biol 60:455–484

Frommolt R, Werner S, Paulsen H, Goss R, Wilhelm C, Zauner S, Maier UG, Grossman AR, Bhattacharya D, Lohr M (2008) Ancient recruitment by chromists of green algal genes encoding enzymes for carotenoid biosynthesis. Mol Biol Evol 25: 2653–2667

Giordano M, Beardall J, Raven JA (2005) CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. Annu Rev Plant Biol 56:99–131

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704

Gupta RS, Mathews DW (2010) Signature proteins for the major clades of cyanobacteria. BMC Evol Biol 10:24–44

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98

Igarashi D, Miwa T, Seki M, Kobayashi M, Kato T, Tabata S, Shinozaki K, Ohsumi C (2003) Identification of photorespiratory glutamate:glyoxylate aminotransferase (GGAT) gene in *Arabidopsis*. Plant J 33:975–987

Kaplan A, Reinhold L (1999) CO_2 concentrating mechanisms in photosynthetic microorganism. Annu Rev Plant Physiol Plant Mol Biol 50:539–570

Keeling PJ (2010) The endosymbiotic origin, diversification and fate of plastids. Philos Trans R Soc Lond [Biol] 365:729–748

Kroth PG, Chiovitti A, Gruber A, Martin-Jezequel V, Mock T, Parker MS, Stanley MS, Kaplan A, Caron L, Weber T, Maheswari U, Armbrust EV, Bowler C (2008) A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. PLos One 3:e1426

Le Corguille G, Pearson G, Valente M, Viegas C, Gschloessl B, Corre E, Bailly X, Peters AF, Jubin C, Vacherie B, Cock JM, Leblanc C (2009) Plastid genomes of two brown algae, *Ectocarpus siliculosus* and *Fucus vesiculosus*: further insights on the evolution of red-algal derived plastids. BMC Evol Biol 9:253

Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol 25:1307–1320

Liepman AH, Olsen LJ (2001) Peroxisomal alanine: glyoxylate aminotransferase (AGT1) is a photorespiratory enzyme with multiple substrates in *Arabidopsis thaliana*. Plant J 25:487–498

Liepman AH, Olsen LI (2004) Genomic analysis of aminotransferases in *Arabidopsis thaliana*. Crit Rev Plant Sci 23:73–89

Lutziger I, Oliver DJ (2000) Molecular evidence of a unique lipoamide dehydrogenase in plastids: analysis of plastidic lipoamide dehydrogenase from *Arabidopsis thaliana*. FEBS Lett 484:12–16

Macheroux P, Mulrooney SB, Williams CH, Massey V (1992) Direct expression of active spinach glycolate oxidase in *Escherichia coli*. Biochim Biophys Acta 1132:11–16

Maruyama S, Matsuzaki M, Misawa K, Nozaki H (2009) Cyanobacterial contribution to the genomes of the plastid-lacking protists. BMC Evol Biol 9:197

Matsuzaki M, Misumi O, Shin I, Maruyama S, Takahara M, Miyagishima SY, Mori T, Nishida K, Yagisawa F, Nishida K, Yoshida Y, Nishimura Y, Nakao S, Kobayashi T, Momoyama Y, Higashiyama T, Minoda A, Sano M, Nomoto H, Oishi K, Hayashi H, Ohta F, Nishizaka S, Haga S, Miura S, Morishita T, Kabeya Y, Terasawa K, Suzuki Y, Ishii Y, Asakawa S, Takano H, Ohta N, Kuroiwa H, Tanaka K, Shimizu N, Sugano S, Sato N, Nozaki H, Ogasawara N, Kohara Y, Kuroiwa T (2004) Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae* 10D. Nature 428:653–657

McClung CR, Hsu M, Painter JE, Gagne JM, Karlsberg SD, Salome PA (2000) Integrated temporal regulation of the photorespiratory pathway. Circadian regulation of two *Arabidopsis* genes encoding serine hydroxymethyltransferase. Plant Physiol 123:381–392

Mereschkowsky C (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. Biol Centralbl 25:593–604

Moustafa A, Beszteri B, Maier UG, Bowler C, Valentin K, Bhattacharya D (2009) Genomic footprints of a cryptic plastid endosymbiosis in diatoms. Science 324:1724–1726

Mulkidjanian AY, Koonin EV, Makarova KS, Mekhedov SL, Sorokin A, Wolf YI, Dufresne A, Partensky F, Burd H, Kaznadzey D, Haselkorn R, Galperin MY (2006) The cyanobacterial genome core and the origin of photosynthesis. Proc Natl Acad Sci USA 103:13126–13131

Nelissen B, Vandepeer Y, Wilmotte A, Dewachter R (1995) An early origin of plastids within the cyanobacterial divergence is suggested by evolutionary trees based on complete 16S ribosomal-RNA sequences. Mol Biol Evol 12:1166–1173

Nelson EB, Tolbert NE (1970) Glycolate dehydrogenase in green algae. Arch Biochem Biophys 141:102–110

Parker MS, Mock T, Armbrust EV (2008) Genomic insights into marine microalgae. Annu Rev Genet 42:619–645

Raven JA, Cockell CS, La Rocha CL (2008) The evolution of inorganic carbon concentrating mechanisms in photosynthesis. Philos Trans R Soc Lond [Biol] 363:2641–2650

Schnarrenberger C, Martin W (2002) Evolution of the enzymes of the citric acid cycle and the glyoxylate cycle of higher plants. A case study of endosymbiotic gene transfer. Eur J Biochem 269: 868–883

Schwarte S, Bauwe H (2007) Identification of the photorespiratory 2-phosphoglycolate phosphatase, PGLP1, in *Arabidopsis*. Plant Physiol 144:1580–1586

Somerville CR (2001) An early *Arabidopsis* demonstration. Resolving a few issues concerning photorespiration. Plant Physiol 125: 20–24

Somerville CR, Ogren WL (1980) Photorespiration mutants of *Arabidopsis thaliana* deficient in serine-glyoxylate aminotransferase activity. Proc Natl Acad Sci USA 77:2684–2687

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W—improving the sensitivity of progressive multiple sequence

alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22: 4673–4680

Timm S, Nunes-Nesi A, Pamik T, Morgenthal K, Wienkoop S, Keerberg O, Weckwerth W, Kleczkowski LA, Fernie AR, Bauwe H (2008) A cytosolic pathway for the conversion of hydroxypyruvate to glycerate during photorespiration in *Arabidopsis*. Plant Cell 20:2848–2859

Tolbert NE (1997) The C-2 oxidative photosynthetic carbon cycle. Annu Rev Plant Physiol Plant Mol Biol 48:1–23

Voll LM, Jamai A, Renne P, Voll H, McClung CR, Weber APM (2006) The photorespiratory *Arabidopsis shm1* mutant is deficient in SHM1. Plant Physiol 140:59–66

Winkler U, Stabenau H (1995) Isolation and characterization of peroxisomes from diatoms. Planta 195:403–407

Worden AZ, Lee JH, Mock T, Rouze P, Simmons MP, Aerts AL, Allen AE, Cuvelier ML, Derelle E, Everett MV, Foulon E, Grimwood J, Gundlach H, Henrissat B, Napoli C, McDonald SM, Parker MS, Rombauts S, Salamov A, Von Dassow P, Badger JH, Coutinho PM, Demir E, Dubchak I, Gentemann C, Eikrem W, Gready JE, John U, Lanier W, Lindquist EA, Lucas S, Mayer KFX, Moreau H, Not F, Otillar R, Panaud O, Pangilinan J, Paulsen I, Piegu B, Poliakov A, Robbens S, Schmutz J, Toulza E, Wyss T, Zelensky A, Zhou K, Armbrust EV, Bhattacharya D, Goodenough UW, Van de Peer Y, Grigoriev IV (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. Science 324:268–272

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

Zelitch I, Schultes NP, Peterson RB, Brown P, Brutnell TP (2009) High glycolate oxidase activity is required for survival of maize in normal air. Plant Physiol 149:195–204