Chlorophyll Biosynthesis Gene Evolution Indicates Photosystem Gene Duplication, Not Photosystem Merger, at the Origin of Oxygenic Photosynthesis

Filipa L. Sousa[1,*,‡], Liat Shavit-Grievink[1,4,‡], John F. Allen[2,3], and William F. Martin[1]

[1]Institute of Molecular Evolution, University of Düsseldorf, Düsseldorf, Germany

[2]School of Biological and Chemical Sciences, Queen Mary, University of London, London, United Kingdom

[3]Department of Genetics, Evolution and Environment, University College London, London, United Kingdom

[4]Present address: Edmond and Lily Safra Center for Brain Sciences, Hebrew University of Jerusalem, Jerusalem, Israel

‡These authors contributed equally to this work.

*Corresponding author: E-mail: Filipa.Sousa@hhu.de.

Accepted: December 17, 2012

Abstract

An open question regarding the evolution of photosynthesis is how cyanobacteria came to possess the two reaction center (RC) types, Type I reaction center (RCI) and Type II reaction center (RCII). The two main competing theories in the foreground of current thinking on this issue are that either 1) RCI and RCII are related via lineage divergence among anoxygenic photosynthetic bacteria and became merged in cyanobacteria via an event of large-scale lateral gene transfer (also called "fusion" theories) or 2) the two RC types are related via gene duplication in an ancestral, anoxygenic but protocyanobacterial phototroph that possessed both RC types before making the transition to using water as an electron donor. To distinguish between these possibilities, we studied the evolution of the core (bacterio)chlorophyll biosynthetic pathway from protoporphyrin IX (Proto IX) up to (bacterio)chlorophyllide *a*. The results show no dichotomy of chlorophyll biosynthesis genes into RCI- and RCII-specific chlorophyll biosynthetic clades, thereby excluding models of fusion at the origin of cyanobacteria and supporting the selective-loss hypothesis. By considering the cofactor demands of the pathway and the source genes from which several steps in chlorophyll biosynthesis are derived, we infer that the cell that first synthesized chlorophyll was a cobalamin-dependent, heme-synthesizing, diazotrophic anaerobe.

Key words: redox-switch, selective-loss, manganese, cyanobacteria, fusion.

Introduction

The origin of oxygenic photosynthesis introduced a new high-potential electron acceptor into microbial ecosystems (Holland 2006) and enzymatic reaction mechanisms (Raymond et al. 2002), marking the onset of pivotal changes in geochemical cycles and biochemical pathways. Oxygen first began to accumulate in the atmosphere approximately 2.4 billion years ago, and its subsequent accumulation in the oceans was slower, such that full oxic conditions were only reached approximately 635–580 Ma (Arnold et al. 2004; Lyons 2007; Canfield et al. 2008; Scott et al. 2008; Lyons et al. 2009; Sahoo et al. 2012). This "new chemistry" had far-reaching impact on the evolutionary process.

Chlorophyll-based photosynthesis arose among eubacteria, where it is currently found among six phyla. Chlorobia (green sulfur bacteria [GSB]), firmicutes (heliobacteria), and acidobacteria have anoxygenic Type I reaction centers (RCIs), whereas chloroflexi (green nonsulfur bacteria, GNSB of filamentous anoxygenic phototrophs, FAPs) and some proteobacterial organisms (purple nonsulfur bacteria and purple sulfur oxidizing bacteria [Dahl et al. 2005]) perform anoxygenic photosynthesis with Type II reaction centers (RCIIs) (Xiong and Bauer 2002a; Bryant et al. 2007; Maqueo Chew and Bryant 2007; Blankenship 2010). Only cyanobacteria and, via endosymbiosis, photosynthetic eukaryotes (Margulis and Bermudes 1985) perform oxygenic photosynthesis, having both the Photosystem I (PSI) and the Photosystem II (PSII) complexes that are homologous to RCI and RCII, respectively (Michel and Deisenhofer 1988; Hauska et al. 2001; Neerken and Amesz 2001). Although at the sequence level there is

almost no detectable similarity between RCI and RCII, their structural and cofactor arrangements are unquestionably homologous and clearly indicate common ancestry (Schubert et al. 1998; Barber et al. 2000; Sadekar et al. 2006).

Chlorophyll (Chl) and bacteriochlorophyll (Bch) are required for oxygenic and anoxygenic photosynthesis, respectively, where they serve two essential functions. As light-harvesting antennae, they act as photon funnels, absorbing light and channeling its energy to reaction centers (RCs), where chlorophylls perform their second function: photochemical charge separation to create strong oxidants and reductants, sending low-potential electrons through the electron transport chain (ETC). In anoxygenic phototrophs, bacteriochlorophylls are the main pigments. In contrast, chloroplasts and cyanobacteria possess only chlorophylls, which can, however, also be residually present in the RC of some anoxygenic photosynthetic bacteria (Kobayashi et al. 2000; Hohmann-Marriott and Blankenship 2011; Sarrou et al. 2012; Tsukatani et al. 2012). With exception of chlorophyll *c*, chemically, both chlorophylls and bacteriochlorophylls are chlorins; reduced magnesium-containing cyclic tetrapyrroles with an additional fifth isocyclic ring. Their main differences concern the level of unsaturation affecting the system of conjugated double bonds. Chlorophyll has a single bond between carbons C17 and C18 (IUPAC numbering), and bacteriochlorophyll has an additional C7–C8 single bond (Maqueo Chew and Bryant 2007; Niedzwiedzki and Blankenship 2010). There are 11 major types of bacteriochlorophyll and chlorophylls (Hohmann-Marriott and Blankenship 2011) differing with respect to the substituents of the tetrapyrrole ring, hence their differing absorption properties, which allow organisms to specialize toward different spectral niches, most notably as a function of depth in the water column (Glaeser et al. 2002; Manske et al. 2005; Gomez Maqueo Chew et al. 2007; Kiang, Siefert, et al. 2007; Kiang et al. 2007; Stomp et al. 2007; Chen et al. 2010).

The enzymes involved in chlorophyll metabolism are the only set of photosynthesis-related proteins common to all phototrophs (Mulkidjanian et al. 2006), and thoughts on chlorophyll evolution have a long history. Central to the topic is the Granick hypothesis (Granick 1965), which posits that the evolution of the chlorophyll biosynthetic pathway followed the sequential inventions of new enzymes to generate more stable products. This premise has been widely used to study the evolution of the chlorophyll pathway, in particular as a proxy for the evolution of photosynthesis itself (Olson and Pierson 1987a, 1987b; Xiong and Bauer 2002a, 2002b; Gupta 2012). However, other hypotheses are still discussed, and the issue is debated (Lockhart et al. 1996; Blankenship 2001). In 2000, Xiong et al. studied phylogenies for 9 of the 17 enzymes then known to be involved in the (bacterio)chlorophyll pathway. They suggested that (bacterio)chlorophylls first arose within purple bacteria (proteobacteria) and that the pathway's emergence involved the recruitment and duplication of homologous enzymes such as nitrogenase subunits and cobalt chelatase from cobalamin biosynthesis (Xiong et al. 1998; Xiong and Bauer 2002a, 2002b). More recently, and based on phylogenies and sequence signatures of the (B)ChlNBL complex, which is responsible for catalyzing the last step of the (bacterio)chlorophyll core pathway, an origin of chlorophyll in the Gram-positive heliobacteria lineage has been suggested (Gupta 2012).

Here, we address the evolution of the chlorophyll biosynthesis pathway to distinguish between competing theories for the origin of two photosystems in cyanobacteria. The RCs of photosystems I and II are clearly related at the level of three-dimensional structure (Nitschke and Rutherford 1991; Schubert et al. 1998; Baymann et al. 2001; Allen and Puthiyaveetil 2005; Allen et al. 2011; Hohmann-Marriott and Blankenship 2011), and the issue is how they came to reside within the ancestral cyanobacterial genome. The "fusion" (or "merger") hypothesis asserts that the two photosystems diverged during the evolution of anoxygenic photosynthetic lineages and became merged in the founding cyanobacterium via lateral gene transfer (LGT) (Mathis 1990; Meyer 1994; Xiong and Bauer 2002a; Hohmann-Marriott and Blankenship 2011). In contrast, the "duplication" hypothesis asserts that the photosystems diverged within a protocyanobacterial ancestor and subsequently underwent vertical inheritance and export, via LGT, to diverse anoxygenic photosynthetic lineages (Olson and Pierson 1987a, 1987b; Baymann et al. 2001; Olson 2001; Allen 2005; Mulkidjanian et al. 2006). Were the merger hypothesis correct, the genes of chlorophyll biosynthesis in anoxygenic photosynthetic lineages should reflect early lineage splittings and hence the same deep divergence as the photosynthetic RCs. In the case that the two photosystems arose via duplication in a protocyanobacterium, the chlorophyll biosynthetic pathway should not reflect ancient lineage splittings, that is, there should be no deep dichotomy into RCI and RCII-specific chlorophyll biosynthetic pathways types. To distinguish between these possibilities, we have studied the evolution of the core (bacterio)chlorophyll biosynthetic pathway from Proto IX up to (bacterio)chlorophyllide *a*, which is the last chemical intermediate common to all chlorophyll types.

Materials and Methods

Data

The list of organisms was extracted from the GOLD table as accessed on the 16 June 2010 (Bernal et al. 2001). Genomes of photosynthetic prokaryotes were downloaded from RefSeq database release 41 of 9 May 2010 (Pruitt et al. 2005). In addition, the recently sequenced genome of *Candidatus Chloracidobacterium thermophilum* was also included (Garcia Costas et al. 2012). Homologous proteins involved in (bacterio)chlorophyll biosynthesis were identified by Basic

Local Alignment Search Tool (BLAST) (Altschul et al. 1997) within the data set of downloaded genomes (17 eukaryotes and 87 prokaryotes, supplementary table S1, Supplementary Material online) using the proteins shown in supplementary table S2, Supplementary Material online, as queries. The BLAST lists were filtered for *E* values better than 10^{-10} and amino acid identities of ≥25%. Paralogs from the cobalamin pathway or later steps in the bacteriochlorophyll *c* pathway were removed. Because some enzymes of chlorophyll biosynthesis (BchE/ChlE, BciA, and light-dependent NADPH:protochlorophyllide oxidoreductase [LPOR]) belong to protein families with members catalyzing reactions in other pathways, possible paralogs were identified by reBLASTing against the entire National Center for Biotechnology Information (NCBI) refseq database.

Sequence Alignments and Phylogenetic Analysis

Proteins were aligned using MUSCLE (Edgar 2004) using its default parameters. Trees were reconstructed using PhyML 3.0 (Guindon and Gascuel 2003) with the WAG + I + G model and four rate categories. Alignments are available upon request and nonconcatenated phylogenetic trees are given as supplementary figures S1–S12, Supplementary Material online.

Topology Comparisons

Topology testing was performed using the approximately unbiased (AU) test and the Shimodaira–Hasegawa (SH) test (Shimodaira and Hasegawa 1999; Shimodaira 2002). Where justified, initial constrained trees with an imposed separation of (*B*)*Chls* genes from RCI- and RCII-containing organisms were created with an in-house perl script and optimized in PhyML 3.0 (Guindon and Gascuel 2003). The site likelihoods of the constrained and the nonconstrained trees were calculated using PhyML 3.0 and the AU and the SH tests performed using CONSEL v0.1k (Shimodaira and Hasegawa 2001).

Results

Taxonomic Distribution of Photosynthesis and Bacteriochlorophyll Biosynthesis Genes

The taxonomic distribution of the photosynthetic apparatus in a sample of sequenced prokaryotic genomes is presented in figure 1*A*. Photosynthesis is absent among archaebacteria and present in only 6 of the 24 bacterial phyla currently recognized by NCBI taxonomy. Within those six phyla, photosynthesis is, in general, sparsely distributed. All cyanobacterial genomes and, with one exception—the recently sequenced *Ignavibacterium album* bacterium (an early diverging member of the chlorobia phylum) (Iino et al. 2010; Liu et al. 2012), all chlorobial genomes sequenced so far have genes coding for the photosynthetic apparatus. *Ignavibacterium* is a facultative aerobe, lacks photosynthesis genes, and is currently the only nonphototrophic chlorobium to have been cultured (Liu et al. 2012). Nevertheless, the dense distribution of photosynthetic genes in cyanobacteria and chlorobia suggests their presence in the ancestors of these phyla. The same can also be assumed for chloroflexi, at least at the class level. The presence of 1 to 3 chlorophyll biosynthesis genes in *Dehalococcoides* sp. BAV1, *Dehalococcoides* sp. CBDB1, and *Herpetosiphon aurantiacus* probably indicates either LGT or past presence of the pathway and its subsequent loss. In contrast, within proteobacteria and the firmicutes, even at the class level, a sparse distribution of photosynthetic representatives is found, indicating either late acquisition(s) of photosynthesis or many independent losses of it. Because photosynthetic genes are often encoded in superoperons (Alberti et al. 1995; Xiong et al. 1998) or lineage-specific small photosynthetic gene clusters (Xiong et al. 2000), LGT scenarios can, in principle, account for the patchy distribution of photosynthesis among eubacteria. This would also be consistent with the presence of photosynthesis genes, including both photosystems I and II, in the marine phage metagenome (Lindell et al. 2004; Sharon et al. 2009). However, scenarios of vertical inherence have also been discussed, at least for the evolution of RCI-containing organisms (Mix et al. 2005).

Genes involved in chlorophyll biosynthesis were identified in our data set of 104 complete sequenced genomes from photosynthetic organisms (see Materials and Methods). Their taxonomic distribution is presented in figure 1*B*. The enzymes of the core chlorophyll biosynthesis pathway are present in all photosynthetic organisms, regardless of which RC type the organism uses. The few differences that are observed with respect to gene presence or absence concern O_2-dependent and O_2-independent alternative enzymes for some steps (see later). Clearly, a single pathway of chlorophyll biosynthesis was present in the ancestral photosynthetic bacterium.

Mg Chelatase

The first unique intermediate of chlorophyll biosynthetic pathway, Mg-protoporphyrin IX (fig. 2*A*, step 1), is generated by the insertion of Mg^{2+} into protoporphyrin IX. Biochemical and genetic analysis identified a Class I ATP-dependent magnesium chelatase, composed of three subunits BchH, BchI, and BchD, that catalyzes a reaction (Walker and Willows 1997; Willows and Beale 1998; Bollivar 2006) consisting of an activation and a chelation step. In the presence of both ATP and Mg^{2+}, an AAA+ motor complex is assembled from a hexameric (Willows et al. 2004) or heptameric (Reid et al. 2003) BchI ring connected to a hexameric BchD ring (Lundqvist et al. 2010) (the activation step). Proto IX binds to the BchH catalytic subunit, and its transient interaction with the formed AAA+ motor complex leads to the insertion of Mg^{2+} in the tetrapyrrole macrocycle (the chelation step).

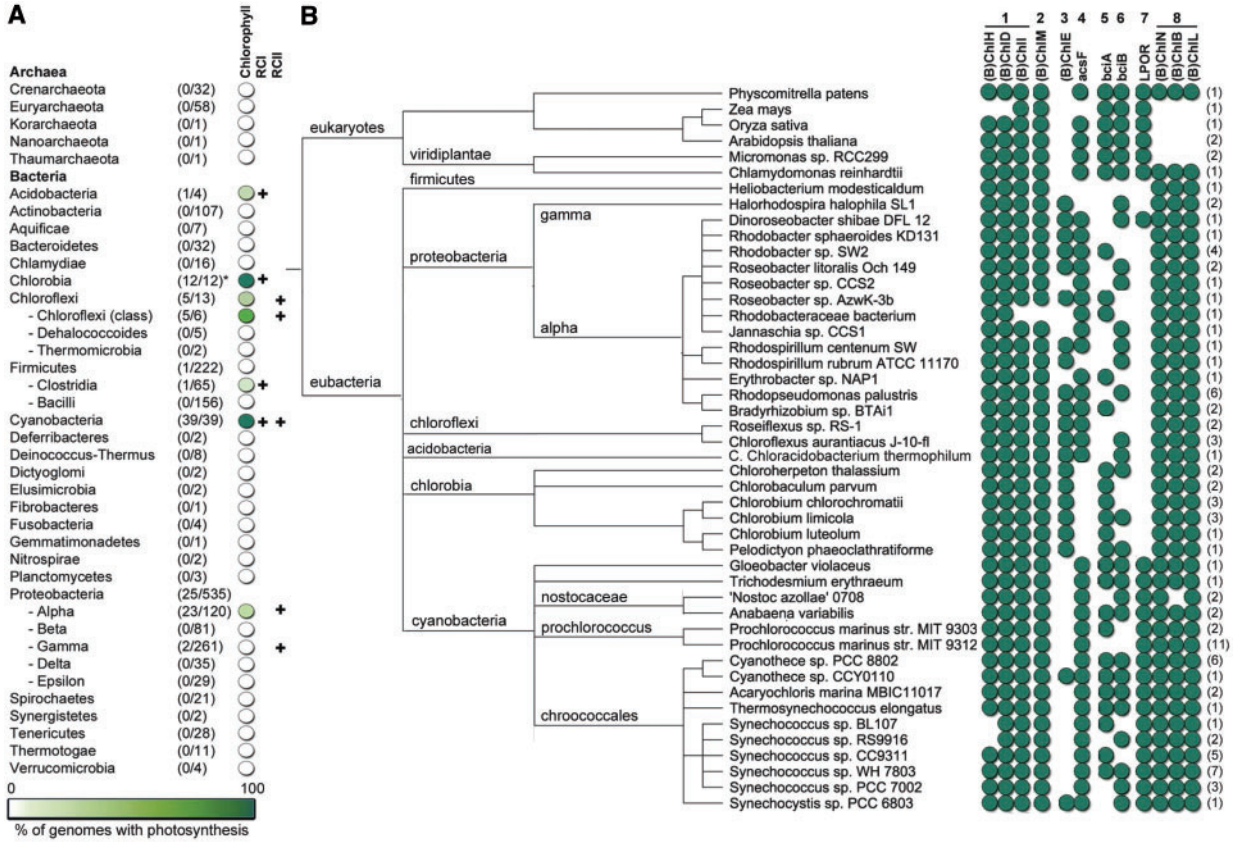

Fig. 1.—Distribution of photosynthesis in prokaryotes. (*A*) Distribution of photosynthesis in completely sequenced genomes of prokaryotic organisms organized according to their phyla and class. The color gradient of each circle reflects the normalization of species within a certain taxon that have genes coding for photosynthetic apparatus versus the total number of organisms from that taxa whose genome has been completely sequenced. Numbers in brackets represent the photosynthetic species within that taxon and the total number of species with completely sequenced genomes. **Ignavibacterium album*, the only known nonphotosynthetic member of chlorobia phylum was not included in our data set. (*B*) Taxonomic distribution (presence or absence) of the different genes involved in the biosynthesis of (bacterio)chlorophyll from Mg-protoporphyrin IX to chlorophyllide *a*. Columns correspond to genes involved in the different steps of chlorophyll photosynthesis. A modified taxonomic tree is represented on the left where leafs are grouped by taxonomy and same pattern of presence and absence of genes. Numbers between brackets represent number of organisms with the same pattern belonging to same taxonomic group: *Physcomitrella patens*; *Zea mays*; *Oryza sativa*; *Arabidopsis thaliana* (2): *A. thaliana* and *Vitis vinifera*; *Micromonas* sp. RCC299 (2): *Micromonas* sp. RCC299 and *Ostreococcus lucimarinus* CCE9901; *Chlamydomonas reinhardtii; Heliobacterium modesticaldum* Ice1; *Halorhodospira halophile* SL1 (2): *H. halophile* SL1 and *Allochromatium vinosum* DSM180; *Dinoroseobacter shibae* DFL 12; *Rhodobacter sphaeroides* KD131; *Rhodobacter* sp. SW2 (4): *Rhodobacter* sp. SW2, *R. sphaeroides* ATCC17025, *R. sphaeroides* ATCC17029, and *R. sphaeroides* 241; *Roseobacter litoralis* Och149 (2): *Ros. litoralis* Och149 and *Ros. denitrificans* OCh114; *Roseobacter* sp. CCS2; *Roseobacter* sp. AzwK-3b; *Rhodobacterales bacterium* HTCC2083; *Jannaschia* sp. CCS1; *Rhodospirillum centenum* SW; *Rho. rubrum* ATCC11170; *Erythrobacter* sp. NAP1; *Rhodopseudomonas palustris* (6): BisB18, HaA2, CGA009, BisA53, BisB5, and TIE-1; *Bradyrhizobium* sp. BTAi1 (2): sp. BTAi1 and sp. ORS278; *Roseiflexus* sp. RS-1 (2): *Roseiflexus* sp. and *Roseiflexus castenholzii* DSM13941; *Chloroflexus aurantiacus* J-10-fl (3): *Chloroflexus aurantiacus* J-10-fl, *Chloroflexus aggregans* DSM9485, and *Chloroflexus* sp. Y-400-fl; *C. Chloracidobacterium thermophilum*: *Candidatus Chloracidobacterium thermophilum; Chloroherpeton thalassium* (2)*: Chloroherpeton thalassium* ATCC35110 and *Prosthecochloris aestuarii* DSM271; *Chlorobaculum parvum* (2): NCIB 8327 and *Chlorobaculum tepidum* TLS; *Chlorobium chlorochromatii* (3): *Chlorobium chlorochromatii* CaD3, *Chlorobium phaeobacteroides* BS1, and *P. vibrioformis* DSM 265; *Chlorobium limicola* (3)*: Chlorobium limicola* DSM245, *Chlorobium phaeobacteroides* DSM266, and *Chlorobium ferrooxidans* DSM13031; *Chlorobium luteolum* DSM273; *Pelodictyon phaeoclathratiforme* BU-1; *Gloeobacter violaceus* PCC7421; *Trichodesmium erythraeum* IMS101; "*Nostoc azollae*" 0708 (2): "*Nostoc azollae*" 0708 and *Nostoc punctiforme* PCC73102; *Anabaena variabilis* (2): ATCC29413 and *Nostoc* sp. PCC7120; *Prochlorococcus marinus* str. MIT9303 (2): str. MIT9303 and MIT 9313; *Prochlorococcus marinus* str. MIT9312 (11): str. MIT9312, subsp. pastoris. str. CCMP1986, subsp. pastoris. str. CCMP1375, str. NATL1A, str. NATL2A str. MIT9202, str. MIT9215, str. MIT9211, str. MIT9515, str. AS9601, and str. MIT9301; *Cyanothece* sp. PCC8802 (6): sp. PCC8802, sp. PCC7424, sp. PCC7425, sp. ATCC51142, sp. PCC7822, and sp. PCC8801, *Cyanothece* sp. CCY0110; *Acaryochloris marina* MBIC11017 (2): *Aca. marina* MBIC11017 and *Microcystis aeruginosa* NIES843; *Thermosynechococcus elongatus* BP-1; *Synechococcus* sp. BL107; *Synechococcus* sp. RS9916 (2): sp. RS9916 and sp. WH7805; *Synechococcus* sp. CC9311 (5): sp. CC9311, sp. WH8109, sp. WH8102, sp. CC9902, and sp. CC9605; *Synechococcus* sp. WH7803(7): sp. WH7803, sp. JA-3-3Ab, *elongatus* PCC7942, *elongatus* PCC6301, sp. 2-3Ba2-13, sp. PCC7335, and sp. WH5701; *Synechococcus* sp. PCC7002 (3): sp. PCC7002, sp. RS9917, and sp. RCC307; and *Synechocystis* sp. PCC6803.

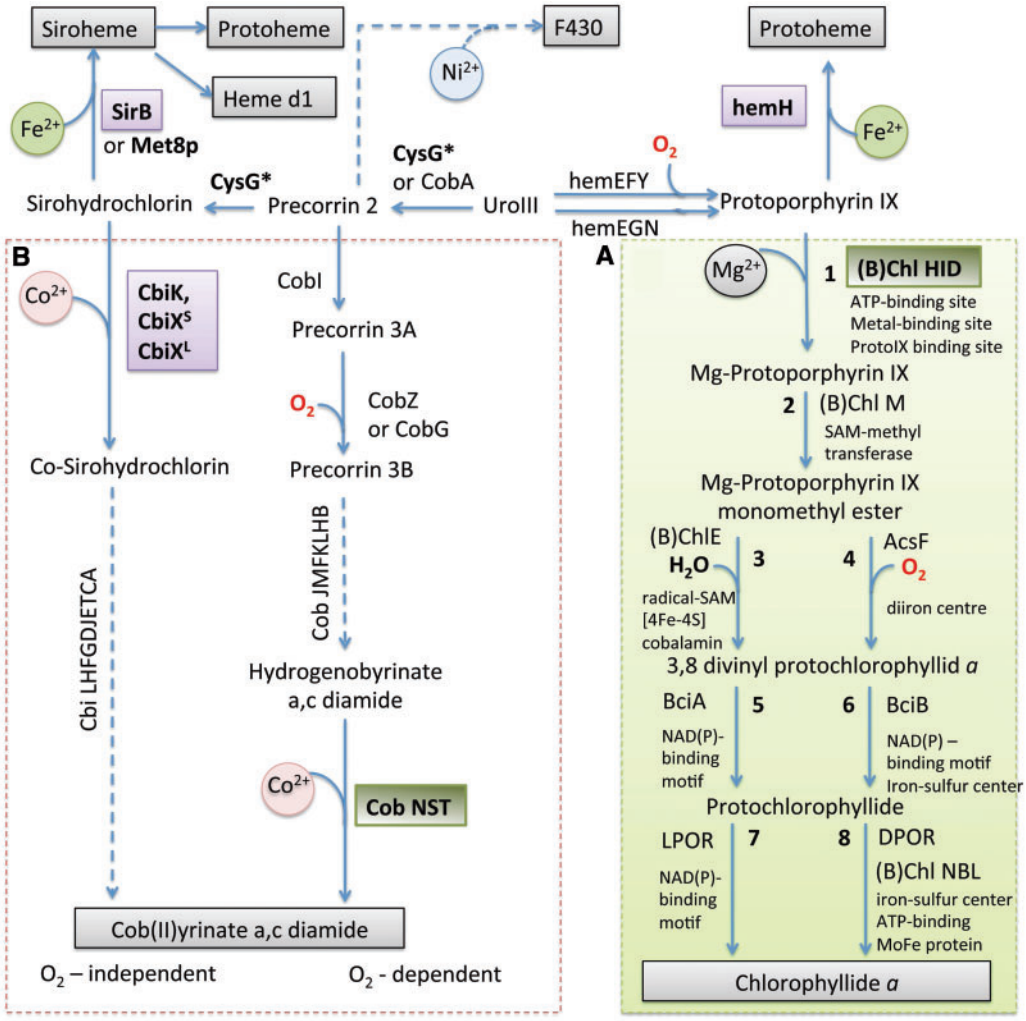

Fig. 2.—Schematic representation of the tetrapyrrole biosynthesis pathway. Based on information from Martens et al. (2002), McGoldrick et al. (2005), Heinemann et al. (2008), Storbeck et al. (2010) and Bali et al. (2011). The different metal chelatases are represented in bold and colored according to their sequence and/or structural similarity. The numbers correspond to the different steps of the common core of (bacterio)chlorophyll pathway. UroIII, the last common intermediate of all tetrapyrroles, can either be methylated by CysG (or CobA) to form precorrin-2 or suffer decarboxylation and oxidation to form Protoporphyrin IX via the O_2-dependent route (HemE, HemF, HemY) or the O_2-independent route (HemE, HemG, HemN). Precorrin-2 is the branching point for F430 (a Ni^{2+}-containing tetrapyrrole), the cobalamin O_2-dependent route, and Sirohydrochlorin. Iron can be inserted into Sirohydrochlorin by SirB or by multifunctional oxidases such as Met8p or CysG to form siroheme. Siroheme can be further modified to form heme d1 or heme, according to the alternative archaeal heme synthesis pathway. Protoporphyrin IX is the branch point to hemes and chlorophylls. Iron is inserted into protoporphyrin IX (classical pathway). (*A*) Specific (bacterio)chlorophyll steps: the incorporation of Mg^{2+} into protoporphyrin IX occurs by the action of (B)ChlHID complex (step 1), resulting in the formation of Mg-protoporphyrin. A methyl transferase (B)ChlM subsequently converts Mg-protoporphyrin into Mg-protoporphyrin IX monomethyl ester (step 2), which is the substrate for a cyclase reaction, involving water (B)ChlE (step 3) or molecular oxygen (AcsF, step 4). The cyclase forms divinyl protochlorophyllide *a*, an intermediate that contains a fifth ring (ring E), characteristic of all (bacterio)chlorophylls. The next reaction consists in the reduction of the 8-vinyl group by an 8-vinyl reductase, BciA (step 5) or BciB (step 6). Finally, the reduction of the IV pyrrole ring of protochlorophyllide to form chlorophyllide is performed by the action of the light-dependent LPOR (step 7) or the light-independent (DPOR) complex formed by (B)Chl LNB. (*B*) Alternative pathways for cobalamin synthesis. In the O_2-independent cobalamin pathway, cobalt is inserted into Sirohydrochlorin to form the Co-Sirohydrochlorin intermediate, via one of the three homologous chelatases CibK, $CibX^S$, or $CibX^L$, homologs to the SirB chelatase and the HemH chelatase from the classical heme pathway. Co-Sirohydrochlorin is then sequential converted into Cob(II)yrinate a,c diamide by the action of 10 enzymes (Cbi LHFGDJETCA). In the O_2-dependent pathway, CobI performs the methylation of precorrin-2 originating precorrin-3A that will be oxidized by one of the two O_2-dependente enzymes (CobZ or CobG) originating precorrin-3B. A series of reactions involving *cob JMFKLHB* genes occur until hydrogenobyrinate a,c diamide is formed. At this stage, Co^{2+} ion is inserted into the ring by the CobNST (homologous to BChl HID chelatase from the (bacterio)chlorophyll pathway) forming Cob(II)yrinate a,c diamide, the merger point between the two cobalamin pathways.

The Mg-chelatase complex has both sequence and structural homology with the Class I cobalt chelatase of the O_2-dependent cobalamin biosynthetic pathway (Lundqvist et al. 2009). The CobN, CobS, and CobT subunits of the trimeric cobalt chelatase are homologous with the BchH/ChlH, BchI/ChlI, and BchD/ChlD subunits of magnesium chelatase, respectively (Fodje et al. 2001; Lundqvist et al. 2009). This homology has been interpreted as reflecting ancient duplication and divergence (Lundqvist et al. 2009). Although there is a broad taxonomic distribution of the *CobN* gene among cobalamin-dependent organisms, a much narrower distribution is observed for the genes that compose the AAA+ motor complex (*CobS* and *CobT*) (Rodionov et al. 2003).

The Catalytic Subunit BchH/ChlH

Chlorobaculum tepidum, similar to most members of GSB, has three paralogous genes for this subunit, named *BchS*, *BchH,* and *BchT*. Single- and double-mutant experiments (Gomez Maqueo Chew et al. 2009) showed that strains with only the *BchH* or the *BchS* gene retained sufficient Mg-chelatase activity to be viable, but its activity was maximal if the *BchS* gene was present together with *BchH*. This was also confirmed by biochemical characterization of the recombinant enzymes (Johnson and Schmidt-Dannert 2008). However, in all mutagenesis experiments, a decrease in the production of bacteriochlorophyll *c*, the main photosynthetic pigment of GSB, was detected. This suggests that the different isoforms function in end-product regulation and/or substrate channeling of the Bch *c* intermediates (Johnson and Schmidt-Dannert 2008; Gomez Maqueo Chew et al. 2009).

The existence of *BchH/ChlH* paralogs is not unique to the GSB. All chloroflexi have at least one additional gene coding this subunit (Gomez Maqueo Chew et al. 2009), and the same is also true for some cyanobacterial species (Lohr et al. 2005). Isoforms of the gene also occur in some eukaryotes (Lohr et al. 2005) although, in this case, different duplication events are possibly at their origin.

The schematic diagram summarizing the BchH/ChlH phylogenetic tree is presented in figure 3*A*. A clear dichotomy between the different subunit isoforms can be observed. The upper clade contain sequences similar to the *BchH* gene of *Chlorobaculum tepidum* TLS. With the exception of *Chlorobium chlorochromatii*, all organisms from the seven phyla considered contain at least one copy the *BchH/ChlH* gene. In this group, three distinct groups are retrieved: 1) chloroflexales and chlorobia, 2) acidobacteria and proteobacteria, and 3) heliobacteria and cyanobacteria with photosynthetic eukaryotes branching as basal members. Recent events of gene duplication can also be observed in the chloroflexales ancestor and at the species level of three land plants.

Notably, there is no relationship between the type of RC and the organization of the sequences within the tree, indicating a single origin of this pathway. Moreover, a close relationship between chlorobia and chloroflexi sequences is observed, a relationship found in most of our phylogenetic analysis. Based inter alia on the observation that fundamental aspects of core carbon and energy metabolism in chlorobia and chloroflexi are different, it has been suggested that lateral transfer of genes underlying chlorosome structure and chlorophyll biosynthesis might have occurred between these groups (Raymond et al. 2002). However, the close phylogenetic proximity of chlorophyll biosynthesis genes among organisms that contain different RC types is not confined to these two groups.

The lower clade of the tree contains sequences similar to the *Chlorobaculum tepidum* TLS *BchS* and *BchT* genes, belonging to chlorobia, chloroflexi, cyanobacteria, and green algae. In this clade, two distinct groups are retrieved, one comprising the cyanobacterial and plastid sequences and the other with chloroflexi BchH2 and chlorobial BchS and BchT sequences. A duplication event in the ancestor of the chlorobia appears to be the likely origin of the two isoforms.

The role of these subunits has been proposed to be either end-product regulation and/or substrate channeling of the Bch *c* intermediates (Johnson and Schmidt-Dannert 2008; Gomez Maqueo Chew et al. 2009). Both chlorobia and chloroflexi have Bch *c* pigments, but the cyanobacteria synthesize only chlorophyll *a*, raising the question of what the role of these cyanobacterial homologs might be. The *Gloeobacter violaceus* ChlH2 sequence alone falls outside these two main clades; furthermore, it is more similar to homologs in nonphotosynthetic lineages (euryarchaeotes and firmicutes) than to its photosynthetic homologous (data not shown). This suggests that *Gloeobacter* ChlH2 is not involved in chlorophyll biosynthesis.

The presence of BchH/ChlH isoforms in three out of the six considered bacterial phyla and their monophyly within the two clades may suggest their presence in the ancestral chlorophyll pathway, in which case the isoforms would have been lost (or not LGT acquired) in proteobacteria, acidobacteria, and heliobacteria.

The AAA^+ Motor: BchI/ChlI and BchD/ChlD

The BchI/ChlI subunit is a member of the AAA^+-ATPase family of proteins, which includes proteins of diverse function (Fodje et al. 2001). With the exception of some land plants (Kobayashi et al. 2008; Huang and Li 2009) and green algae (Lohr et al. 2005), there is a consensus in the literature that photosynthetic organisms only have one BchI/ChlI subunit. However, a study of two recombinant isoforms of BchI from *Prosthecochloris vibrioformis* showed that both had Mg chelatase activity in vitro (Petersen et al. 1998), and in the recently sequenced genome of *Chloroflexus aurantiacus*, three genes were annotated as *BchI* (Tang et al. 2011). In our search, several isoforms belonging to both chlorobia and chloroflexi were retrieved. These enzymes are distinct from the aerobic

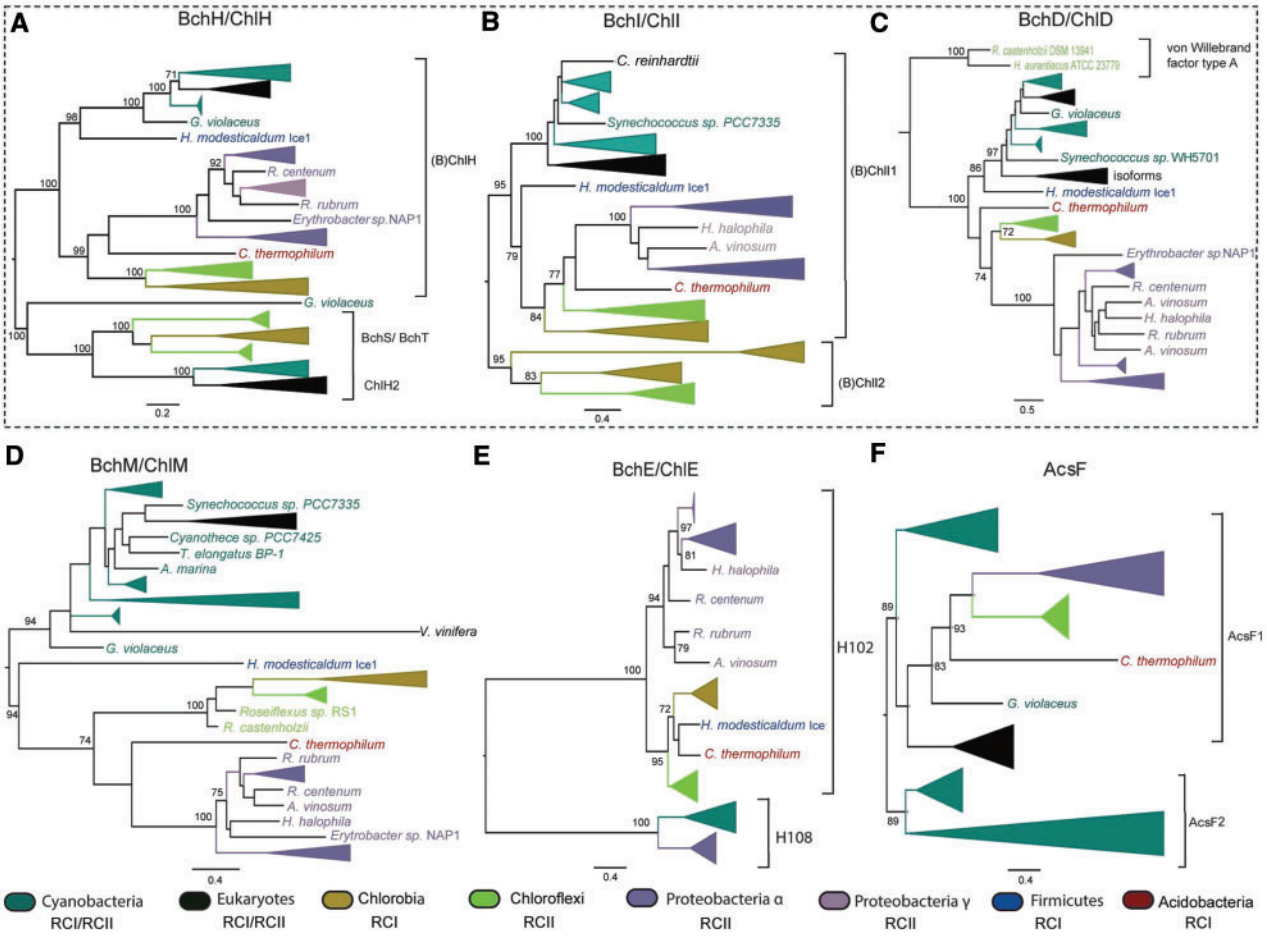

Fig. 3.—Phylogenetic relationships among photosynthetic bacteria based on the different (bacterio)chlorophyll genes. (*A*) *BchH/ChlH*, (*B*) *BchI/ChlI*, (*C*) *BchD/ChlD*, (*D*) *BchM/ChlM*, (*E*) *BchE/ChlE,* and (*F*) *AcsF*.

cobalamin chelatase *CobS* gene and were included in the analysis (fig. 3*B*).

The BchI/ChlI tree is divided into two clades, one comprising *BchI/ChlI* genes among all photosynthesizers and one containing the different isoforms from GSB and GNSB, which we call the GSB/GNSB clade. In the former, two major groups can be observed, one with heliobacteria, chlorobia, chloroflexi, acidobacteria, and proteobacteria (branching in that order) and a second one, where the cyanobacteria and eukaryotic sequences cluster together. The isoforms present in the GSB/GNSB clade, as in the case of *G. violaceus* ChlH2 sequence, have high similarity with sequences similar to hypothetical magnesium chelatases from nonphotosynthetic organisms. Although the results from *P. vibrioformis* indicate that both isoforms have identical activities in vitro, their sequences show higher similarity with sequences from organisms that do not synthetize bacteriochlorophyll, questioning their role in chlorophyll biosynthesis. Specifically, within delta-proteobacteria, proprionobacterales, firmicutes, and euryarchaeota, some organisms possess isoforms of the *BchI/ChlI* gene in addition to *CobS*. Other cobalamin producers possess only the *BchI/ChlI* gene and lack *CobS*. In these organisms, the *BchI/ChlI* genes probably substitute the missing *CobNST* genes (Xiong et al. 1998; Lohr et al. 2005). The distributions of *BchI/ChlI* and *CobS* homologs suggest a way in which pre-existing building blocks could have been recruited into the assembly of the ancestral chlorophyll and O_2-dependent cobalamin pathway.

The N-terminus of *BchD/ChlD* gene, which is also a member of the AAA^+-ATPase family of proteins (Fodje et al. 2001), has a segment of approximately 260 amino acid residues with sequence homology to BchI/ChlI (Masuda et al. 1999). BchD/ChlD has a proposed structural role in magnesium chelatase, functioning as a platform for the convergence of the other two subunits (Axelsson et al. 2006). The BchD/ChlD phylogenetic tree in figure 3*C* was rooted by two sequences of von Willebrand factor Type A, which is considered to be an ancient protein domain (Ponting et al. 1999). The overall BchD/ChlD tree topology is very similar to the upper part of the BchH/ChlH tree with the exception of the position of *C. Chloracidobacterium thermophilum*. There are

two distinct BchD/ChlD clades. One contains acidobacteria, the GSB/GNSB clade, and proteobacteria. In the second, heliobacteria, cyanobacteria, and eukaryotic sequences cluster. There are two eukaryotic clades, one composed of green algae isoforms that branch between heliobacteria and cyanobacteria, and a second in which at least one copy from every eukaryotic organism is represented in a sister group of the *G. violaceus* sequence. However, this branch is poorly supported. As in the case of *BchI/ChlI*, copies of this gene are also present in nonphotosynthetic organisms.

Mg Protoporphyrin IX Methyltransferase

The next step in the pathway of synthesis is the insertion of a methyl group by the enzyme Mg-protoporphyrin IX methyl transferase (BchM/ChlM) to form the Mg-protoporphyrin IX monomethyl ester (MPE) (fig. 2*A*, step 2). This enzyme belongs to the Class I *S*-adenosylmethionine-(SAM-)dependent methyltransferase family (Gibson and Hunter 1994) and transiently interacts with BchH/ChlH (Johnson and Schmidt-Dannert 2008). With the exception of *Arabidopsis thaliana*, all photosynthetic organisms have only one *BchM/ChlM* gene. There is strong support for monophyly of the major photosynthetic taxa, and in this case, because cyanobacteria and eukaryotes cluster together, while heliobacteria branch closer to the chlorobia/chloroflexi sister groups and acidobacteria closer to proteobacteria (fig. 3*D*).

Mg-Protoporphyrin IX Monomethyl Ester Oxidative Cyclases

The next step in the pathway is the formation of the isocyclic ring by the conversion of MPE to 3,8-divinyl protochlorophyllide *a*. This complex reaction entails the cyclization of the ring and the introduction of a keto group at carbon 13^1. This step is catalyzed by either of two different and unrelated enzymes: the oxygen-dependent enzyme AcsF (fig. 2*A*, step 4) and the oxygen-independent enzyme (BchE), both of which catalyze the reaction (fig. 2*A*, step 3). The arguably more ancient reaction of the oxygen-independent cyclase, BchE, uses water as electron donor and is widely distributed among heliobacteria, proteobacteria, GNSB, and GSB. Although all cyanobacteria have the AcsF enzyme, some, including *Synechococcus* sp. PCC 6803, also harbor a functional *ChlE* gene (Minamizaki et al. 2008) that can complement *Rhodobacter capsulatus BchE*-deficient strains (Bollivar 2006). BchE/ChlE is a radical-SAM enzyme with homology to adenosylcobalamin-binding domain. In *R. capsulatus* and *Chlorobium limicola strain 1230,* the enzyme appears to be vitamin B12 (cobalamin) dependent (Fuhrmann et al. 1993; Gough et al. 2000). However, these two sequences lack the canonical signature motif (DXHXXG) that typically contains the cobalt coordinating histidine residue (Drennan et al. 1994). The BchE/ChlE alignment reveals, however, a highly conserved histidine (H102 *Chlorobaculum tepidum* TLS numbering) that might function as the missing cobalt ligand, and, in sequences where H102 is missing, another histidine residue (H108 *Synechocystis* sp. *PCC6803* numbering) is present close by, whereas no functional role of either in B12 binding has been shown. The two clades of the phylogenetic BchE tree (fig. 3*E*) can be divided by the position of this residue

Chlorobaculum tepidum has seven *BchE* homologous genes, whereby only three (*BchE, BchR,* and *BchQ*) are involved in bacteriochlorophyll biosynthesis (Gomez Maqueo Chew et al. 2007). BchR and BchQ act as methyl transferases performing specific methylations in the later steps of Bacteriochlorophyll *c* pathway (Gomez Maqueo Chew et al. 2007) and probably arose from later lineage-specific duplication events. These paralogs and all *BchE*-related genes with higher sequence similar to homologs from nonphotosynthetic organisms were removed from the phylogenetic analysis.

The BchE tree reveals two well-supported major clades, one containing the *Chlorobaculum tepidum* and *R. capsulatus* sequences and another where cyanobacterial homologs and proteobacterial BchE isoforms are sister groups (fig. 3*E*). Within the first clade, there is a clear separation between proteobacteria sequences from the other taxa. In this tree, the sequences belonging to heliobacteria and acidobacteria cluster between chloroflexi and chlorobia.

AcsF is the enzyme from the O_2-dependent pathway. It is present in cyanobacteria, photosynthetic eukaryotes, and chloroflexi, as well as in some photosynthetic proteobacteria. This oxygen-dependent enzyme most likely had its origin within the cyanobacteria and was vertically inherited and horizontally transferred to other photosynthetic organisms. Until the discovery of photosynthesis in Acidobacteria, this enzyme was only present in RCII-containing organisms. The presence of an O_2-dependent enzyme in an RCI phototroph is interesting due to the usual oxygen sensitivity of the iron–sulfur centers of RCI (Tsukatani et al. 2012). However, the *C. Chloracidobacterium thermophilum* RCI center is not oxygen sensitive, probably because of adaptation of the organism to the aerobic lifestyle (Tsukatani et al. 2012). The AcsF phylogenetic tree can be divided into two major clades, both containing cyanobacterial sequences (fig. 3*F*). Nostocales and *Cyanothece* taxa have two AcsF isoforms named here AcsF1 and AcsF2. At first sight, the dichotomy of the tree might seem to be a reflection of this duplication event. Nevertheless, none of the clades contain sequences from all represented cyanobacteria organisms. Specifically, the major branch contains AcsF1 from all photosynthetic phyla chosen, but no *Prochlorococcus* sequences are present. These cluster in the lower branch of the tree, with the other cyanobacterial AcsF2 isoforms. This could suggest that the oxygenic photosynthetic ancestor (or at least the extinct ancestor of nostocales and *Cyanothece* and possibly *Prochlorococcus*) had two AcsF isoforms, which were selectively lost in some organisms, with the AcsF1 isoform having been laterally transferred to

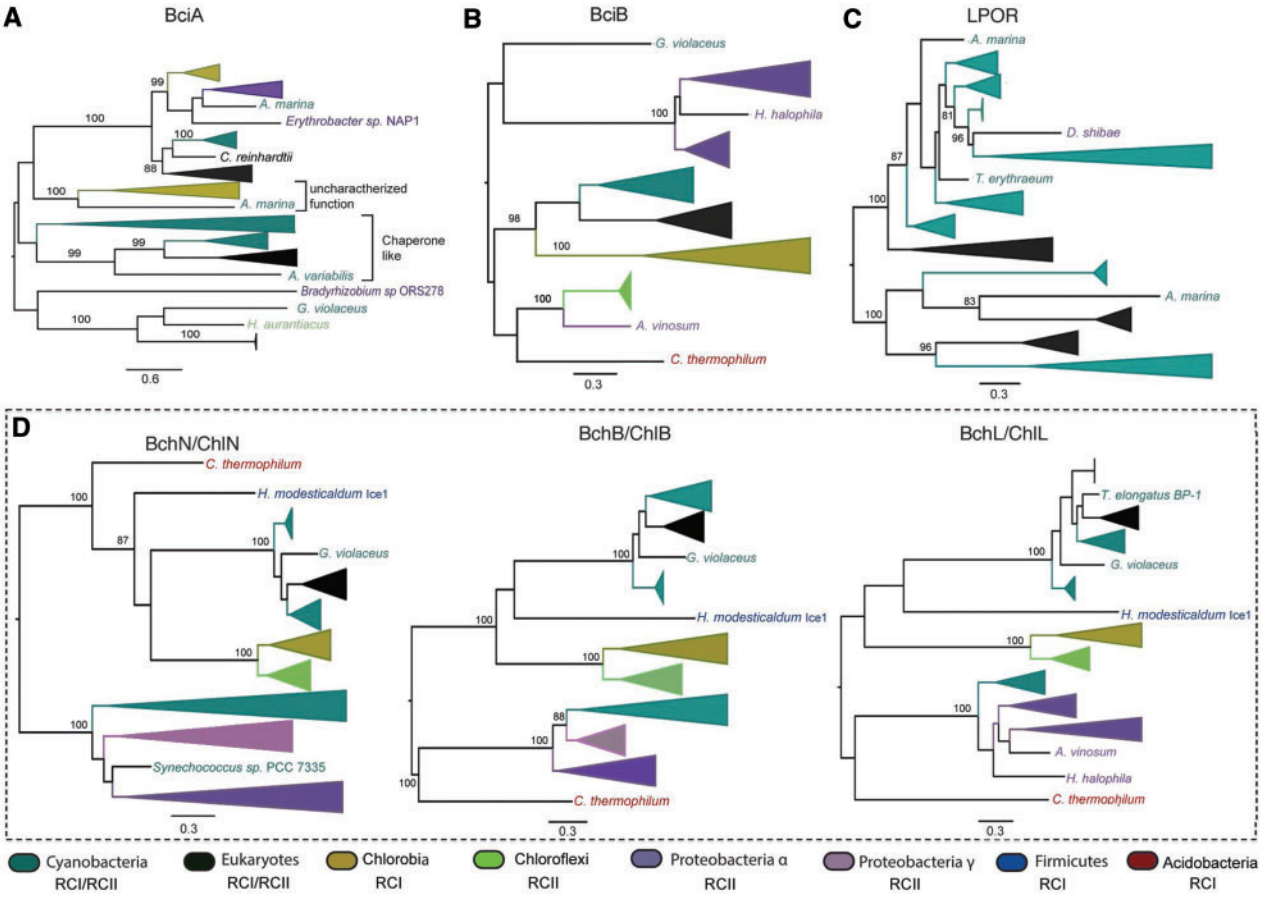

Fig. 4.—Phylogenetic relationships among photosynthetic bacteria based on the different (bacterio)chlorophyll genes. (*A*) *BciA*, (*B*) *BciB*, (*C*) *LPOR*, and (*D*) *DPOR* with *BchN/ChlN*, *BchB/ChlB*, and *BchL/ChlL* gene trees.

other photosynthetic lineages. The *C. Chloracidobacterium thermophilum* sequence is basal to the group of noncyanobacterial AcsF sequences and has higher sequence similarity with the chloroflexi clade.

(Di)Vinyl Reductases

With exception of some *Prochlorococcus* strains (Chisholm et al. 1992) and heliobacteria (bacteriochlorophyll *g*), (bacterio)chlorophylls have an ethyl group at position 8 of ring II. The reduction of the 8-vinyl group by a 8-vinyl reductase enzyme may occur at various steps of the chlorophyll biogenesis pathway (Tripathy and Rebeiz 1988; Kim and Rebeiz 1996; Adra and Rebeiz 1998; Kolossov and Rebeiz 2001) either in a sequential or in a parallel route (Parham and Rebeiz 1995; Nagata et al. 2005), thereby generating some of the observed heterogeneity of chlorophylls and their intermediates. For clarity, here we show a sequential representation, keeping the other possible alternatives in mind.

There are at least two unrelated enzymes that catalyze this reaction: BciA and BciB (or slr1923) (Islam et al. 2008; Ito et al. 2008; Bryant et al. 2012). BciA (fig. 2*A*, step 5) belongs to the short-chain dehydrogenases/reductases family having a ligand-binding site and a $NAD(P)^+$/NAD(P)H-binding site. It is present in some organisms belonging to chlorobia, cyanobacteria, eukaryotes, and proteobacteria. Mutagenesis and biochemical experiments identified the *BciA* gene in *Chlorobaculum tepidum, A. thaliana,* and *Oryza sativa* (Nagata et al. 2005; Nakanishi et al. 2005; Chew and Bryant 2007; Wang et al. 2010). As seen in figure 4*A*, these orthologo genes form a well-supported clade that can be subdivided into two groups, one comprising GSB and proteobacteria sequences and another with cyanobacteria and eukaryotic sequences. The only exception is the *Acaryochloris marina* sequence that clusters within the proteobacterial clade.

Because of the broad distribution of NAD(P)H-binding sites (Rossman folds) in proteins, database searching retrieved many additional hits (lower part of the tree). BciA shares sequence similarity with a proposed chaperone for quinone binding in PSII (Ermakova-Gerdes and Vermaas 1999) and also with some sequences of unknown function among

chlorobia. Interestingly, another *Aca. marina* isoform is positioned with the functionally undetermined chlorobial sequences. To root the tree, a group of five sequences belonging to photosynthetic and nonphotosynthetic organisms was used.

The *BciB* gene (fig. 2*B*, step 6) is present in all phyla considered here, and, with exception of gamma-proteobacteria, the groups are all monophyletic (fig. 4*B*). BciB has high homology with coenzyme F420-reducing hydrogenase only found in methanogenic archaea (Islam et al. 2008). The *G. violaceus* sequence, basal to the proteobacterial clade, represents a second copy (an isoform) of the *BciB* gene in the *Gloeobacter* genome and was probably a later acquisition via lateral transfer from proteobacteria. The BciB tree is one of the few cases where chloroflexi and chlorobia are not sister groups, with chloroflexi closer to *Allochromatium vinosum* and *C. Chloracidobacterium thermophilum* sequences. Also, BciB shows separation between the major proteobacterial clade and the other photosynthetic taxa.

Protochlorophyllide Reductases

The last step of the common metabolic pathway of (bacterio)chlorophyll consists of the reduction of the IV pyrrole ring of protochlorophyllide to form chlorophyllide. Two distinct enzymes have evolved to catalyze this reaction, the LPOR and the light-independent protochlorophyllide reductase (DPOR) composed of three different subunits, BchL/ChlL, BchN/ChlN, and BchB/ChlB.

LPOR (fig. 2*A*, step 7) belongs to the short chain alcohol dehydrogenases family and is present in oxygenic photosynthetic organisms (Dahlin et al. 1999). This enzyme most likely originated in cyanobacteria and was acquired by eukaryotes at the origin of plastids (Yang and Cheng 2004). Because of its involvement in the development of etiolated and greening tissues (Fujita 1996), LPOR from different organisms has been studied extensively in terms of mechanism, function, localization, and expression (summarized in Reinbothe et al. [2010]). Several isoforms of this nuclear encoded enzyme are present in eukaryotic phototrophs (named PorA, PorB, and PorC in *A. thaliana*) with differential expression and function along *A. thaliana* development (Masuda et al. 2003). Besides the isoforms present in eukaryotes, additional putative LPORs belonging to cyanobacteria and *Dinoroseobacter shibae* DFL 12 were retrieved via database searching. The *Dinoroseobacter* gene clusters within the *Synechococcus* clade and is probably a recent gene acquisition. The tree in figure 4*C* can be divided into two major clades, each one containing cyanobacterial and eukaryotic homologs. The upper part of the tree contains biochemically characterized LPORs. The several PorA, PorB, and PorC enzymes found in angiosperms cluster together, indicating duplication and functional or regulatory specialization during plant evolution.

The lower part of the LPOR tree contains cyanobacterial isoforms and eukaryotic isoforms, the latter, so far, uncharacterized. Curiously, in this subtree, angiosperm (*Vitis vinifera* and *Oryza sativa*) and algal sequences (*Micromonas* sp. and *Ostreococcus lucimarinus*) form two statistical significant groups, the former clustering with *Aca. marina* and *Cyanothece* sequences and the later with marine cyanobacteria (*Synechococcus* and *Prochlorococcus*). It is possible that these two major LPOR isoforms existed before the origin of plastids and were acquired then, with subsequent differential loss. This suggestion is in agreement with the presence of multiple LPOR isoforms in modern cyanobacteria.

The DPOR (fig. 2*A*, step 8) is present in all photosynthetic phyla, having been lost in most of the angiosperm species (Fong and Archibald 2008). Its three subunits, (B)ChlLNB, have homology with a second chlorin reductase complex BchXYZ present only in bacteriochlorophyll-containing organisms. This complex is responsible for the reduction of the C7–C8 bond, converting chlorophyllide into bacterochlorophyllide. (B)ChlLNB also has homology with the nitrogenase complex (nifHDK) responsible for nitrogen fixation in some bacterial and archaeal taxa (Leigh 2000; Dedysh et al. 2004; Raymond et al. 2004). The nifHDK complex is proposed to have an ancient origin and to be at the base of the (B)ChlLNB complex (Dorr et al. 2003; Martin 2012).

Two groups of cyanobacteria are observed in (B)ChlLNB trees (fig. 4*D*). Sequences from marine cyanobacteria (*Synechococcus* and *Prochlorococcus*) cluster with proteobacterial homologs, whereas the remaining cyanobacterial sequences and the two eukaryotes known to have (B)ChlLNB (*Physcomitrella patens* and *Chlamydomonas reinhardtii*) cluster next to the chlorobia/chloroflexi clades. Events of LGT between cyanobacteria and proteobacteria are well documented (Badger and Price 2003; Beiko et al. 2005; Zhaxybayeva et al. 2006), especially for the closely related *Synechococcus* and *Prochlorococcus* genera. However, the polarity of this LGT (who was donor, who was recipient) is still not established (Bryant et al. 2012; Gupta 2012). In the different DPOR trees, the heliobacterial sequences align either with nonmarine cyanobacteria or with the chlorobia/cloroflexi group, whereas the acidobacterial sequence is basal either to the proteobacterial clade (BchB/ChlB and BchL/ChlL) or to the *Heliobacterium*-containing clade (BchN/ChlN).

Discussion

Individually, the trees for proteins underlying chlorophyll biosynthesis are complex. Although they do not all tend strongly to reflect a single underlying topology, they do have aspects in common. Their main implication in the context of this article is as follows: taken as a whole, the trees for chlorophyll biosynthesis appear to distinguish between competing hypotheses for the presence of two serially linked photosystems at the

origin of water-splitting photosynthesis. This behavior is statistically supported by the AU and SH tests, with rejection of trees where separation of RCI and RCII organisms was imposed.

Two Cyanobacterial RCs: Gene Duplication, Not Lineage Merger

The most important observation from this study is that there is no coevolutionary pattern linking chlorophyll biosynthesis gene phylogeny with either RCIs or RCIIs; in other words, we observe neither Type I- or Type II-specific chlorophyll biosynthesis genes nor Type I- or Type II-specific chlorophyll biosynthesis gene phylogenies. Because RCs cannot undergo evolution in the absence of chlorophyll, this lack of coevolutionary pattern linking chlorophyll biosynthesis to the divergence of RCI and RCII allows us to exclude the widely discussed possibility that RCI and RCII diverged via lineage splitting and became reunited in cyanobacteria via a large-scale gene transfer event (Mathis 1990; Blankenship 1992; Meyer 1994; Blankenship and Hartman 1998; Xiong et al. 1998, 2000; Blankenship 2001; Xiong and Bauer 2002a, 2002b; Blankenship 2010), of the kind that Hohmann-Marriott and Blankenship (2011) call "fusion theories." Indeed, the consistently close proximity of the GSB (chlorobia, with RCI) and green nonsulfur bacteria (chloroflexi, with RCII) in chlorophyll biosynthesis trees argues strongly against the view that there was a deep evolutionary split in chlorophyll biosynthesis corresponding to a lineage split between RCI and RCII, one that would be expected to have pulled (B)Chl genes in tow. This is all the more true given the tendency for the chlorophyll biosynthesis genes from proteobacteria (with RCII) to branch in close proximity to their cyanobacterial homologs, rather than with homologs from green nonsulfur bacteria. In addition, the chlorophyll biosynthesis genes of *C. Chloracidobacterium thermophylum* (with RCI) usually branch in close proximity to RCII-containing taxa (proteobacteria or chloroflexi).

If we thus exclude ancient lineage splitting and later (re)union in cyanobacteria at the evolutionary origin of two RC types coexisting in the same cell, the simplest competing alternative, and one widely discussed in the literature (Olson and Pierson 1987a; Vermaas 1994; Olson 2001; Allen 2005; Mulkidjanian et al. 2006), is that gene duplication giving rise to RCI and RCII within the same genome gave rise to the photosystem configuration in oxygenic photosynthesis. This possibility is easily reconciled with chlorophyll biosynthesis phylogenies. An immediately ensuing question is: in which lineage did the gene duplication take place? Occam's razor clearly favors the premise that the photosystem genes underwent duplication in an ancestral cyanobacterium—a protocyanobacterium—because cyanobacteria are the only group where genes for both RC types have remained present and expressed. It is possible to assume that the duplication took place elsewhere, but there is no obvious alternative location.

What Use Might Two Photosystems Be?

Olson and Pierson (1987a, 1987b) have suggested an ancient gene duplication event in a cyanobacterium for the origin of the two photosystems. However, that formulation accounted for the distribution of chlorophyll-based photosynthesis and RC types among prokaryotic groups exclusively by vertical inheritance and differential loss. Today, with the number of bacterial lineages having grown, the distribution of photosynthesis and RCs (fig. 1) has become much more sparse than in 1987, such that differential loss alone is unlikely to account in full for these distributions, especially because photosynthesis genes are observed to be mobile in the marine phage metagenome, and because substantial amounts of LGT have indeed been shown for photosynthetic lineages (Raymond et al. 2002; Huang and Gogarten 2007; Shi and Falkowski 2008). Hence, a mechanistic mixture of some vertical inheritance, some differential loss, and some lateral transfer, relative amounts of which might differ from gene to gene across (B)Chl synthesis, needs to be invoked, because the trees for (B)Chl biosynthesis genes are insufficiently similar for artifacts of phylogenetic reconstruction to account for their differences.

However, how much and what kind of LGT might be required to explain the distribution of photosynthesis among prokaryotes? Various possibilities have recently been discussed, for example, by Bryant et al. (2012) and Gupta (2012). However, that question is not the focus of this article. Rather, the question at our focus is how two photosystems came to reside within a single genome so as to give rise to oxygenic photosynthesis. Gene duplication in a protocyanobacterium is the alternative most compatible with present data. If there were two photosystems, what were they doing?

This question has to do with the transition from anoxygenic photosynthesis to water splitting. Blankenship and Hartman (1998) suggested that hydrogen peroxide (H_2O_2) might have been an initial electron donor far more chemically accessible than water, for a linear two-photosystem ETC. Mulkidjanian et al. (2006) suggested that H_2 might have been the initial electron donor for a two-photosystem ETC. Nisbet and coworkers (Nisbet et al. 1995; Nisbet and Sleep 2001) suggest that the manganese complex evolved either to handle excess of peroxide or as a toxic weapon against competitors.

An alternative suggestion (Mathis 1990; Allen and Puthiyaveetil 2005) differs from the foregoing models with respect to the presumed function of the two photosystems. It posits that the two photosystems in the protocyanobacterium operated in a temporally regulated manner (Allen 2005), for example, as an H_2S-oxidizing and NAD^+-reducing RCI when H_2S was available as it occurs in modern *Chlorobium* or as a light-driven proton pump (cyclic electron flow

thorough RCII as in *Rhodobacter*) when H_2S was not available. Before the water-splitting complex had evolved, what would a bacterium with two different and specialized, nonoxygenic photosystems have done with them? Probably just what modern bacteria do: express them when needed, with the help of a regulatory switch. This model implies relatively strict regulation of the RC genes, because in the event that regulation failed, for example, through mutation of the regulatory protein constituting the redox switch, both RCIs and RCIIs became expressed in the absence of H_2S, the protocyanobacterium would be exposed to a lethal level of oxidative stress (there is no way to turn off assembled photosystems), unless it could extract electrons from an environmentally available donor.

In principle, such a donor could have been H_2O_2, or H_2, or an organic compound such as succinate. However, it also could have been aqueous $Mn^{II/III}$ cations, which have the interesting property of giving off low potential electrons under ultraviolet (UV) radiation (photooxidation) (Anbar and Holland 1992; Hakala et al. 2006; Allen and Martin 2007; Russell et al. 2008). Before the accumulation of atmospheric oxygen (and hence ozone), UV was a larger component of the solar radiation that reaches the Earth's surface than it is today. Thus, the presence of light for photosynthesis would have also meant the presence of UV radiation for $Mn^{II/III}$ photooxidation. Photooxidation of aqueous, environmental $Mn^{II/III}$ could have thus literally "pushed" electrons into photooxidized chlorophyll of RCII. That would have led to a lethal log-jam of electrons, or over-reduction, in the electron-transport cycle—unless the protocyanobacterium was simultaneously expressing RCI. This complex would redirect the surplus membrane-bound electrons stemming from photosystem II into CO_2-reduction and thereby create precisely the flow of electrons seen in cyanobacteria today: a linear flow through two photosystems.

Obviously, such an ETC would hardly have been perfect from the beginning. And clearly there is a difference between tapping environmental Mn atoms photooxidatively, one at a time, and the establishment of the resident, catalytic Mn_4Ca center in cyanobacterial RCII water-splitting complex. However, the overall contours of the circuit would have been right, from manganese at RCII and going on through RCI to NADH and ultimately to CO_2. The key to water splitting would then have entailed the transition from exploiting an environmental supply of soluble manganese, where each electron-donating Mn ion would reach photosystem II by simple diffusion, to holding four manganese atoms (and a calcium of as-yet-unknown function) in place. Dissolved Mn^{2+} has long been known to donate electrons to photosystem II, and thus to reconstitute noncyclic electron transport, in isolated chloroplasts biochemically depleted of the capacity to oxidize water to oxygen (Cheniae 1970). Following adsorption of Mn to its donor side, a fine tuning of photosystem II by natural selection to optimize its reduction/oxidation potential could have allowed it to oxidize a now biologically portable manganese reservoir four times in a row.

It is notable that the Mn atoms of the water-splitting complex are bound directly by the proteins of the photosystem II RC, without an intervening protein or electron carrier. This suggests that no major evolutionary invention was required for photosystem II to tap environmental Mn as an electron source. In line with that, it has recently been shown (Allen et al. 2012) that an engineered, Mn-binding RCII of *R. sphaeroides* will produce O_2 from O_2^- in the presence of Mn in a light-dependent reaction in which photodamage is impeded in comparison with that in a wild-type, Mn-free RC. Allen et al. (2012) interpret this observation as an important clue to the origin of oxygenic photosynthesis.

In what sort of setting could suitably high concentrations of $Mn^{II/III}$ have accumulated for such series of events to transpire? Mn is not a good candidate for high enough concentrations to make the model viable in an open ocean setting. However, in a locally circumscribed freshwater setting, sufficiently high $Mn^{II/III}$ concentrations could, in principle, have accumulated.

In this context, it is of particular interest that Blank and Sáchez-Baracaldo (2010) recently provided evidence that oxygenic photosynthesis arose in a freshwater environment, based on the basal phylogenetic position of freshwater cyanobacteria and the derived phylogenetic position of marine cyanobacteria.

Related Pathways and the Cell That Invented Chlorophyll

Chlorophylls coordinate Mg^{2+} and belong to the tetrapyrrole family, which includes cobalamin (Co^{2+}), heme (Fe^{2+}), siroheme (Fe^{2+}), heme d1 (Fe^{2+}), and F_{430} (Ni^{2+}) (Heinemann et al. 2008; Warren and Smith 2009; Zappa et al. 2010). All tetrapyrrole biosynthetic pathways are related in the sense that they start from the universal precursor, δ-aminolevulinic acid, and share three enzymatic steps that generate the tetrapyrrole macrocycle uroporphyrinogen III (UroIII) (fig. 5). Proto IX is the common precursor for chlorophyll synthesis and heme synthesis via the classical pathway that occurs in eubacteria and eukaryotes (Heinemann et al. 2008). Precorrin-2 is the common precursor for cobalamin, heme synthesis via the alternative pathway recently discovered in archaebacteria (Storbeck et al. 2010; Bali et al. 2011), siroheme (a cofactor in some nitrite and all sulfite reductases (Tripathy et al. 2010), heme d1 (a cofactor only present in the bacterial cd_1 nitrite reductases (Allen et al. 2005), and F_{430} pathways. F_{430} is a cofactor that is critical to methanogenesis and has only been found in methanogens so far (Thauer 1998). In addition to two routes for heme synthesis, there are also two routes for cobalamin synthesis, one O_2 dependent (often called the late pathway due to the late insertion of Mg^{2+}) and one O_2 independent (early insertion of Mg^{2+}) (Martens et al. 2002; Warren et al. 2002). The O_2-dependent cobalamin pathway involves the Co^{2+} chelatase CobNST, the three subunits of

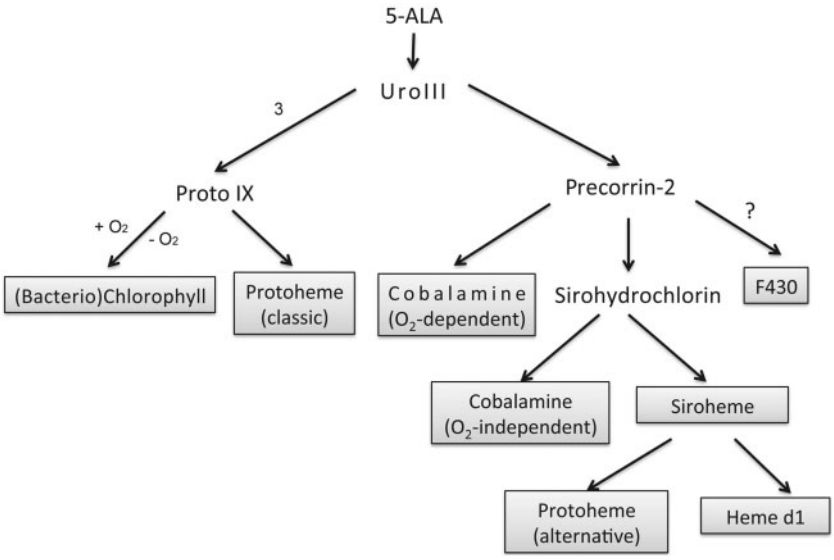

Fig. 5.—Schematic representation of the bifurcations present in the global tetrapyrrole pathway. The biosynthetic pathway of all tetrapyrroles begins with the condensation of eight molecules of 5-amino-levulinate to uroporphyrinogen-II (UroIII), the first cyclic tetrapyrrole. This compound can be sequentially converted to protoporphyrin IX in three enzymatic steps. This is the last common precursor for (bacterio)chlorophyll synthesis and heme synthesis via the classical pathway (present in most eubacteria and eukaryotes). Alternatively, UroIII can be converted to precorrin-2, the branching point of F430 synthesis, of the O_2-dependent cobalamin pathway, and of the formation of Sirohydrochlorin. At the level of Sirohydrochlorin, the O_2-independent cobalamin pathways diverge from the formation of siroheme. Siroheme may be used as a cofactor, further transformed into heme d1 or transformed into heme according to the alternative heme pathway present in archaea and some eubacteria.

which are related to (B)ChlHID. It has been suggested that (B)ChlHID arose from CobNST (Xiong et al. 2000), but the converse might be more likely given the presence of one O_2-dependent step in the late cobalamin pathway (Martens et al. 2002; McGoldrick et al. 2005). In the early (O_2 independent) cobalamin pathway, the CbiK/CbiX^L/CbiX^S chelatases are related to the SirB and HemH chelatases of the siroheme and classical heme pathways, respectively (Schubert et al. 1999; Brindley et al. 2003; Romao et al. 2011). The chelatase for the F_{430} pathway is unknown, but genes related to *(B)ChlHID* are present in methanogens and encode candidates for the Ni^{2+}-chelatase. In some organisms, the insertion of iron and cobalt into sirohydrochlorin is performed by multifunctional chelatases (CysG and Met8p) (Spencer et al. 1993; Fazzio and Roth 1996; Schubert et al. 2002) (Class III).

What did the cell that invented chlorophyll biosynthesis have in terms of tetrapyrrole pathways? It definitely contained cobalamin and might have had a cobalamin biosynthesis pathway (the O_2-independent type, obviously), because the O_2-independent route to 3,8-divinyl protochlorophyllide *a* via (B)ChlE, a cobalamin-dependent reaction (Fuhrmann et al. 1993; Gough et al. 2000). Of course, it could also have just been cobalamin dependent; some cyanobacteria synthesize cobalamin, and many, however, acquire it from the environment via high affinity importers (Tang et al. 2012). The cell that invented chlorophyll also had a classical heme pathway, because barring chlorophyll, the three steps from UroIII to Proto IX are specific to classical heme synthesis (Heinemann et al. 2008) and hence preceded the (B)Chl pathway. The first cell with chlorophyll probably also fixed nitrogen, because the subunit of DPOR, (B)ChlNBL, is related to nitrogenase subunits (Xiong et al. 1998). We propose that the cell that invented chlorophyll was a cobalamin-dependent, heme-synthesizing, diazotrophic anaerobe.

Conclusions

The phylogeny and evolution of the chlorophyll biosynthesis core pathway were analyzed both at the gene and pathway level. The lack of coevolution of chlorophyll biosynthesis genes with RCIs or RCIIs permits us to exclude the widely discussed possibility that RCI and RCII diverged via lineage splitting and became reunited in cyanobacteria via a large-scale gene transfer (fusion) event. Moreover, it can be concluded that the primordial photosynthetic organism performed nitrogen fixation, synthesized heme, and was cobalamin dependent.

Note Added in Proof

In support of the genetic mobility of anoxygenic photosysnthesis, photosynthesis is encoded on a plasmid in Roseobacter littoralis (Kalhöfer et al. 2012).

Supplementary Material

Supplementary figures S1–S12 and tables S1 and S2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

The authors thank Bob Blankenship for critical comments on the text. This work was supported by the Alexander von Humboldt Foundation (a postdoctoral fellowship to L.S.-G.), the ERC (to W.F.M. and F.L.S.), and the Leverhulme Trust (as Research Grant F07476AQ to J.F.A.).

Literature Cited

Adra AN, Rebeiz CA. 1998. Chloroplast biogenesis-81: transient formation of divinyl chlorophyll a following a 2.5 ms light flash treatment of etiolated cucumber cotyledons. Photochem Photobiol. 68:852–856.

Alberti M, Burke DH, Hearst JE. 1995. Structure and sequence of the photosynthesis gene cluster. In: Blankenship RE, Madigan MT, Bauer CE, editors. Anoxygenic photosynthetic bacteria. Advances in photosynthesis and respiration, Vol. 2.. Dordrecht (The Netherlands): Springer. p. 1083–1106.

Allen JF. 2005. A redox switch hypothesis for the origin of two light reactions in photosynthesis. FEBS Lett. 579:963–968.

Allen JF, de Paula WB, Puthiyaveetil S, Nield J. 2011. A structural phylogenetic map for chloroplast photosynthesis. Trends Plant Sci. 16: 645–655.

Allen JF, Martin W. 2007. Evolutionary biology: out of thin air. Nature 445: 610–612.

Allen JF, Puthiyaveetil S. 2005. *Choroflexus aurantiacus* and the origin of oxygenic, two-light reaction photosynthesis in failure to switch between type I and type II reaction centres. In: Est A, Bruce D, editors. Photosynthesis: fundamental aspects to global perspectives. Lawrence (KS): Alliance Communications Group. p. 753–756.

Allen JP, et al. 2012. Light-driven oxygen production from superoxide by Mn-binding bacterial reaction centers. Proc Natl Acad Sci U S A. 109: 2314–2318.

Allen JW, et al. 2005. Why isn't "standard" heme good enough for c-type and d1-type cytochromes? Dalton Trans. 21:3410–3418.

Altschul SF, et al. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389–3402.

Anbar AD, Holland HD. 1992. The photochemistry of manganese and the origin of banded iron formations. Geochim Cosmochim Acta. 56: 2595–2603.

Arnold GL, Anbar AD, Barling J, Lyons TW. 2004. Molybdenum isotope evidence for widespread anoxia in mid-proterozoic oceans. Science 304:87–90.

Axelsson E, et al. 2006. Recessiveness and dominance in barley mutants deficient in Mg-chelatase subunit D, an AAA protein involved in chlorophyll biosynthesis. Plant Cell 18:3606–3616.

Badger MR, Price GD. 2003. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. J Exp Bot. 54:609–622.

Bali S, et al. 2011. Molecular hijacking of siroheme for the synthesis of heme and d1 heme. Proc Natl Acad Sci U S A. 108: 18260–18265.

Barber J, Morris E, Buchel C. 2000. Revealing the structure of the photosystem II chlorophyll binding proteins, CP43 and CP47. Biochim Biophys Acta. 1459:239–247.

Baymann F, Brugna M, Muhlenhoff U, Nitschke W. 2001. Daddy, where did (PS)I come from? Biochim Biophys Acta. 1507: 291–310.

Beiko RG, Harlow TJ, Ragan MA. 2005. Highways of gene sharing in prokaryotes. Proc Natl Acad Sci U S A. 102:14332–14337.

Bernal A, Ear U, Kyrpides N. 2001. Genomes OnLine Database (GOLD): a monitor of genome projects world-wide. Nucleic Acids Res. 29: 126–127.

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

Blankenship RE. 1992. Origin and early evolution of photosynthesis. Photosynth Res. 33:91–111.

Blankenship RE. 2001. Molecular evidence for the evolution of photosynthesis. Trends Plant Sci. 6:4–6.

Blankenship RE. 2010. Early evolution of photosynthesis. Plant Physiol. 154:434–438.

Blankenship RE, Hartman H. 1998. The origin and evolution of oxygenic photosynthesis. Trends Biochem Sci. 23:94–97.

Bollivar DW. 2006. Recent advances in chlorophyll biosynthesis. Photosynth Res. 90:173–194.

Brindley AA, Raux E, Leech HK, Schubert HL, Warren MJ. 2003. A story of chelatase evolution: identification and characterization of a small 13-15-kDa "ancestral" cobaltochelatase ($CbiX^S$) in the archaea. J Biol Chem. 278:22388–22395.

Bryant D, Costas AG, Heidelberg J, Ward D. 2007. *Candidatus Chloracidobacterium thermophilum*: an aerobic phototrophic Acidobacterium with chlorosomes and type 1 reaction centers. Photosynth Res. 91:269–270.

Bryant DA, et al. 2012. Comparative and functional genomics of anoxygenic green bacteria from the taxa Chlorobi, Chloroflexi, and Acidobacteria. In: Burnap R, Vermaas W, editors. Functional genomics and evolution of photosynthetic systems. Vol. 33. Dordrecht (The Netherlands): Springer. p. 47–102.

Canfield DE, et al. 2008. Ferruginous conditions dominated later neoproterozoic deep-water chemistry. Science 321:949–952.

Chen M, et al. 2010. A red-shifted chlorophyll. Science 329:1318–1319.

Cheniae GM. 1970. Photosystem-II and O_2 evolution. Ann Rev Plant Physiol. 21:467–498.

Chew AG, Bryant DA. 2007. Characterization of a plant-like protochlorophyllide a divinyl reductase in green sulfur bacteria. J Biol Chem. 282: 2967–2975.

Chisholm SW, et al. 1992. *Prochlorococcus marinus* Nov Gen-Nov Sp—an oxyphototrophic marine prokaryote containing divinyl chlorophyll-a and chlorophyll-b. Arch Microbiol. 157:297–300.

Dahl C, et al. 2005. Novel genes of the dsr gene cluster and evidence for close interaction of Dsr proteins during sulfur oxidation in the phototrophic sulfur bacterium *Allochromatium vinosum*. J Bacteriol. 187: 1392–1404.

Dahlin C, et al. 1999. The role of protein surface charge in catalytic activity and chloroplast membrane association of the pea NADPH: protochlorophyllide oxidoreductase (POR) as revealed by alanine scanning mutagenesis. Plant Mol Biol. 39:309–323.

Dedysh SN, Ricke P, Liesack W. 2004. NifH and NifD phylogenies: an evolutionary basis for understanding nitrogen fixation capabilities of methanotrophic bacteria. Microbiology 150:1301–1313.

Dorr M, et al. 2003. A possible prebiotic formation of ammonia from dinitrogen on iron sulfide surfaces. Angew Chem Int Ed Engl. 42: 1540–1543.

Drennan CL, Huang S, Drummond JT, Matthews RG, Lidwig ML. 1994. How a protein binds B12: a 3.0 A X-ray structure of B12-binding domains of methionine synthase. Science 266:1669–1674.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.

Ermakova-Gerdes S, Vermaas W. 1999. Inactivation of the open reading frame slr0399 in *Synechocystis* sp. PCC 6803 functionally complements mutations near the Q(A) niche of photosystem II—a possible role of slr0399 as a chaperone for quinone binding. J Biol Chem. 274: 30540–30549.

Fazzio TG, Roth JR. 1996. Evidence that the CysG protein catalyzes the first reaction specific to B-12 synthesis in *Salmonella typhimurium*, insertion of cobalt. J Bacteriol. 178:6952–6959.

Fodje MN, et al. 2001. Interplay between an AAA module and an integrin I domain may regulate the function of magnesium chelatase. J Mol Biol. 311:111–122.

Fong A, Archibald JM. 2008. Evolutionary dynamics of light-independent protochlorophyllide oxidoreductase genes in the secondary plastids of cryptophyte algae. Eukaryot Cell. 7:550–553.

Fuhrmann S, Overmann J, Pfennig N, Fischer U. 1993. Influence of Vitamin-B12 and light on the formation of chlorosomes in green-colored and brown-colored *Chlorobium* species. Arch Microbiol. 160:193–198.

Fujita Y. 1996. Protochlorophyllide reduction: a key step in the greening of plants. Plant Cell Physiol. 37:411–421.

Garcia Costas AM, et al. 2012. Complete genome of *Candidatus Chloracidobacterium thermophilum*, a chlorophyll-based photoheterotroph belonging to the phylum acidobacteria. Environ Microbiol. 14: 177–190.

Gibson LC, Hunter CN. 1994. The bacteriochlorophyll biosynthesis gene, *bchM*, of *Rhodobacter sphaeroides* encodes S-adenosyl-L-methionine: Mg protoporphyrin IX methyltransferase. FEBS Lett. 352:127–130.

Glaeser J, Baneras L, Rutters H, Overmann J. 2002. Novel bacteriochlorophyll e structures and species-specific variability of pigment composition in green sulfur bacteria. Arch Microbiol. 177:475–485.

Gomez Maqueo Chew A, Frigaard NU, Bryant DA. 2007. Bacteriochlorophyllide c C-8^2 and C-12^1 methyltransferases are essential for adaptation to low light in *Chlorobaculum tepidum*. J Bacteriol. 189:6176–6184.

Gomez Maqueo Chew A, Frigaard NU, Bryant DA. 2009. Mutational analysis of three bchH paralogs in (bacterio-)chlorophyll biosynthesis in *Chlorobaculum tepidum*. Photosynth Res. 101:21–34.

Gough SP, Petersen BO, Duus JO. 2000. Anaerobic chlorophyll isocyclic ring formation in *Rhodobacter capsulatus* requires a cobalamin cofactor. Proc Natl Acad Sci U S A. 97:6908–6913.

Granick S. 1965. Evolution of heme and chlorophyll. In: Bryson G, Vogel HJ, editors. Evolving genes and proteins. New York: Academic Press. p. 67–88.

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. 2012. Origin and spread of photosynthesis based upon conserved sequence features in key bacteriochlorophyll biosynthesis proteins. Mol Biol Evol. 29:3397–3412.

Hakala M, Rantamaki S, Puputti EM, Tyystjarvi T, Tyystjarvi E. 2006. Photoinhibition of manganese enzymes: insights into the mechanism of photosystem II photoinhibition. J Exp Bot. 57:1809–1816.

Hauska G, Schoedl T, Remigy H, Tsiotis G. 2001. The reaction center of green sulfur bacteria. Biochim Biophys Acta. 1507:260–277.

Heinemann IU, Jahn M, Jahn D. 2008. The biochemistry of heme biosynthesis. Arch Biochem Biophys. 474:238–251.

Hohmann-Marriott MF, Blankenship RE. 2011. Evolution of photosynthesis. Ann Rev Plant Biol. 62:515–548.

Holland HD. 2006. The oxygenation of the atmosphere and oceans. Philos Trans R Soc Lond B Biol Sci. 361:903–915.

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

Huang YS, Li HM. 2009. *Arabidopsis* CHLI2 Can Substitute for CHLI1. Plant Physiol. 150:636–645.

Iino T, et al. 2010. *Ignavibacterium album* gen. nov., sp nov., a moderately thermophilic anaerobic bacterium isolated from microbial mats at a terrestrial hot spring and proposal of Ignavibacteria classis nov., for a novel lineage at the periphery of green sulfur bacteria. Int J Syst Evol Microbiol. 60:1376–1382.

Islam MR, et al. 2008. slr1923 of *Synechocystis* sp. PCC6803 is essential for conversion of 3,8-divinyl(proto)chlorophyll(ide) to 3-monovinyl (proto)chlorophyll(ide). Plant Physiol. 148:1068–1081.

Ito H, Yokono M, Tanaka R, Tanaka A. 2008. Identification of a novel vinyl reductase gene essential for the biosynthesis of monovinyl chlorophyll in *Synechocystis* sp. PCC6803. J Biol Chem. 283:9002–9011.

Johnson ET, Schmidt-Dannert C. 2008. Characterization of three homologs of the large subunit of the magnesium chelatase from *Chlorobaculum tepidum* and interaction with the magnesium protoporphyrin IX methyltransferase. J Biol Chem. 283:27776–27784.

Kalhöfer D, Thole S, Voget S, Lehmann R, Liesegang H, Wollher A, Daniel R, Simon M, Brinkhoff T. 2012. Comparative genome analysis and genome-guided physiological analysis of *Roseobacter litoralis*. BMC Genomics 12:324.

Kiang NY, Siefert J, Govindjee, Blankenship RE. 2007. Spectral signatures of photosynthesis. I. Review of Earth organisms. Astrobiology 7: 222–251.

Kiang NY, et al. 2007. Spectral signatures of photosynthesis. II. Coevolution with other stars and the atmosphere on extrasolar worlds. Astrobiology 7:252–274.

Kim JS, Rebeiz CA. 1996. Origin of chlorophyll alpha biosynthetic heterogeneity in higher plants. J Biochem Mol Biol. 29:327–334.

Kobayashi M, et al. 2000. The primary electron acceptor of green sulfur bacteria, bacteriochlorophyll 663, is chlorophyll a esterified with Delta 2,6-phytadienol. Photosynth Res. 63:269–280.

Kobayashi K, et al. 2008. Functional analysis of *Arabidopsis thaliana* isoforms of the Mg-chelatase CHLI subunit. Photochem Photobiol Sci. 7: 1188–1195.

Kolossov VL, Rebeiz CA. 2001. Chloroplast biogenesis 84: Solubilization and partial purification of membrane-bound [4-vinyl] chlorophyllide a reductase from etiolated barley leaves. Anal Biochem. 295:214–219.

Leigh JA. 2000. Nitrogen fixation in methanogens: the archaeal perspective. Curr Issues Mol Biol. 2:125–131.

Lindell D, et al. 2004. Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. Proc Natl Acad Sci U S A. 101: 11013–11018.

Liu Z, et al. 2012. Complete genome of *Ignavibacterium album*, a metabolically versatile, flagellated, facultative anaerobe from the phylum chlorobi. Front Microbiol. 3:185.

Lockhart PJ, Larkum AWD, Steel MA, Waddell PJ, Penny D. 1996. Evolution of chlorophyll and bacteriochlorophyll: The problem of invariant sites in sequence analysis. Proc Natl Acad Sci U S A. 93: 1930–1934.

Lohr M, Im CS, Grossman AR. 2005. Genome-based examination of chlorophyll and carotenoid biosynthesis in *Chlamydomonas reinhardtii*. Plant Physiol. 138:490–515.

Lundqvist J, et al. 2009. The AAA(+) motor complex of subunits CobS and CobT of cobaltochelatase visualized by single particle electron microscopy. J Struct Biol. 167:227–234.

Lundqvist J, et al. 2010. ATP-induced conformational dynamics in the AAA plus motor unit of magnesium chelatase. Structure 18:354–365.

Lyons TW. 2007. Palaeoclimate: oxygen's rise reduced. Nature 448: 1005–1006.

Lyons TW, Anbar AD, Severmann S, Scott C, Gill BC. 2009. Tracking euxinia in the ancient ocean: a multiproxy perspective and proterozoic case study. Ann Rev Earth Planet Sci. 37:507–534.

Manske AK, Glaeser J, Kuypers MAM, Overmann J. 2005. Physiology and phylogeny of green sulfur bacteria forming a monospecific

phototrophic assemblage at a depth of 100 meters in the Black Sea. Appl Env Microbiol. 71:8049–8060.
Maqueo Chew AG, Bryant DA. 2007. Chlorophyll biosynthesis in bacteria: the origins of structural and functional diversity. Ann Rev Microbiol. 61:113–129.
Margulis L, Bermudes D. 1985. Symbiosis as a mechanism of evolution: status of cell symbiosis theory. Symbiosis 1:101–124.
Martens JH, Barg H, Warren MJ, Jahn D. 2002. Microbial production of vitamin B12. Appl Microbiol Biotechnol. 58:275–285.
Martin WF. 2012. Hydrogen, metals, bifurcating electrons, and proton gradients: the early evolution of biological energy conservation. FEBS Lett. 586:485–493.
Masuda T, et al. 1999. Magnesium insertion by magnesium chelatase in the biosynthesis of zinc bacteriochlorophyll a in an aerobic acidophilic bacterium *Acidiphilium rubrum*. J Biol Chem. 274: 33594–33600.
Masuda T, et al. 2003. Functional analysis of isoforms of NADPH:protochlorophyllide oxidoreductase (POR), PORB and PORC, in *Arabidopsis thaliana*. Plant Cell Physiol. 44:963–974.
Mathis P. 1990. Compared structure of plant and bacterial photosynthetic reaction centers—evolutionary implications. Biochim Biophys Acta. 1018:163–167.
McGoldrick HM, et al. 2005. Identification and characterization of a novel vitamin B-12 (Cobalamin) biosynthetic enzyme (CobZ) from *Rhodobacter capsulatus*, containing flavin, heme, and Fe-S cofactors. J Biol Chem. 280:1086–1094.
Meyer TE. 1994. Evolution of photosynthetic reaction centers and light-harvesting chlorophyll proteins. Biosystems 33:167–175.
Michel H, Deisenhofer J. 1988. Relevance of the photosynthetic reaction center from purple bacteria to the structure of Photosystem-II. Biochemistry 27:1–7.
Minamizaki K, Mizoguchi T, Goto T, Tamiaki H, Fujita Y. 2008. Identification of two homologous genes, *chlAI* and *chlAII*, that are differentially involved in isocyclic ring formation of chlorophyll *a* in the cyanobacterium *Synechocystis* sp. PCC 6803. J Biol Chem. 283: 2684–2692.
Mix LJ, Haig D, Cavanaugh CM. 2005. Phylogenetic analyses of the core antenna domain: investigating the origin of photosystem I. J Mol Evol. 60:153–163.
Mulkidjanian AY, et al. 2006. The cyanobacterial genome core and the origin of photosynthesis. Proc Natl Acad Sci U S A. 103: 13126–13131.
Nagata N, Tanaka R, Satoh S, Tanaka A. 2005. Identification of a vinyl reductase gene for chlorophyll synthesis in *Arabidopsis thaliana* and implications for the evolution of *Prochlorococcus* species. Plant Cell 17: 233–240.
Nakanishi H, et al. 2005. Characterization of the *Arabidopsis thaliana* mutant pcb2 which accumulates divinyl chlorophylls. Plant Cell Physiol. 46:467–473.
Neerken S, Amesz J. 2001. The antenna reaction center complex of heliobacteria: composition, energy conversion and electron transfer. Biochim Biophys Acta. 1507:278–290.
Niedzwiedzki DM, Blankenship RE. 2010. Singlet and triplet excited state properties of natural chlorophylls and bacteriochlorophylls. Photosynth Res. 106:227–238.
Nisbet EG, Cann JR, Vandover CL. 1995. Origins of photosynthesis. Nature 373:479–480.
Nisbet EG, Sleep NH. 2001. The habitat and nature of early life. Nature 409:1083–1091.
Nitschke W, Rutherford AW. 1991. Photosynthetic reaction centres: variations on a common structural theme? Trends Biochem Sci. 16: 241–245.
Olson JM. 2001. "Evolution of Photosynthesis" (1970), re-examined thirty years later. Photosynth Res. 68:95–112.
Olson JM, Pierson BK. 1987a. Evolution of reaction centers in photosynthetic prokaryotes. Int Rev Cytol. 108:209–248.
Olson JM, Pierson BK. 1987b. Origin and evolution of photosynthetic reaction centers. Origins Life Evol Biosph. 17:419–430.
Parham R, Rebeiz CA. 1995. Chloroplast biogenesis 72: a [4-vinyl]-chlorophyllide a reductase assay using divinyl chlorophyllide a as an exogenous substrate. Anal Biochem. 231:164–169.
Petersen BL, et al. 1998. Reconstitution of an active magnesium chelatase enzyme complex from the *bchI*, *-D*, and *-H* gene products of the green sulfur bacterium *Chlorobium vibrioforme* expressed in *Escherichia coli*. J Bacteriol. 180:699–704.
Ponting CP, Aravind L, Schultz J, Bork P, Koonin EV. 1999. Eukaryotic signalling domain homologues in archaea and bacteria. Ancient ancestry and horizontal gene transfer. J Mol Biol. 289:729–745.
Pruitt KD, Tatusova T, Maglott DR. 2005. NCBI reference sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. 33:D501–D504.
Raymond J, Siefert JL, Staples CR, Blankenship RE. 2004. The natural history of nitrogen fixation. Mol Biol Evol. 21:541–554.
Raymond J, Zhaxybayeva O, Gogarten JP, Gerdes SY, Blankenship RE. 2002. Whole-genome analysis of photosynthetic prokaryotes. Science 298:1616–1620.
Reid JD, Siebert CA, Bullough PA, Hunter CN. 2003. The ATPase activity of the ChlI subunit of magnesium chelatase and formation of a heptameric AAA+ ring. Biochemistry 42:6912–6920.
Reinbothe C, et al. 2010. Chlorophyll biosynthesis: spotlight on protochlorophyllide reduction. Trends Plant Sci. 15:614–624.
Rodionov DA, Vitreschak AG, Mironov AA, Gelfand MS. 2003. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. J Biol Chem. 278:41148–41159.
Romao CV, et al. 2011. Evolution in a family of chelatases facilitated by the introduction of active site asymmetry and protein oligomerization. Proc Natl Acad Sci U S A. 108:97–102.
Russell MJ, Allen JF, Milner-White EJ. 2008. Inorganic complexes enabled the onset of life and oxygenic photosynthesis. In: Allen JF, Gantt E, Golbeck JH, Osmond B, editors. Photosynthesis. Energy from the sun: 14th International Congress on Photosynthesis. Heidelberg (Germany): Springer. p. 1187–1192.
Sadekar S, Raymond J, Blankenship RE. 2006. Conservation of distantly related membrane proteins: photosynthetic reaction centers share a common structural core. Mol Biol Evol. 23:2001–2007.
Sahoo SK, et al. 2012. Ocean oxygenation in the wake of the Marinoan glaciation. Nature 489:546–549.
Sarrou I, et al. 2012. Purification of the photosynthetic reaction center from *Heliobacterium modesticaldum*. Photosynth Res. 111: 291–302.
Schubert HL, Raux E, Wilson KS, Warren MJ. 1999. Common chelatase design in the branched tetrapyrrole pathways of heme and anaerobic cobalamin synthesis. Biochemistry 38:10660–10669.
Schubert WD, et al. 1998. A common ancestor for oxygenic and anoxygenic photosynthetic systems: a comparison based on the structural model of photosystem I. J Mol Biol. 280:297–314.
Schubert HL, et al. 2002. The structure of *Saccharomyces cerevisiae* Met8p, a bifunctional dehydrogenase and ferrochelatase. EMBO J. 21:2068–2075.
Scott C, et al. 2008. Tracing the stepwise oxygenation of the Proterozoic ocean. Nature 452:456–459.
Sharon I, et al. 2009. Photosystem I gene cassettes are present in marine virus genomes. Nature 461:258–262.
Shi T, Falkowski PG. 2008. Genome evolution in cyanobacteria: the stable core and the variable shell. Proc Natl Acad Sci U S A. 105: 2510–2515.
Shimodaira H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst Biol. 51:492–508.

Shimodaira H, Hasegawa M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol Biol Evol. 16:1114–1116.
Shimodaira H, Hasegawa M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17:1246–1247.
Spencer JB, Stolowich NJ, Roessner CA, Scott AI. 1993. The *Escherichia coli Cysg* gene encodes the multifunctional protein, siroheme synthase. FEBS Lett. 335:57–60.
Stomp M, Huisman J, Stal LJ, Matthijs HC. 2007. Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. ISME J. 1:271–282.
Storbeck S, et al. 2010. A novel pathway for the biosynthesis of heme in archaea: genome-based bioinformatic predictions and experimental evidence. Archaea 2010:175050.
Tang K, Jiao N, Liu K, Zhang Y, Li S. 2012. Distribution and functions of TonB-dependent transporters in marine bacteria and environments: implications for dissolved organic matter utilization. PLoS One 7: e41204.
Tang KH, et al. 2011. Complete genome sequence of the filamentous anoxygenic phototrophic bacterium *Chloroflexus aurantiacus*. BMC Genomics 12:334.
Thauer RK. 1998. Biochemistry of methanogenesis: a tribute to Marjory Stephenson. 1998 Marjory Stephenson Prize Lecture. Microbiology 144:2377–2406.
Tripathy BC, Rebeiz CA. 1988. Chloroplast biogenesis-60—conversion of divinyl protochlorophyllide to monovinyl protochlorophyllide in green (ing) *Barley*, a dark monovinyl Light divinyl plant-species. Plant Physiol. 87:89–94.
Tripathy BC, Sherameti I, Oelmuller R. 2010. Siroheme: an essential component for life on Earth. Plant Signal Behav. 5:14–20.
Tsukatani Y, Romberger SP, Golbeck JH, Bryant DA. 2012. Isolation and characterization of homodimeric type-I reaction center complex from *Candidatus Chloracidobacterium thermophilum*, an aerobic chlorophototroph. J Biol Chem. 287:5720–5732.
Vermaas WFJ. 1994. Evolution of *heliobacteria*—implications for photosynthetic reaction-center complexes. Photosynth Res. 41:285–294.
Walker CJ, Willows RD. 1997. Mechanism and regulation of Mg-chelatase. Biochem J. 327:321–333.
Wang P, et al. 2010. Divinyl chlorophyll(ide) a can be converted to monovinyl chlorophyll(ide) a by a divinyl reductase in rice. Plant Physiol. 153: 994–1003.
Warren M, Smith A. 2009. Tetrapyrroles: birth, life, and death. New York: Springer Science & Business Media.
Warren MJ, Raux E, Schubert HL, Escalante-Semerena JC. 2002. The biosynthesis of adenosylcobalamin (vitamin B12). Nat Prod Rep. 19: 390–412.
Willows RD, Beale SI. 1998. Heterologous expression of the *Rhodobacter capsulatus BchI*, *-D*, and *-H* genes that encode magnesium chelatase subunits and characterization of the reconstituted enzyme. J Biol Chem. 273:34206–34213.
Willows RD, Hansson A, Birch D, Al-Karadaghi S, Hansson M. 2004. EM single particle analysis of the ATP-dependent BchI complex of magnesium chelatase: an AAA+ hexamer. J Struct Biol. 146: 227–233.
Xiong J, Bauer CE. 2002a. Complex evolution of photosynthesis. Ann Rev Plant Biol. 53:503–521.
Xiong J, Bauer CE. 2002b. A cytochrome *b* origin of photosynthetic reaction centers: an evolutionary link between respiration and photosynthesis. J Mol Biol. 322:1025–1037.
Xiong J, Fischer WM, Inoue K, Nakahara M, Bauer CE. 2000. Molecular evidence for the early evolution of photosynthesis. Science 289: 1724–1730.
Xiong J, Inoue K, Bauer CE. 1998. Tracking molecular evolution of photosynthesis by characterization of a major photosynthesis gene cluster from *Heliobacillus mobilis*. Proc Natl Acad Sci U S A. 95:14851–14856.
Yang J, Cheng Q. 2004. Origin and evolution of the light-dependent protochlorophyllide oxidoreductase (LPOR) genes. Plant Biol. 6: 537–544.
Zappa S, Li K, Bauer CE. 2010. The tetrapyrrole biosynthetic pathway and its regulation in *Rhodobacter capsulatus*. Adv Exp Med Biol. 675: 229–250.
Zhaxybayeva O, Gogarten JP, Charlebois RL, Doolittle WF, Papke RT. 2006. Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. Genome Res. 16:1099–1108.

Associate editor: Yves Van De Peer