Origin and Evolution of Water Oxidation before the Last Common Ancestor of the Cyanobacteria

Tanai Cardona,*[1] James W. Murray,[1] and A. William Rutherford[1]
[1]Department of Life Sciences, Imperial College London, London, United Kingdom
*Corresponding author: E-mail: t.cardona@imperial.ac.uk.
Associate editor: Manolo Gouy

Abstract

Photosystem II, the water oxidizing enzyme, altered the course of evolution by filling the atmosphere with oxygen. Here, we reconstruct the origin and evolution of water oxidation at an unprecedented level of detail by studying the phylogeny of all D1 subunits, the main protein coordinating the water oxidizing cluster (Mn_4CaO_5) of Photosystem II. We show that D1 exists in several forms making well-defined clades, some of which could have evolved before the origin of water oxidation and presenting many atypical characteristics. The most ancient form is found in the genome of *Gloeobacter kilaueensis* JS-1 and this has a C-terminus with a higher sequence identity to D2 than to any other D1. Two other groups of early evolving D1 correspond to those expressed under prolonged far-red illumination and in darkness. These atypical D1 forms are characterized by a dramatically different Mn_4CaO_5 binding site and a Photosystem II containing such a site may assemble an unconventional metal cluster. The first D1 forms with a full set of ligands to the Mn_4CaO_5 cluster are grouped with D1 proteins expressed only under low oxygen concentrations and the latest evolving form is the dominant type of D1 found in all cyanobacteria and plastids. In addition, we show that the plastid ancestor had a D1 more similar to those in early branching *Synechococcus*. We suggest each one of these forms of D1 originated from transitional forms at different stages toward the innovation and optimization of water oxidation before the last common ancestor of all known cyanobacteria.

Key words: water oxidation, Photosystem II, PsbA, cyanobacteria, oxygenic photosynthesis.

Introduction

Light-driven water oxidation by Photosystem II evolved only once in an ancestor of the phylum Cyanobacteria. The exact timing for the origin of oxygenic photosynthesis is debated, from shortly before (Kopp et al. 2005; Rasmussen et al. 2008) to several hundred million years prior to the Great Oxygenation Event, about 2.4 billion years ago (Waldbauer et al. 2009; Flannery and Walter 2012; Lyons et al. 2014). Oxygenic photosynthesis involves the light-driven oxidation of water at a complex tetramanganese cluster (Mn_4CaO_5) within Photosystem II. In this reaction, four electrons are extracted from two water molecules, with the release of four protons into the thylakoid lumen and O_2 as a byproduct. All cyanobacteria discovered and studied until now have descended from a common ancestor with a photosynthetic apparatus that is already highly specialized for oxygenic photosynthesis. Even the earliest diverging cyanobacteria of the genus *Gloeobacter* have a Photosystem II very similar to those of late evolving members of the phylum (Nakamura et al. 2003; Koyama et al. 2008; Saw et al. 2013). This presupposes a period of evolutionary innovation that must have existed before the last common ancestor of the phylum Cyanobacteria, a hidden or extinct biota that turned a simpler photosynthetic reaction center into the sophisticated photosystem required for water oxidation. Although several hypotheses have been proposed to explain the origin of water oxidation and the Mn_4CaO_5 cluster (Blankenship and

Hartman 1998; Dismukes et al. 2001; Sauer and Yachandra 2002; Rutherford and Faller 2003; Johnson et al. 2013), it is still unclear how an ancestral bacterium evolved the capacity to split water.

Photosystem II is a large and complex redox enzyme composed of at least 17 protein subunits and a large number of cofactors specialized in light harvesting, photoprotection, and electron and proton transfer reactions including water oxidation chemistry (Cardona et al. 2012; Dau et al. 2012; Muh et al. 2012; Cox et al. 2013). The reaction center cofactors involved in charge separation and water oxidation are coordinated by a pair of homologous protein subunits, known as D1 (PsbA) and D2 (PsbD). The D1 protein provides most of the ligands to the Mn_4CaO_5 cluster where water oxidation occurs (Ferreira et al. 2004; Umena et al. 2011). The D1 and D2 proteins are also homologous to the L and M subunits of anoxygenic Type II reaction centers found in the phyla Proteobacteria and Chloroflexi. Consequently, Photosystem II and anoxygenic Type II reaction centers must have evolved from a primordial homodimeric Type II reaction center (Michel and Deisenhofer 1988; Beanland 1990; Nitschke and Rutherford 1991).

Unlike the genomes of anoxygenic phototrophic bacteria, the genomes of most cyanobacteria contain multiple genes encoding reaction center proteins. These include several *psbA* genes, some of which show significant sequence variation. The total number of *psbA* gene copies and variant forms of

Open Access

D1 varies from species to species. These additional copies and variant forms probably endow the bacterium with the ability to adapt to a very wide range of conditions. At least two different photoprotective mechanisms have been identified involving the additional gene copies of D1. The first mechanism is simply the stress-induced enhancement of the expression of additional copies of the *psbA* gene encoding identical D1 proteins and this presumably enhances the repair process (Mulo et al. 2009, 2012). The second mechanism entails the D1 subunit being replaced by a different form that is better adapted to the new conditions. The best known example of the second mechanism is the expression of D1 variants under high-light in which glutamate replaces glutamine at position 130 (Kulkarni and Golden 1994; Tichy et al. 2003). The glutamate provides a stronger hydrogen bond (H-bond) to the pheophytin electron acceptor, $Pheo_{D1}$, making its redox potential more positive (Giorgi et al. 1996; Sugiura et al. 2010). This provides tolerance to highlight by diminishing the formation of damaging chlorophyll triplet states and therefore decreasing photoinhibition (Merry et al. 1998; Cser and Vass 2007; Ogami et al. 2012; Sugiura et al. 2013).

In some cyanobacteria a different *psbA* gene has been shown to be transcribed under microaerobic conditions (Summerfield et al. 2008; Sicora et al. 2009); this protein, which is commonly referred to as D1′, has approximately 85% sequence similarity to other D1 copies within the organism. Photosystem II containing the D1′ isolated from *Thermosynechococcus elongatus* BP-1 showed modifications in the proton-coupled electron transfer on specific steps of the enzyme cycle (Sugiura et al. 2012, 2013). The functional significance of this under low oxygen concentrations remains to be elucidated (Sugiura et al. 2012) but could represent a protective adaptation for growth with an overreduced electron transfer chain.

Recently, it was pointed out that an extremely divergent form of D1 occurred in 13 of the cyanobacterial genomes sequenced at that time (Murray 2012). These so-called "rogue" D1 (rD1) made a consistent group of D1 sequences, diverging early and with extensive modifications both at the Mn_4CaO_5 cluster and at the exchangeable quinone (Q_B) binding site. These rD1 had about 65% sequence similarity to the conventional D1 proteins. Murray (2012) suggested that the rD1 variants could represent a "missing link" between anoxygenic and oxygenic Type II reaction centers. The transcription of a *psbA* gene encoding a rD1 was found to be enhanced in the N_2-fixing *Cyanothece* sp. ATCC 51142, *Cyanothece* sp. PCC 7822 and *Chrocosphaera watsonii* WH 8501 in the night, when oxygen sensitive enzymes such as nitrogenase and the uptake hydrogenase are most active (Toepel et al. 2008; Shi et al. 2010; Zhang and Sherman 2012; Wegener et al. 2014). In a recent transcriptomic analysis in *Anabaena variabilis* ATCC 29413 it was shown that heterotrophic growth (in the dark and on fructose) enhanced the transcription of a *psbA* gene encoding a rD1, both in vegetative cells and heterocysts (Park et al. 2013). Another rD1 was studied in the chlorophyll *d*-containing cyanobacterium *Acaryochloris marina* MBIC11017, in which the gene for the rD1 was abundantly transcribed after prolonged exposure to

dibromothymoquinone (DBMIB), an inhibitor of the Cytochrome b_6f complex, although no direct evidence of its natural function was provided (Kiss et al. 2012). It was suggested that rD1 in general may not be functional in water splitting and this may be to permit N_2 fixation or other O_2 sensitive processes to occur (Toepel et al. 2008; Murray 2012). This was also discussed for *A. marina* despite it being non-diazotrophic (unlike other *Acaryochloris* strains), with the suggestion that this strain might have lost the ability to fix N_2 in recent evolutionary time (Kiss et al. 2012; Murray 2012).

In addition to these rD1 variants, another type of *psbA* gene was identified at that time in *Synechococcus* sp. PCC 7335 and this was suggested to be the earliest diverging D1 form with a sequence similarity of approximately 55% to any conventional D1. This was designated a "super-rogue" D1 (Murray 2012). New cyanobacterial genomes have shown that these super-rogue D1 sequences are not unique to *Synechococcus* 7335, but are also present in few other strains. In a recent study, Gan et al. (2014) showed that the expression of a gene encoding a super-rogue D1 was enhanced after prolonged far-red light exposure. Furthermore, the strains containing a super-rogue D1 are also capable of synthetizing chlorophylls *d* and *f* as part of that far-red light acclimation response (Gan et al. 2014).

In this article, we have analyzed the phylogeny of all D1 proteins to understand the origin and evolution of oxygenic photosynthesis. We identified at least five distinct types of D1 and their peculiar characteristics have allowed us to reconstruct a sequential evolutionary scenario for the origin of water oxidation in Photosystem II.

Results

Diversity of D1 Proteins

Figure 1 shows a rooted phylogenetic tree for D1 in cyanobacteria; it is complex and has some unexpected features (for a fully expanded tree see supplementary fig. S1, Supplementary Material online). The D1 tree does not follow conventional cyanobacterial phylogenies, in which the earliest diverging cyanobacterial species, such as *Gloeobacter violaceus* PCC 7421, *Synechococcus* sp. JA-3-3Ab and JA-2-3B′a, branch basally (Gupta 2010; Criscuolo and Gribaldo 2011; Larsson et al. 2011; Shih et al. 2013). This is due to the large number of the *psbA* gene duplications that has occurred; some of which appear to be very ancient whereas others appear to be more recent, as described in more detail below. In this data set we found an average of 2.83 D1 sequences per D2 sequence, where some unicellular marine cyanobacteria such as *Prochlorococcus* and *Synechococcus* had a single copy of the gene, whereas most cyanobacterial species have three to six distinct copies of D1 (fig. 2). The two species with the greatest number of D1 proteins were found to be *Leptolyngbia* sp. PCC 7375 and *Leptolyngbia* sp. Heron Island J with seven and eight distinct D1, respectively.

The phylogenetic tree shows that D1 exists in several different forms. The earliest diverging form of D1 is found only in the recently sequenced genome of *Gloeobacter kilaueensis*
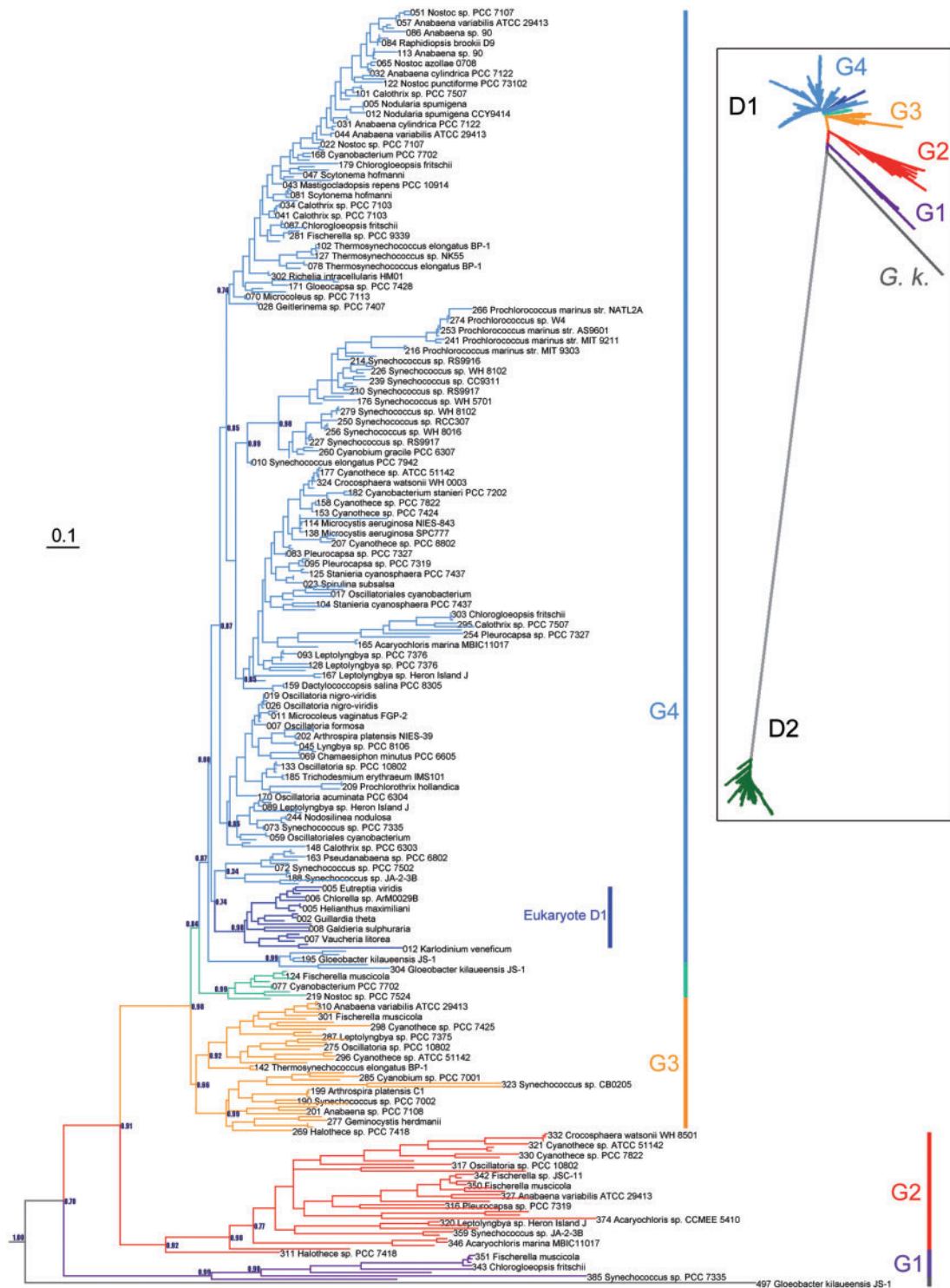
FIG. 1. Rooted maximum likelihood phylogeny of D1 proteins showing the evolutionary relationships of D1 sequences. D2 sequences were used as an outgroup (inset). Branch supports are expressed as approximate likelihood ratio test aLRT probabilities. *G.k.* refers to the atypical sequence from *Gloeobacter kilaueensis* JS-1; G1, G2, G3, and, G4, refer to D1 sequences from Group 1 to Group 4. Scale bar represents 10% amino acid substitution.

JS-1 (Saw et al. 2013). The genome of this organism contains six *psbA* genes encoding four different D1 primary sequences, one of which is unlike any other D1 protein found in the data set and it appears to have many ancient traits. This type of D1 lacks all of the ligands for the Mn_4CaO_5 cluster with the exception of E189. It is thus highly unlikely that this protein, if

incorporated into a Photosystem II complex, could support water oxidation. The sequence shares 51% identity with the well-characterized D1 from *T. elongatus* (PsbA1). The C-terminus of this sequence seems to show hybrid traits between D1 and D2 (see fig. 3). In particular, within the last 25 amino acids there are five positions that are strictly conserved in D2
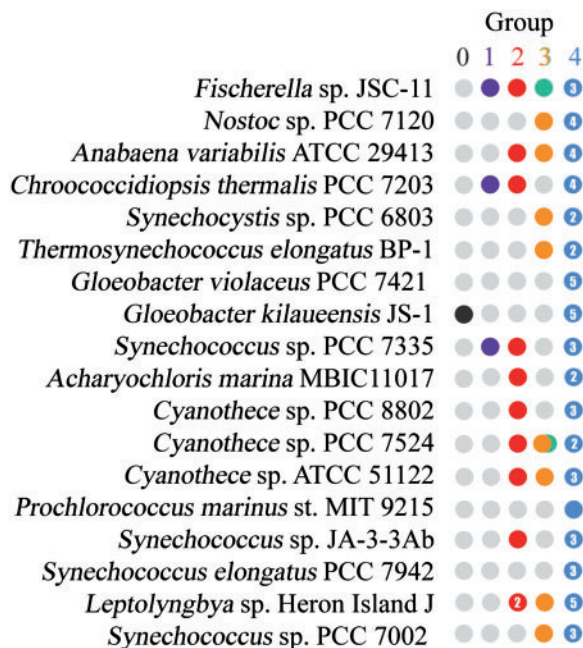
FIG. 2. Distribution of variant forms of D1 in a selected number of cyanobacteria. Gray or colored circles denote absence or presence of a gene from a particular group in the genome of that bacterium, respectively. The inner number represents the number of *psbA* genes of that group of D1. No number means that there is only a single copy of that gene. Group 0 marks the atypical sequence only found in *Gloeobacter kilaueensis* JS-1.

but are not present in any other D1 sequence. In contrast, almost no sequence homology can be retrieved after the fifth transmembrane helix between any other D1 sequence and D2, although the protein folds are still relatively similar to each other. This may be taken as evidence that the very divergent D1 sequence in *G. kilaueensis* is indeed the earliest evolving D1 found within today's cyanobacterial diversity and therefore its phylogenetic position is not caused by long-branch attraction artifacts.

Immediately after the D1 sequences from *G. kilaueensis* we find another early branching group of D1 sequences, designated here Group 1 (G1, purple branches in fig. 1). This group is composed of eight sequences present in eight different species, one of which is the super-rogue D1 described by Murray (2012). This group contains all the D1 sequences found in the gene cluster associated with the far-red light acclimation response described by Gan et al. (2014). The sequence identity in this group ranges from 55% to 63% in comparison with the PsbA1 from *T. elongatus*. These super-rD1 or Group 1 sequences represent a well-supported group of related sequences, with severe changes around the water oxidizing complex and at the acceptor side. As is the case for the early branching sequence from *G. kilaueensis*, it is very unlikely that a Photosystem II containing a Group 1 sequence could support water oxidation and oxygen evolution as seen in conventional Photosystem II. Despite Group 1 constituting just a small fraction of the total number of D1 sequences, they are present in strains from all five sections of the traditional classification (Rippka et al. 1979).

Next, we observe a large monophyletic group of D1 sequences, also with divergent characteristics; here they will be designated Group 2 (G2, red lines in fig. 1). In this data set, 36 out of a total 360 cyanobacterial D1 sequences belong to Group 2. These sequences correspond to the rD1 sequences described by Murray (2012). Although they are less common than conventional D1 sequences, totaling 10% of the sequences, they are present in organisms from all five sections of the traditional cyanobacteria classification too (Rippka et al. 1979), including the unicellular *Synechococcus* sp. JA-3-3Ab and JA-2-3B′a, some *Cyanothece* strains as well as some *Acaryochloris*, some filamentous nondifferentiating strains such as *Leptolyngbia* sp. PCC 7375 and *Halothece* sp. PCC 7418, and also a few heterocyst-forming cyanobacteria such as *Fischerella* sp. JSC-11 and *A. variabilis*. The sequences belonging to Group 2 have a sequence identity ranging from 63% to 73% when compared with the PsbA1 sequence from *T. elongatus*, and also have significant changes both at the donor and acceptor side, which should modify the normal functioning of a fully assembled Photosystem II.

The next step in the evolution sees the appearance of water oxidation activity as we know it in conventional Photosystem II, because from this point onward every sequence has all the required ligands to assemble a Mn_4CaO_5 cluster (fig. 1 orange, green and blue branches). These D1 sequences that can support conventional water oxidation are the vast majority in the data set, making a total of 87.5% of the sequences. They all cluster together in a well-supported clade, yet there are some distinct subclades. The earliest branching of these are designated here as Group 3 (G3, fig. 1, orange branches). This group is composed of 39 sequences and includes some sequences that have been shown to be expressed under microaerobic conditions: such us the PsbA2 from *T. elongatus*, the PsbA0 from *Nostoc* sp. PCC 7120, and the PsbA1 from *Synechocystis* sp. PCC 6803. Therefore, it is possible that Group 3 represents an early form of D1 capable of water splitting that arose when little or no oxygen was present in the atmosphere, and thus shows optimal function under low oxygen concentrations in extant cyanobacteria. In a previous study, Sicora et al. (2009) suggested based on a few available D1′ sequences that these were characterized by the presence of three particular amino acid changes that are not shared with the well-known D1 form: G80, F158, and T289 changed to A80, L158, and A286 in the D1′, respectively. The sequence alignments revealed that indeed, all Group 3 sequences have these amino acids in common and, with a few specific exceptions, they are not shared with the standard D1. Remarkably, although Group 3 sequences are more like the dominant form of D1, they also share these substitutions with the early branching Group 1, and Group 2 sequences.

The latest evolving group of D1 is the well-characterized D1 form (G4, fig. 1, blue lines), such as that present in the crystal structures of Photosystem II. This group will be called Group 4 and every single genome from a cyanobacterium so far sequenced has at least one Group 4 sequence (fig. 2). The only exception is the symbiotic cyanobacterium

(a)
```
N. 7120    D2    159 YPLGQSSWFFAPSFGVAAIFRFLLFLQGFHNWTLNPFHM 197
T. elon    D2    160 YPLGQSSWFFAPSFGVAAIFRFLLFFQGFHNWTLNPFHM 198

G. kila    G0-D1  161 YPIGQGAFASGMPLGITGTFNWMLQLQAEHDILNHPLHM 199

C. ther    G1-D1  183 YPVGQGFSAGMPLGISGTFNFMLRFQADHNILMSPFHV 221
F. JSC11   G1-D1  183 YPVGQGSLSAGMPLGISGTFHFMLQFQADHNILMNPLHQ 221

C. ther    G2-D1  164 YPIGQGSFADGLPLGIAGTFNFMLAVQADHNILMHPFHM 202
C. 51142   G2-D1  164 YPLGQGSFSEGMPLGISGTFHFMLAFQADHNILMHPLHM 202

T. elon    G3-D1  161 YPIGQGSFSDGLMLGISGTFNFMIVFQAEHNILMHPFHM 199
N. 7120    G3-D1  161 YPIGQGSFSDGMMLGISGTFNFMIVFSPEHNILMHPFHM 199
S. 7002    G3-D1  161 YSIGQGSFSDGLPLGISGTFNFMFVLQAEHNVLMHPFHM 199

S. 6803    G4-D1  161 YPIGQGSFSDGMPLGISGTFNFMIVFQAEHNILMHPFHM 199
T. elon    G4-D1  161 YPIGQGSFSDGMPLGISGTFNFMIVFQAEHNILMHPFHQ 199
S. 7942    G4-D1  161 YPIGQGSFSDGMPLGISGTFNFMFVFQAEHNILMHPFHM 199
                       Yz/Yᴅ        •              •        P_{D1}/P_{D2}
```

(b)
```
N. 7120    D2    233 AFNPTQSEETYSMVTAN--RFWSQIF-GIAFSNKRWLHF 268
T. elon    D2    234 AFNPTQAEETYSMVTAN--RFWSQIF-GIAFSNKRWLHF 269

G. kila    G0-D1  231 AAGAGDLIETYSLWGGY-NYVWRLAISPVGLVNPRTIAA 267

C. ther    G1-D1  263 GYKLGQKRPTYSFRAAQ-LYLWRLIWRGTSFPNSRRLHF 300
F. JSC11   G1-D1  260 GYKLGQQHPTYNFRSAQ-VYLWHLIWHRVSFPNSRKLHF 297

C. ther    G2-D1  239 GYKLGQSQMTYHYLAGHYGFLGRLLVPWFASRNHRAFHF 277
C. 51142   G2-D1  239 GYHFGQQEATYNLVAGHLGYLGRLLIPSLGFKNSRSVHF 277

T. elon    G3-D1  236 GYKFGQEEETYNIVAAH-GYFGRLIFQYASFNNSRSLHF 273
N. 7120    G3-D1  236 GYKFGQEEETYNIVAAH-GYFGRLIFQYASFNNSRSLHF 273
S. 7002    G3-D1  236 GYRFGQEEETYNIVAAH-GYFGRLIFQYASFNNSRALHF 273

S. 6803    G4-D1  236 GYKFGQEEETYNIVAAH-GYFGRLIFQYASFNNSRSLHF 273
T. elon    G4-D1  236 GYKFGQEEETYNIVAAH-GYFGRLIFQYASFNNSRSLHF 273
S. 7942    G4-D1  236 GYKFGQEEETYNIVAAH-GYFGRLIFQYASFNNSRSLHF 273
                       Bicarb.           Q_B/Q_A      Fe²⁺
```

(c)
```
N. 7120    D2    318 ILLNEGIRAWMAPQDQPH-EKFVFPEEVLPRGNAL---- 351
T. elon    D2    319 LLLNEGIRAWMAPQDQPH-ENFVFPEEVLPRGNAL---- 352

G. kila    G0-D1  317 KRADAGIDGMMQPGTIAVADAVQYTAEWLPNRPLR---- 351

C. ther    G1-D1  349 DWANLGIDMARDRQLHQFPSDLMAVSNE---------- 376
F. JSC11   G1-D1  346 DWASLGIKILDERHIYDFPENLTAGEVVPWK-------- 376

C. ther    G2-D1  326 NRANLGIQAMHAVNTHHFPNIIAGGGIQPVNIASIL--- 361
C. 51142   G2-D1  326 NRANLGIRSMHAPNTHHFPLTLASGKATVIS-------- 356

T. elon    G3-D1  322 NRANIGIEVMHERNAHNFPLDLASGELAPVAMIAPSIEA 360
N. 7120    G3-D1  322 NRANLGIEVMHERNAHNFPLDLASGEVQPIALTAPAIAS 360
S. 7002    G3-D1  322 NRANLGIEAMHERNVHNFPLDLAAGDQAPVALQAPAING 360

S. 6803    G4-D1  322 NRANIGFEVMHERNAHNFPLDLASGEQAPVALTAPAVNG 360
T. elon    G4-D1  322 NRANLGMEVMHERNAHNFPLDLASAESAPVAMIAPSING 360
S. 7942    G4-D1  322 NRANLGMEVMHERNAHNFPLDLAAGEATPVALTAPAING 360
                        ••        • •
```

Fig. 3. Sequence alignment of selected regions from a number of representative D1 and D2 sequences. (*a*) Sequence alignment showing the region from the redox tyrosines, Y_Z and Y_D, to the chlorophyll ligands, P_{D1} and P_{D2}. To the right of the alignment, the corresponding region in the structure of Photosystem II is shown (PDB ID: 3WU2). The amino acids that have been marked in bold in the alignment are shown as sticks in the structure. (*b*) Sequence alignment showing the acceptor side of D1 and D2 and corresponding structural region. (*c*) Sequence alignment of the C-terminus of D1 and D2 and corresponding structural region. Amino acids marked with dots show those that offer ligands to the Mn_4CaO_5 cluster in standard D1. *N.* 7120 denotes *Nostoc* sp. PCC 7120; *T. elon*, *Thermosynechococcus elongatus* BP-1; *G. kila*, *Gloeobacteri kilaueensis* JS-1; *C. ther*, *Chroococcidiopsis thermalis* PCC 7203; *F.* JSC-11, *Fischerella* sp. JSC-11; *C.* 51142, *Cyanothece* sp. ATCC 51142; *S.* 7002, *Synechococcus* sp. PCC 7002; *S.* 6803, *Synechocystis* sp. PCC 6803; and *S.* 7942, *Synechococcus* sp. PCC 7942.

Candidatus Atelocyanobacterium thalassa (UCYN-A), which has lost all Photosystem II genes (Zehr et al. 2008). Group 4 includes all the sequences known to be expressed under high-light conditions; however, "high-light" isoforms do not make a consistent clade, instead it seems clear that they have originated multiple times after relatively recent events of gene duplications, and therefore they represent repetitive cases of convergent evolution (supplementary fig S2, Supplementary Material online). Within Group 4 sequences, the earliest branching D1 belonged to the genus *Gloeobacter*, followed next by the early branching *Synechococcus* sp. JA-3-3Ab and JA-2-3B′a. After these, we observe an expansion of D1 forms covering the full diversity of cyanobacteria.

Similar topologies for the phylogeny of D1 were retrieved with different phylogenetic methods (supplementary fig. S3, Supplementary Material online). In addition, Bayesian inference assuming heterogeneity in the substitution process across sites and over time (CAT + GTR + Γ) was performed to test for the possibility of long-branch attraction artifacts. In all cases, the divergent sequence from *G. kilaueensis* was the earliest branching, followed by Group 1 and Group 2, before the evolution of the D1 forms with conventional ligands to the Mn_4CaO_5 cluster. Therefore this topology seems to be robust. The results suggest that the origin of the divergent D1 forms predates the evolution of the dominant D1 proteins in Photosystem II and suggests a concise evolutionary scenario for the origin of oxygenic photosynthesis.

Characteristics of Atypical D1 Sequences

As we have seen, there are at least three types of early branching D1 proteins that seem to have evolved before the well-known forms of D1. These are: the earliest diverging form of D1 only found in *G. kilaueensis*, Group 1 sequences (super-rogue), and Group 2 sequences (rogue). These sequences together will be referred to as atypical D1 while those from Group 3 and Group 4, characterized by a complete set of ligands to the Mn_4CaO_5 cluster, will be referred to as standard D1. For clarity and in order to relate amino acid positions to the structural data available (Umena et al. 2011), we will refer to structurally homologous positions among the different D1 sequences using the numbering of the *T. elongatus* PsbA1 protein. The reader should be aware that numerous gaps and insertions exist among the entire diversity of D1 sequences, shifting the actual numbering for many of the individual sequences. Here, we shall describe some of the more obvious amino acid variations that seem likely to be relevant to the structure and function of the variant forms of Photosystem II that would result from the expression of these atypical D1 subunits.

Differences around the Mn₄CaO₅ Cluster

The ligands to the Mn_4CaO_5 cluster are conserved in all standard D1 sequences (Group 3 and Group 4). In contrast, all atypical D1 sequences show multiple changes in these key groups, figure 3. Indeed the rD1 protein sequences as a group was originally defined based on the lack of conservation to the ligands to the Mn_4CaO_5 cluster (Murray 2012). In the original work, the changes in the ligands and their potential impact where described in some detail. However, the additional sequences included here make it worthwhile reiterating and reassessing these important changes. Figure 4 shows structural homology models of the Mn_4CaO_5 cluster ligands for the atypical sequences in comparison with the standard D1 (Umena et al. 2011).

Aspartate 170 (D170), which provides a bidentate ligand bridging Mn4 and the Ca^{2+} and which is considered to be the high affinity site for Mn^{2+} binding during photoactivation (Diner and Nixon 1992; Nixon and Diner 1992), is a serine in the atypical D1 from *G. kilaueensis*. In Group 1 it is found as alanine or serine, with glutamate, a potential ligand, present in only one sequence. In the majority of Group 2 sequences it is found as glutamate, with aspartate only present in a few sequences. The change from D170 to E170 still allows for assembly of the cluster and similar rates of O_2 evolution have been measured in site-directed mutants (Nixon and Diner 1992). On the other hand, mutants where D170 was mutated to either alanine or serine were unable to assemble a Mn_4CaO_5 cluster or evolve O_2 (Diner and Nixon 1992; Nixon and Diner 1992). This suggest that while the most ancient forms of atypical D1 may not be able to bind a metal cluster, Group 2 D1 might still allow the assembly of a metallic cluster of some sort.

Glutamate 189 (E189), which in standard D1 provides a single ligand to Mn1, is still conserved in the anoxygenic variant from *G. kilaueensis*. In Group 1 and most Group 2 sequences E189 is found as aspartate, although there is some

variation in the latter group, also N, R, Q, or A can be found in some of the sequences. In total 17 mutants have been constructed at this position in *Synechocystis* 6803 including all the changes mentioned above (Debus 2001) and of those, only E189Q and E189R have been shown, surprisingly, to retain some O_2 evolution and photoautotrophic growth. The conservative change from glutamate to aspartate unexpectedly abolished water oxidation (Chu et al. 1995a; Clausen et al. 2001).

Histidine 332 (H332), a ligand to Mn1 in all standard D1 is found as methionine in the sequence from *G. kilaueensis*, it is also absent in all Group 1 sequences with changes to Q, G, R, S, and D. In contrast, Group 2 sequences also have a histidine in this position, with only two exceptions out of 36 sequences, where a tyrosine and asparagine are found instead. Most mutations at this position inhibit photoautotrophic growth with the exceptions of H332Q and H332S, where only 10–15% of the wild-type O_2 evolution rate was measured (Nixon and Diner 1994; Chu et al. 1995b). Given the amino acid differences described above, it is unlikely that the atypical D1 from *G. kilaueensis* and from Group 1 can bind a Mn cluster that resembles that in the standard protein.

Glutamate 333 (E333), a bidentate ligand that bridges Mn4 and Mn3 in standard D1, is found as a glutamine in the sequence from *G. kilaueensis*. It is found as aspartate or glutamate in Group 1 sequences, but this ligand is absent in all Group 2 sequences, where it is found as alanine in the majority of the sequences, less commonly as serine, and in one exceptional case as glycine, none of which can fulfill the liganding role of E333. Indeed the E333A mutant has been constructed in *Synechocystis* 6803 and it is incapable of photoautotrophic growth and cannot assemble a Mn_4CaO_5 cluster (Chu et al. 1995b).

Aspartate 342 (D342), a bidentate carboxylate ligand that bridges Mn1 and Mn2 in conventional D1, is an alanine in the atypical sequence from *G. kilaueensis*, in Group 1 sequences it is found mostly as asparagine, while in Group 2 there is a wide range of amino acids, mostly but not exclusively with nonpolar side chains (L, V, M) that cannot play a similar liganding role. In a few sequences N, S, or T is present. At least five mutants at this position have been constructed in *Synechocystis* 6803, all rendering the organism unable to grow photoautotrophically, with the exception of the conservative D342E mutation (Nixon and Diner 1994; Chu et al. 1995b).

Alanine 344 (A344) is the C-terminal amino acid of processed D1 and the terminal carboxylate itself provides a bidentate ligand that bridges Mn2 and the Ca^{2+}. A344 is found as a glutamine in the atypical sequence from *G. kilaueensis*. In Group 1 sequences a T, D, or K occupy the equivalent position. In Group 2 alanine is conserved in a number of sequences but also serine is common. It is worth mentioning, that for some Group 2 sequences this alanine is the last amino acid in the sequence. In standard D1, a C-terminal extension is cleaved at A344 as part of the assembly of the Photosystem II complex and this is required to allow full assembly of the Mn_4CaO_5 cluster (Nixon et al. 1992). The protease responsible for this is specific for alanine, so
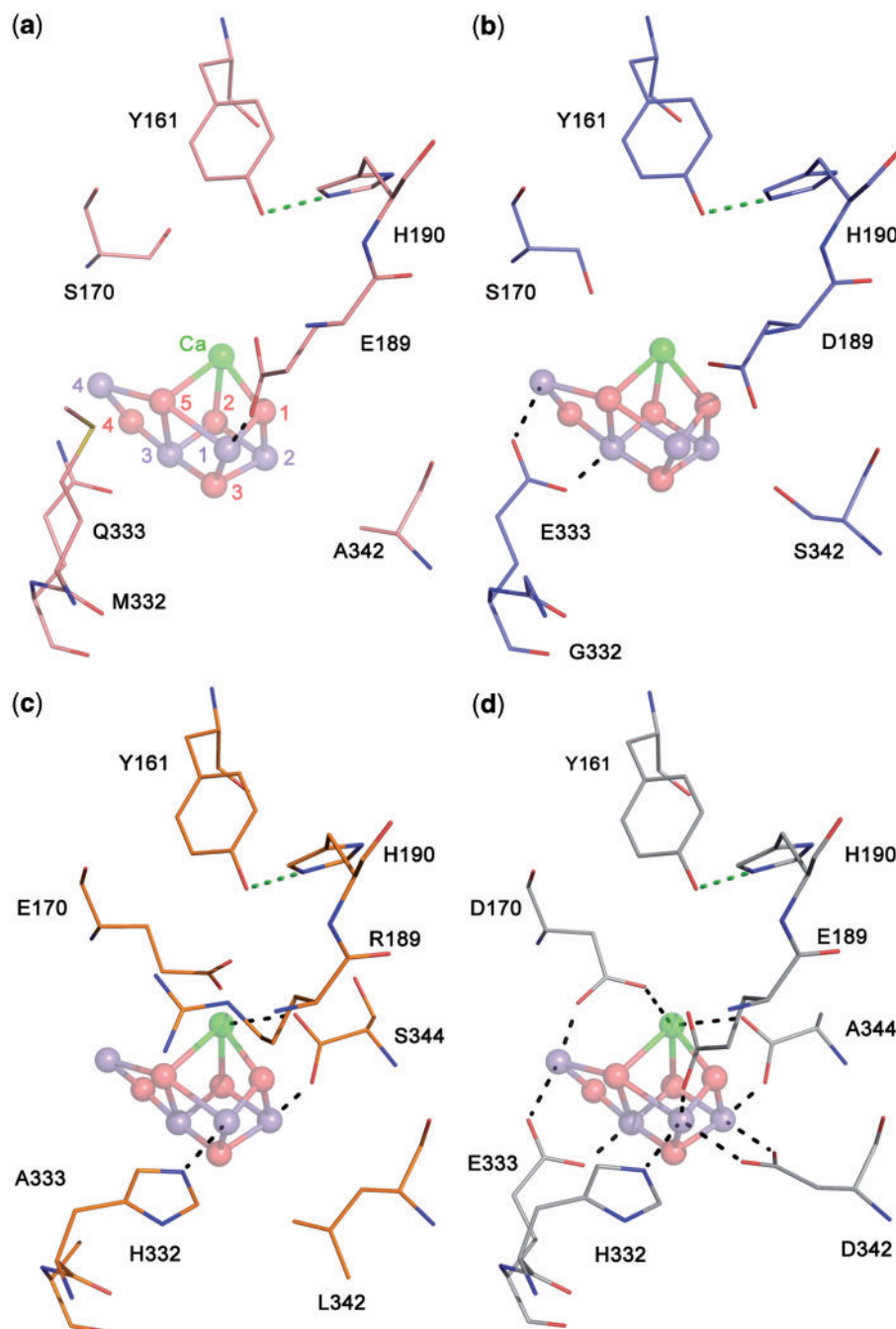
Fig. 4. Molecular models of the ligand sphere of the Mn_4CaO_5 cluster in the atypical D1 forms. (a) A model of the atypical D1 from *G. kilaueensis*, assuming that the C-terminus could fold in a D1-like manner. (b) A model from a Group 1 D1 and (c) a Group 2 D1 from *Synechococcys* sp. PCC 7335. (d) The water oxidizing complex of Photosystem II from *Thermosynechococcus vulcanus* at 1.9 Å resolution. Numbers in purple and red represent the numeration for Mn and O atoms respectively.

the absence of alanine at this position, along with the wide range of variations in the normal conserved sequence prior to the alanine, may indicate that atypical D1 proteins are not processed or at least not processed in the same way. This could mean that the terminal bidentate ligand to the Mn cluster would be absent in most of the sequences. Lack of processing of D1 results in loss of water oxidation and there are some reports that suggest there is residual binding of one or two Mn ions (Metz and Bishop 1980; Roose and Pakrasi 2004).

Histidine 337 (H337) donates an H-bond to O3, at the corner of the cubane part of the cluster and is strictly conserved in every standard sequence and also in Group 2 sequences. It is absent in the sequence from *G. kilaueensis*, where an alanine is found instead. In Group 1, there is a tyrosine at this position in most of the sequences, but also glutamate and proline in two of them. The H337Y and H337E mutations have been studied before and rendered the organism unable to grow photoautotrophically. Only a small fraction of O_2 evolution was detected in the H337E mutant (Chu et al. 1995b).

Differences around Channels

Key residues that could play a role in proton or water channels show variation in the earliest evolving D1 proteins. In the second coordination sphere on the Mn_4CaO_5 cluster, D61 is found within 4 Å of O4 and is conserved in every standard D1 sequence. It has been proposed based on mutagenesis studies that this residue modulates the redox state of the higher states of the water oxidation cycle (Qian et al. 1999), but in the light of the crystal structure (Umena et al. 2011) it has been suggested to be involved in substrate access to the cluster (Vassiliev et al. 2012; Debus 2014). It is conserved in the atypical sequence from *G. kilaueensis* and in Group 1 sequences. It is also conserved in some Group 2 sequences, but glutamate at this position is more common. The D61E mutant has been constructed before, and although it did not significantly alter the rate of O_2 evolution, it destabilized the intermediates of photoactivation (Qian et al. 1999).

Glutamate 65 (E65) has been identified by computational studies as being directly involved in a hydrogen bonding network for the release of protons during the catalytic cycle (Ishikita et al. 2006; Murray and Barber 2006; Linke and Ho 2013). This position is an arginine in the atypical sequence from *G. kilaueensis* and in Group 1 sequence it is R, S, A, or V. On the other hand, E65 is conserved in all Group 2 sequences and in every standard D1. The mutant E65A studied in *Synechocystis* 6803 showed only 20% of the wild-type O_2 evolution activity (Chu et al. 1995a).

Umena et al. (2011) noted a chain of H-bonded water molecules bridged by N298 and N322 leading from Y_Z-H190 toward the lumen and proposed that it could be a proton exit path. This proposed path implied a mechanism for water oxidation that echoed the hydrogen atom transfer model of Hoganson and Babcock (1997). However, Linke and Ho (2013) argued against this proposal because the side-chain of asparagine is not ionizable, which is consistent with current views on the mechanism of water oxidation that suggests the proton shared by the Y_Z-H190 couple is not translocated to the lumen, but hops between the two residues as Y_Z is oxidized and then rereduced. The positive charge that is retained on H190, near Y_Z, modulates its potential making it positive enough to drive the oxidation of water. Consequently, removal of the proton would lead to a loss of this oxidation potential (Siegbahn and Blomberg 2010). For a recent review see Linke and Ho (2013). In the atypical D1 from *G. kilaueensis*, N298 is found as glycine and N322 as lysine. In Group 1 sequences, N298 is found as glutamate or aspartate, and N322 as aspartate. On the other hand, these two positions are conserved in all Group 2 sequences and every standard D1. Kuroda et al. (2014) recently showed that every mutation at position N298 severely impaired water oxidation, and only 7 out of 19 mutants could grow photoautotrophically, albeit at low light intensities. The N298G mutant showed less than 20% of the wild-type O_2 evolution and the N298E or N298D mutant had no O_2 evolution at all.

In Group 1 sequences the presence of ionizable groups, E298 or D298, in addition to D322, should allow the deprotonation of Y_Z and the migration of the positive charge to the

lumen. In the sequence from *G. kilaueensis*, the small side chain of G298 could allow a water molecule to provide a H-bond to H190 and K322 is ionizable, so the deprotonation of Y_Z through this path is also possible. This suggests that Y_Z oxidation in the earlier versions of Photosystem II could have involved deprotonation via this pathway. This may be taken as evidence that at the evolutionary stage when Group 1 sequences first appeared, water oxidation may not have yet evolved, because the $Y_Z^{\bullet}$ radical was not oxidizing enough to support water oxidation (due to the evacuation of the proton). Alternatively, if at this evolutionary stage water oxidation was possible, it may have occurred by a mechanism involving hydrogen atom abstraction as suggested by Hoganson and Babcock (1997). If this is so, then at the evolutionary stage when Group 2 sequences appeared for the first time, the proton pathways have already started to take the current form, even in the absence of all ligands to the Mn_4CaO_5 cluster. Furthermore, this suggests that the apparent path noted by Umena et al. (2011) could be a relic proton-path that was blocked as an adaptation to increase the oxidizing potential of Y_Z.

Asparagine 335 (N335) was proposed by Linke and Ho (2013) to be involved in a different proton path. In this case N335 is not required to be deprotonated as it has a structural role in a tight network of hydrogen bonds that includes a few water molecules. Similar to the cases above, this position is a glycine in the atypical sequence from *G. kilaueensis* and it is found as P, H, Q in Group 1 sequences. It is, however, strictly conserved in Group 2 sequences and in standard D1. The mutant N335R was constructed by Linke and Ho (2013) and was severely impaired in photoautotrophic growth and O_2 evolution.

Other residues have also been suggested to have roles in proton pathways in D1: for example, E329 (Service et al. 2010) and R334 (Linke and Ho 2013). E329 is an aspartate in the sequence from *G. kilaueensis*; lysine or histidine in Group 1 sequences; and it is found mostly as glutamine or arginine in Group 2. R334 is found as isoleucine in *G. kilaueensis*, it is conserved as arginine in Group 1 sequences, but it is found mostly as proline in Group 2, although V, L, M, and I are also found in a few sequences. The mutants E329Q, R334K, and R334N have been studied and shown impaired water oxidation reactions, although they did not completely abolished O_2 evolution (Service et al. 2010; Linke and Ho 2013). This effect was interpreted as a disruption of the proton pathways.

It seems clear that proton or water channels in the atypical D1 sequence from *G. kilaueensis* and Group 1 sequences differ significantly from those in standard D1 forms. The key residues that define the paths in standard D1 start to appear for the first time in Group 2, but only reach their final form in Group 3 and Group 4 sequences.

Differences at the Acceptor Side

In addition to the differences near the Mn_4CaO_5 cluster, which are described above and in Murray (2012), we observed functionally significant variation in the atypical D1 sequences around the acceptor side.

Glutamine or glutamate at position 130 (Q130 or E130) provides an H-bond with different strengths to $Pheo_{D1}$ thereby modulating its potential. Q130 gives rise to a low potential $Pheo_{D1}$ that is present in the dominant form of D1 expressed under the light intensities normally used to maintain laboratory cultures. Under high light, variants of D1 are expressed with E130 resulting in a higher potential $Pheo_{D1}$ that confers additional photoprotection (Rappaport et al. 2002; Cser and Vass 2007; Sugiura et al. 2010). All standard D1 proteins belonging to Group 3 and 4 contain one or the other of these options (supplementary fig. S2, Supplementary Material online). The atypical sequence from *G. kilaueensis* has a leucine at this position, thus the H-bond would be absent. In most Group 1 sequences, it is found as glutamate, nominally the high-light form. In most sequences in Group 2 the H-bond is absent due to substitution to L, F, or M; however, glutamine is found in a few sequences. The lack of an H-bond would be expected to give rise to a very low potential $Pheo_{D1}$, as shown for mutants constructed in *Synechocystis* 6803 (Giorgi et al. 1996; Merry et al. 1998; Cser and Vass 2007). The variation observed within these atypical D1 sequences implies that they probably have specialized to fulfill quite different roles.

There is significant variation in the stromal loop surrounding the nonheme Fe^{2+} in all atypical D1 forms. Variations occur in the 241-QEEE motif, which provides coordination to the bicarbonate (E244) and a negatively charged patch on the stromal surface of Photosystem II. In the earliest diverging D1 sequence from *G. kilaueensis* this stromal loop seems to be shorter than in any other D1. Moreover, in all Group 1 and in 30 out of 36 Group 2 sequences E244 is substituted by A, P, V, K, or M. This means that Photosystem II containing an atypical D1 may lack bicarbonate, or at least be modified in terms of bicarbonate function. It is noteworthy that anoxygenic Type II reaction centers do not bind bicarbonate. Instead, coordination to the nonheme iron is provided by a glutamate from the M subunit (Cardona et al. 2012).

Serine 264 (S264) provides an H-bond to Q_B through the hydroxyl group (figs. 3 and 5). It is found as glycine in the ancient sequence from *G. kilaueensis*. In Group 1, this serine is conserved in seven out of eight sequences. In Group 2 sequences, this position is found as alanine or glycine. Both S264A and S264G mutations are found in naturally occurring herbicide-resistant strains in plants, and the phenotype shows slower rates of electron transfer from Q_A to Q_B due to a marked increase in the dissociation constant of Q_B (Crofts et al. 1987; Petrouleas and Crofts 2005). Many mutants have been constructed at this position already decades ago in a large variety of organisms with similar effects, reviewed by Oettmeier (1999).

Phenylalanine 265 (F265) provides an H-bond to Q_B through the backbone carbonyl. It is found as leucine in the sequence from *G. kilaueensis*. It is conserved in 7 out of 8 Group 1 sequences, and in 12 out of 36 Group 2 sequences, otherwise it is found as S, W, M, L, or V. In theory, any amino acid could provide an H-bond through the backbone carbonyl. Consequently, it is difficult to predict how these
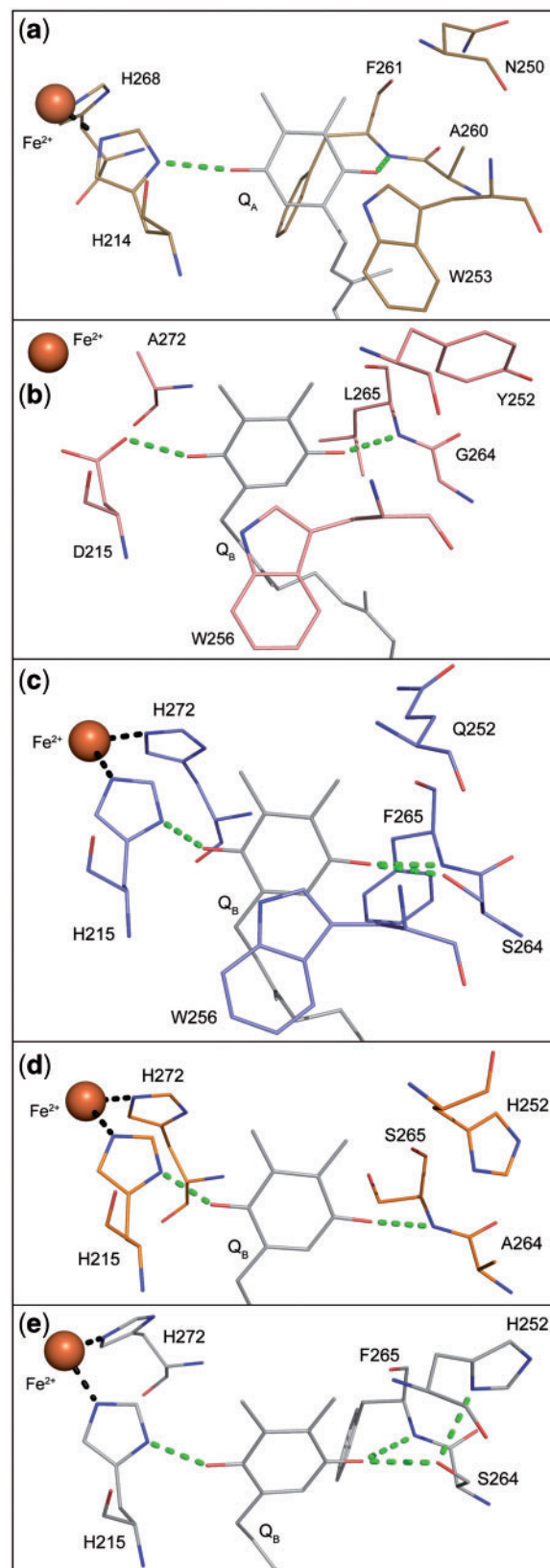
Fig. 5. A comparison of the acceptor side region of Photosystem II in (*a*) the D2 subunit from the crystal structure of *Thermosynechococcus vulcanus*, (*b*) a model of the atypical D1 from *Gloeobacter kilaueensis* JS-1, (*c*) a Group 1 and (*d*) a Group 2 D1 from *Chroococcidiopsis thermalis* PCC 7203, and (*e*) the D1 from the crystal structure of *T. vulcanus*.

variations could alter the binding of Q_B. To our knowledge no mutants at this position have been constructed yet.

Histidine 252 (H252) is thought to be involved in the protonation of Q_B (Crofts et al. 1987; Ishikita and Knapp 2005; Petrouleas and Crofts 2005; Saito et al. 2013). It is absent in the atypical sequence from *G. kilaueensis* where it is found as a tyrosine. It is also absent in Group 1 sequences where it is found as glutamine except in one case where it is an arginine. However, H252 is conserved in all Group 2 and all standard D1 from Group 3 and 4. Several mutants at this position have been constructed before to study the protonation mechanism of Q_B, including H252Q and H252K, none of which could grow photoautotrophically and presented altered rates of electron transfer from Q_A to Q_B (Petrouleas and Crofts 2005). Furthermore, there is a unique insertion in all Group 2 sequences at a position equivalent to G253. The exact amino acid inserted at this position varies and can be found as A, V, L, T, N, Y, or E. Changes to H-bonding networks, especially those directly associated with protonation reactions are expected to modify electron transfer (Crofts et al. 1987; Ishikita and Knapp 2005; Saito et al. 2013). Again, a pattern appears where the earliest branching forms of D1, that from *G. kilaueensis* and Group 1 sequences, differ in key residues involved in protonation reactions. However, in Group 2 sequences, despite having modified donor and acceptor sides, these amino acids are more like in standard D1. If the current model for the mechanism of Q_B protonation is correct (Saito et al. 2013), the difference at this position in the early evolving atypical sequences might imply that protonation of Q_B occurs by a different mechanism.

In addition, the sequence alignment showed that the atypical sequence from *G. kilaueensis* and Group 1 sequences have a tryptophan at position 256 (figs. 3 and 5). This tryptophan is homologous to W253 strictly conserved in all D2 sequences and helps to keep Q_A fixed in position by π-stacking of the indole ring with the plastoquinone head group, but occurring in an offset-stacked manner (Loll et al. 2005; Guskov et al. 2009; Hasegawa and Noguchi 2014; Lambreva et al. 2014). In all standard D1 and also Group 2 sequences, glycine is found at this position. Mutants constructed at D2-W253 have shown severe perturbations of Q_A function. The D2-W253L mutant has been constructed in *Synechocystis* 6803 and was unable to growth photoautotrophically (Vermaas et al. 1990). The mutant D2-W253F was shown to grow photoautotrophically but an increased yield of Q_A^- recombination was observed (Vermaas 1993). Rutherford and Faller (2003) proposed that an ancestral homodimeric Photosystem II could have a quinone binding site with mixed characteristics between Q_A and Q_B. This finding is in agreement with that hypothesis and with the topology of the phylogenetic tree that places the atypical sequence from *G. kilaueensis* and Group 1 basally. It may be, thus, that these early branching sequences descended from an ancestral D1 that still retained some of those hybrid and ancestral functional traits.

At this stage it is very difficult to predict how the unique properties of the atypical sequences affect quinone reduction,

but the mechanism should be different to that in standard D1. The differences seen around the Q_B site might be related to the binding of an alternative quinone. In fact, *G. violaceous* has been shown to use menaquinone as the secondary electron acceptor of Photosystem I (Mimuro et al. 2005) and something similar could be happening in Photosystem II, as both strains from the genus *Gloeobacter* lack several genes of conventional quinone synthesis pathways and might lack the capacity to synthetize plastoquinone-9 (Nakamura et al. 2003; Saw et al. 2013). It could also be that under conditions where the atypical sequences are or could be expressed (e.g., far-red light, anaerobic conditions) the redox potential of plastoquinone is not well-suited, and therefore a different quinone may be used. Alternatively, the differences in the atypical D1 sequences could result in unconventional electron exit pathways. For example, an electron transfer exit pathway from Q_A to a soluble electron acceptor was engineered by a single amino acid mutation on the stromal side of D1 (Larom et al. 2010) and by the addition of an exogenous electron acceptor in the presence of the herbicide (3-(3,4-dichlorophenyl)-1,1-dimethylurea) DCMU (Sugiura and Inoue 1999; Ulas and Brudvig 2011). An alternative electron transfer pathway from Q_B involving flavodiiron proteins has also been reported recently, but yet to be characterized in detail (Zhang et al. 2012).

Finally, some exceptional amino acid substitutions that should drastically affect function are found in a few sequences. In the atypical sequence from *G. kilaueensis* both ligands to the nonheme Fe^{2+} (H215 and H272) are absent. H215 is found as aspartate and H272 as alanine (fig 5). H215 is also lacking in Group 2 sequences from strains of *Crocosphaera watsonii*, in which it mutated to glutamine. Mutations to a ligand of the iron rendered the cyanobacterium unable to grow photoautotrophically and electron transport from Q_A to Q_B was completely inhibited (Vermaas et al. 1994). Therefore, in these cases, it is very unlikely that if the protein is incorporated into a Photosystem II complex, it could coordinate the nonheme Fe^{2+} or bind Q_B. Similarly, H198, which coordinates P_{D1}, is conserved in all D1 except in Group 2 sequences from *Synechococcus* JA-2-3B′a and JA-3-3Ab where it mutated to serine, and in the sequence from *Synechococcus* sp. PCC 7336 where it mutated to glutamine. It is worth noting that mutants of the P_{D1} chlorophyll ligand have also been constructed in *Synechocystis* 6803, namely H198A and H198K, and at least the latter mutant has been shown to grow photoautotrophically, albeit with modified energetics of charge separation (Giorgi et al. 1996; Merry et al. 1998; Cser and Vass 2007). These are clearly secondary derived mutations and should not be considered ancestral as all Type II reaction centers must have originated from completely functional systems. In consequence, even if the atypical sequence from *G. kilaueensis* have accumulated derived mutations that could inactivate Photosystem II, and even if the protein is not use in Photosystem II at all today, it should have originated from a functional D1 protein predating the evolution of the standard forms.

D1 in Different Organisms of Interest

The genome of *T. elongatus* contains three D1 forms, PsbA1 (dominant D1), PsbA2 (microaerobic), and PsbA3 (high-light). The phylogenetic tree shows that both PsbA1 and PsbA3 belong to Group 4 sequences, while PsbA2 belongs to Group 3. It appears however that the change from Q130 to E130 (low potential to high potential $Pheo_{D1}$) and also from E130 to Q130 have occurred multiple times during the evolution of cyanobacterial D1 proteins (supplementary fig. S2, Supplementary Material online). In fact, a glutamate or glutamine is also found in a homologous position in the L subunit of Type II reaction centers from photosynthetic proteobacteria, depending on the species (Michel et al. 1986). This implies that the modulation of the potential of pheophytin, by varying the strength of the H-bond, is a widespread adaptation strategy in all Type II reaction center containing organisms. In a recent article and using a smaller data set not including the atypical sequences, Vinyard et al. (2014) suggested that a few amino acid positions correlated with the presence of either Q130 and E130; namely, positions 124, 151, 152, and 288. We find that these correlations are not strict, but some do apply for a subset of sequences. These positions have been described and discussed in supplementary figure S2, Supplementary Material online.

Synechocystis 6803 also contains three D1 forms. PsbA2 (*slr1311*) and PsbA3 (*sll1867*) are identical sequences, and these are the dominant form of D1 expressed under standard laboratory conditions; with PsbA2 making 90% of the transcript while PsbA3 is responsible for up to 10%. Both become overexpressed under high-light intensity. PsbA1 (*slr1181*) is considered to be the slightly more divergent form, expressed only under microaerobic conditions, and is designated as a D1′ (Summerfield et al. 2008; Mulo et al. 2009; Sicora et al. 2009). PsbA2 and PsbA3 belong to Group 4 and the microaerobic form belongs to Group 3.

The genome of *Gloeobacter violaceus* PCC 7421 has five copies of D1, three of which are identical proteins and annotated as PsbAI, PsbAII, and PsbAIII. PsbAI and PsbAII have been shown to be dominantly expressed under low-light conditions and under high-light an enhancement in expression of these three genes was measured (Sicora et al. 2008). PsbAIV and PsbAV make the two other distinct D1 proteins and are slightly more divergent and thus presumed to be equivalent to D1′ (Sicora et al. 2008; Mulo et al. 2009). However, all five D1 sequences have the E130 version of $Pheo_{D1}$ and all five of them belong to Group 4. Interestingly, these are the earliest branching sequences within Group 4. The genome of *G. kilaueensis* also contains the same five genes as *G. violaceious*, but it has an additional sixth *psbA* gene. This extra gene encodes the atypical D1 that we have discussed above.

Within the heterocystous cyanobacteria, *Nostoc* sp. PCC 7120 also has five *psbA* genes encoding three distinct D1 proteins. One form of D1 contains the low-light D1 variant (Q130), while the three identical copies of D1 contain the high-light form (E130). Both low-light and high-light forms belong to Group 4. The third form (PsbA0 or *alr3742*) is slightly more divergent, and it has been shown to be transcribed under microaerobic conditions, thus a D1′ (Sicora et al. 2009). This D1′ belongs to Group 3. *Anabaena variabilis*, a similar heterocystous cyanobacterium to *Nostoc* 7120 also has equivalent *psbA* genes; however, it has an additional sixth gene encoding a Group 2 D1 protein. This atypical D1 was shown to be expressed under heterotrophic conditions in the dark (Park et al. 2013). Another heterocystous cyanobacterium, *Calothrix* sp. PCC 7507 also has a similar set of genes as *A. variabilis*, but instead of having a Group 2 sequence it has a Group 1. *Fischerella* sp. JSC-11, a true-branching heterocystous cyanobacterium, has six *psbA* genes including both a Group 1 and a Group 2 sequence. In fact, six out of the eight organisms containing a Group 1 sequence in our data set, also have a Group 2 sequence.

Most photosynthetic eukaryotes on the other hand have a single *psbA* gene in the chloroplast genome, with a few exceptions (Erickson et al. 1984; Lidholm et al. 1991). To explore the evolutionary position of D1 found in photosynthetic eukaryotes, 20 representative D1 sequences were included in the analysis from diverse groups (see fig. 1 and supplementary fig. S1, Supplementary Material online). These include sequences from plants, green and red algae, glaucophytes, cryptophytes, stramenopiles, alveolates, euglenids, haptophytes, and a sequence from the chromatophore of the amoeboid *Paulinella chromatophora* known to have acquired its "plastid" in an independent primary endosymbiotic event (Marin et al. 2005). All eukaryotic sequences, with the exception of that from *P. chromatophora*, grouped as a monophyletic clade branching early within Group 4, after the divergence of Group 4 sequences from both species of *Gloeobacter*, and together with the sequence of the early diverging species *Synechococcus* sp. JA-2-3B′a, JA-3-3Ab, and PCC 7336 (Turner et al. 1999; Shih et al. 2013; see fig. 1 and supplementary fig. S1, Supplementary Material online). There is very little variation among eukaryote D1: for example, the D1 from the red algae *Cyanidioschyzon merolae* shares 88% sequence identity with the D1 present in *Arabidopsis thaliana* and 87% identity with that in *Synechococcus* sp. JA-2-3B′a. In other words, the D1 found in photosynthetic eukaryotes has hardly changed since the primary endosymbiotic event. This is in agreement with the monophyly of the primary photosynthetic eukaryotes (Rodriguez-Ezpeleta et al. 2005) and suggests that the D1 retained in the plastid genomes is more similar to that of Group 4 from early branching unicellular cyanobacteria. The origin and evolution of plastids have been reviewed at length before, see, for example, Gould et al. (2008). Our phylogenetic tree for D1 suggests that the cyanobacterium that became the primary endosymbiont first branched after the last common ancestor of the phylum Cyanobacteria, but before the major cyanobacterial radiation (Turner et al. 1999; Rodriguez-Ezpeleta et al. 2005; Criscuolo and Gribaldo 2011; Shih et al. 2013). This contrasts, however, with the recent proposal that the plastid ancestor might have been more closely related to heterocystous cyanobacteria (Deusch et al. 2008; Dagan et al. 2013; Ochoa de Alda et al. 2014). On the other hand, the D1 sequence in *P. chormatophora* grouped together with that of *Synechococcus* sp. WH 5701 and *Cyanobium gracile* PCC 6307 (see supplementary fig.

S1, Supplementary Material online), in a major clade of Group 4 sequences that included most strains of *Synechococcus* and *Prochlorococcus*, consistent with previous work (Marin et al. 2005; Yoon et al. 2009).

Discussion

Evolution of the D1 Proteins of Photosystem II

The phylogenetic analysis presented here suggests that the last common cyanobacterial ancestor (LCCA) contained several forms of D1, the products of gene duplications of an ancestral *psbA* gene, and each of these could give rise to Photosystem II adapted to specific conditions. The data suggests that the LCCA must have had a fully functional Photosystem II containing a Group 4 D1 sequence. It was therefore capable of reducing quinone and oxidizing water using an exchangeable Q_B site and a standard oxygen evolving complex, very similar to those present in conventional Photosystem II existing today. This is because every cyanobacterium known carries in its genome, at the very least, one copy of a Group 4 D1. The only exception is the secondary loss of Photosystem II by UCYN-A. This implies that the Group 1 to Group 3 sequences, and including the atypical sequence from *G. kilaueensis*, had to have originated by gene duplications before the LCCA. It is impossible to determine how many times the ancestral D1 protein duplicated before the LCCA, but the data speak of at least five events of gene duplications prior to the appearance of Group 4 sequences. Perhaps one of the most important results derived from the present analysis is that there are many D1 sequences that might have originated before the evolution of water oxidation as we know it, in an organism that preceded the LCCA. Moreover, the branching pattern of the D1 evolutionary tree reveals a sequence of events that allows the reconstruction of some of the steps from anoxygenic photosynthesis to a water oxidizing Photosystem II like that known today.

The earliest events in the origin of photosynthesis are the evolution of (bacterio)chlorophyll synthesis and the appearance of the first reaction center protein that could be assembled into a photochemical reaction center (Raymond et al. 2004; Bryant and Liu 2013). A gene duplication of this primordial reaction center protein led to the divergence of Type I and Type II reaction centers in their homodimeric forms (Cardona 2014). A second gene duplication allowed the ancestral Type II reaction center protein to diverge into two types; one is ancestral to the L and M subunits found in phototrophic bacteria of the phyla Proteobacteria and Chloroflexi, the other was ancestral to D1 and D2 in a bacterial lineage that would lead to the Cyanobacteria. A third gene duplication eventually led to the divergence of D1 from D2. The ancestral gene encoding a primitive D1 duplicated several times, with each duplication event allowing for the acquisition of further complexity, until the evolution of water oxidation (Beanland 1990; Nitschke and Rutherford 1991; Rutherford and Faller 2003; Hohmann-Marriott and Blankenship 2011).

When comparing the D1 proteins in the early evolving sequences, it is sometimes difficult to determine which characteristics are ancestral and which are derived secondary mutations. However, because the L and M subunits of anoxygenic Type II reaction centers share a significant level of homology with D1 and D2, it is possible to pick out with confidence some of the traits that must have been ancestral. It appears, therefore, that the anoxygenic forms of D1 and D2 might have been primed for the binding of metals, even before D1 and D2 diverged into two different proteins; in other words, in the most primordial homodimeric Photosystem II ancestral complex. This is because it is quite likely that in the ancestral Type II reaction center there was a glutamate at position 170, the high-affinity Mn binding site in oxygenic Photosystem II. There is a glutamate at the homologous position in the M subunit found in the reaction centers from the Proteobacteria and Chloroflexi (see fig. 6), with the sequence from *Roseiflexus* spp. as the exception. There is also a glutamate present in the L subunit of the Chloroflexi, again with the exception of *Roseiflexus* spp. Glutamate is also present in a sequence in Group 1 and in most Group 2 sequences. This suggests that the presence of glutamate at position 170 precedes the appearance of the Y_Z-H190 couple in D1 and its homologous Y_D-H189 in D2. In the L and M subunit there is not an equivalent tyrosine; in contrast, there is a strictly conserved arginine at this position with only just a couple of exceptions. So, while E170 originated in the primordial Type II reaction center, the oxidizable tyrosine-histidine pair evolved in the ancestral homodimeric Photosystem II, before LCCA and before the evolution of water oxidation. When the oxidation potential of $P^{•+}$ became high enough to oxidize tyrosine, a metal near the vicinity of the newly formed tyrosyl radical could become oxidized and stabilized by the carboxylic group of the glutamate residue already present in the reaction center. Because the redox active tyrosine and its hydrogen bonding partner histidine are strictly conserved in every D1 and D2 sequence, this event must have happened in the homodimeric reaction centre (Rutherford and Nitschke 1996; Rutherford and Faller 2003). The key question is, what selective pressures were driving the potential of $P^{•+}$ to become more oxidizing? This is open for debate, but one possibility could be the depletion of electron donors that were more redox accessible in the local environment (Rutherford and Faller 2003).

There are four phenylalanines in the D2 protein filling the cavity where the Mn_4CaO_5 cluster is located in the D1 protein (fig. 6). Two of these phenylalanines, F169 and F188, are in homologous positions to two of the direct ligands to the tetramanganese complex in D1, D170, and E189, respectively. It appears that these changes occurred specifically in the D2 protein to isolate the Y_D-H189 couple and to fill up the space occupied by the catalytic cluster on the D1 side and thus to hinder the binding of metals on the D2 side. For this reason, it seems likely that during the early stages of Photosystem II evolution, metal clusters of some sort might have been coordinated in both sides of the complex in a symmetrical fashion, as suggested previously (Rutherford and Nitschke 1996;
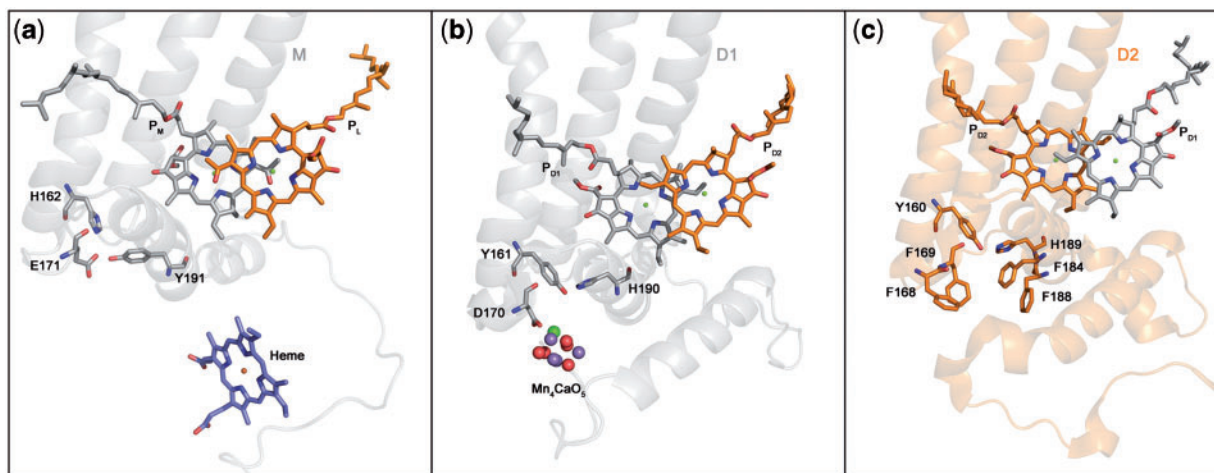
Fig. 6. Comparison of the donor side of the M subunit from *Blastochloris viridis* (PDB ID: 2PRC) with that of D1 and D2. (*a*) The M subunit from *B. viridis* is shown; E171 and H162 sit at a homologous position to (*b*) D170 and H190 of the D1 subunit and (*c*) F169 and H189 of the D2 subunit. Note that in the D2 subunit four phenylalanines are blocking the site where the Mn_4CaO_5 cluster is located in D1.

Rutherford and Faller 2003). These early clusters in the homo-dimeric Photosystem II were coordinated by at least, E170 and E189. Then, the cluster in the D2 side was lost as the complex became more asymmetrical (see Rutherford and Faller [2003] for a discussion of potential selection pressures for this), but the Y_D-H189 couple remained. It is possible that some of the sequences in Group 1 or Group 2 may still be able to assemble metal clusters that resemble those that preceded water oxidation.

From this point, the reaction center evolved toward the retention of Mn atoms at the donor side and the accumulation of positive charges in the initial cluster. This should have driven the evolution of the ligands which are located at the C-terminus of D1. In fact, it was recently shown that this terminal extension might have already been present in the most ancestral Type II reaction center protein, but had a different function in the most ancestral phototrophic bacteria (Cardona 2014). In the atypical sequence from *G. kilaueensis*, none of the ligands at the C-terminus is present, and taking into account the resemblance to D2 and its overall ancestral traits, it can be concluded that by the time this sequence diverged these ligands had not yet evolved. In Group 1 and Group 2 we see the appearance of some of the ligands at the C-terminus, but the full set of ligands is only found in Group 3 and Group 4. Yet, for water oxidation to occur in Photosystem II protonation reactions are also essential both at the donor and acceptor side. The exact mechanisms of proton transfer are not yet fully understood, although great progress has been made since crystal structures became available (Ishikita et al. 2006; Murray and Barber 2006; Umena et al. 2011; Linke and Ho 2013; Saito et al. 2013). The data here showed that in the most ancestral forms of D1, not only the full set of ligands to the Mn_4CaO_5 cluster is incomplete, but the key residues important in the strict control of proton movement had not yet evolved. These start to appear for the first time in Group 2 sequences even though not all the ligands to the standard Mn_4CaO_5 cluster had appeared. In consequence, it seems that the proper protonation pathways

and proton control around Y_Z had to evolve before efficient water oxidation.

The transitional forms of Photosystem II that may have preceded the origin of water oxidation have been lost in all photosynthetic eukaryotes and some species of cyanobacteria: however, there is still a large number of species that retain ancestral forms of D1, such as those in Group 1 and Group 2. Some of these forms of D1 may still be incorporated into a photochemical reaction center and carry out functions that resemble those that existed before efficient oxygenic photosynthesis evolved. What were those ancestral functions of a Photosystem II before water splitting? The earliest reaction center probably performed the light-driven one-electron oxidation of Mn^{2+} to Mn^{3+}. If Mn^{3+} became trapped in the photosystem by binding to the carboxylate side chain of E170 and/or E189, a second oxidation to Mn^{4+} could have been possible. In fact, it was recently suggested based in geochemical evidence, the existence of a photosystem capable of oxidizing manganese, but without O_2 evolution (Johnson et al. 2013). In addition, during photoactivation of Photosystem II in a D170E mutant, Mn^{2+} is oxidized and stabilized as Mn^{4+} (Campbell et al. 2000). It is possible that the ancestral D1 protein that gave rise to the atypical D1 form from *G. kilaueensis* and Group 1 sequences could have supported this type of chemistry.

A more sophisticated, but still primitive, Photosystem II could have catalyzed two-electron chemistry. A popular hypothesis is the oxidation of hydrogen peroxide in a manner similar to catalase (Bader 1994; Rutherford and Nitschke 1996; Blankenship and Hartman 1998). Alternatively, the oxidation of water to hydrogen peroxide is also plausible and it has been observed in Cl^--depleted and Ca^{2+}-depleted Photosystem II (Fine and Frasch 1992; Hillier and Wydrzynski 1993; Klimov et al. 1993; Arato et al. 2004). It is likely that the ancestral D1 protein that gave rise to Group 2 sequences could have supported this type of chemistry, now involving proton movement at the donor side. Nevertheless, it is possible that early transitional forms were capable of some

inefficient water oxidation, even with a mononuclear manganese cluster, and even perhaps in the ancestral homodimeric Photosystem II. In a recent paper, it was shown that a synthetic organometallic compound containing only one Mn atom could perform catalase activity, but with just slight modification of the ligand sphere it could oxidase water (Lee et al. 2014).

Photosystem II Containing Atypical D1 Forms in Extant Cyanobacteria

Questions arise concerning the functional and mechanistic properties of Photosystem II containing an atypical D1 form in cyanobacteria today. The atypical sequence from *G. kilaueensis* is the most different of all D1 sequences. It is characterized by lacking both ligands to the nonheme Fe^{2+}. It should also produce a very low potential pheophytin (L130). In addition, it lacks all ligands to the Mn_4CaO_5 cluster with the exception of E189. On the other hand, it has a conserved Y_Z-H190 couple and a ligand to P_{D1}. If it is incorporated into a Photosystem II complex, this should be neither active in water splitting nor quinone reduction. Today no experimental studies exists that could provide any information about the possible function of this protein and whether it is incorporated into a reaction center or not.

In Group 1 sequences we also observe significant changes around the Q_B site, yet they are likely to still be able to bind a quinone although with altered properties particularly for the protonation reactions (no H252 and altered or no bicarbonate binding). They have the high potential form of pheophytin (E130) and lack the ligands needed for assembling the Mn_4CaO_5 cluster. Last, these sequences appear to lack conventional proton exit pathways. However, this does not rule out completely the possibility that a cluster of some sort could be assembled as there are still several amino acids in the vicinity that could provide ligands.

Gan et al. (2014) recently showed that some cyanobacteria contain a cluster of approximately 21 genes encoding alternative copies of many reaction center proteins, including additional D1, D2, CP43, and CP47 subunits of Photosystem II, and also variant forms of PsaA and PsaB reaction center proteins of Photosystem I, among other subunits. They demonstrated that under prolonged far-red light exposure, an acclimation response was activated that led to the enhanced expression of this gene cluster, the remodeling of both photosystems and the phycobilisome, and the synthesis of chlorophylls *d* and *f*. This gene cluster contains two *psbA* genes, one encoding an atypical D1 from Group 1 and the second encoding a standard D1 from Group 4. This suggests that a Photosystem II carrying a Group 1 D1, containing chlorophylls *d* and *f*, might be adapted to function under far-red light. Gan et al. (2014) showed that the far-red acclimated cells were able to perform water oxidation efficiently, but did not show whether this activity originated from a Photosystem II with a Group 1 D1. Given the properties of the Group 1 D1 it would be surprising if it contributed to efficient water oxidation. Further experimental characterization of the

Photosystem II acclimated to far-red light is required to verify the function of Group 1 D1 proteins.

Group 2 sequences show a Q_B site that resembles those of herbicide resistant mutants. This could result in weakly bound semiquinone, which in turn could oxidize the nonheme Fe^{2+} more readily (Zimmermann and Rutherford 1986). In addition, as is the case for all atypical D1, Group 2 sequences do not appear to be able to coordinate bicarbonate in the same way as standard D1. Depletion or exchange of bicarbonate is known to slow down the rates of electron transfer from Q_A to Q_B and quinol/quinone exchange (Robinson et al. 1984; Sedoud et al. 2011; Muh et al. 2012; Shevela et al. 2012). Added to this, Group 2 sequences should produce a low or very low potential pheophytin (Q130, L130, or M130). A Photosystem II with the Q130L mutation has been shown to have a diminished yield of charge separation (Giorgi et al. 1996) and increased production of singlet oxygen (Vass 2011), which could suggest that this D1 is used under anaerobic conditions or under very low oxygen concentrations. At the donor side, Group 2 sequences contain either D170 or E170, so the oxidation and binding of Mn^{2+} ions at the site might occur (Debus 2001), even perhaps some inefficient water oxidation.

Some experimental evidence has been provided from transcriptomics about the possible role of Group 2 D1. In at least four different species containing a gene encoding a Group 2 D1, it was shown that the gene expression was enhanced in the dark (Toepel et al. 2008; Shi et al. 2010; Zhang and Sherman 2012; Park et al. 2013; Wegener et al. 2014). In unicellular diazotrophic cyanobacteria it was shown that the maximum transcription of the gene correlated with the time when nitrogenase genes were being expressed and translated (Toepel et al. 2008; Shi et al. 2010; Zhang and Sherman 2012; Wegener et al. 2014). In *Cyanothece* sp. ATCC 51142, the transcript abundance increased from less than 0.5% of all *psbA* genes up to 10% (Toepel et al. 2008). In a recent article, the Group 2 D1 sequence from *Cyanothece* 51142 was expressed in a *Synechocystis* 6803 mutant lacking all *psbA* genes (Wegener et al. 2014). Wegener et al. (2014) showed that *Synechosystis* 6803 could assemble a Photosystem II using the Group 2 D1 from *Cyanothece* 51142. They showed that the mutant with the Group 2 D1 could not grow photoautotrophically and had low oxygen evolution activity. Wegener et al. (2014) suggested that a Photosystem II containing a Group 2 D1 during the night could protect nitrogenase from oxygen production driven by moonlight and starlight. In the case of *A. variabilis*, the expression was enhanced after exposure to darkness in the presence of sucrose, and the gene expression was detected in both vegetative cells and heterocysts (Park et al. 2013), suggesting that the expression of the gene is not correlated to N_2 fixation per se, because vegetative cells do not contain nitrogenase. Thus, the role of Group 2 D1 bearing Photosystem II might be correlated with the metabolic state of the cells after prolonged darkness and may be useful with the onset of light after a long dark time, rather than for maintaining an anaerobic environment within the cell. In a fifth strain, a nondiazotrophic cyanobacterium, *Acaryochloris marina*, the expression of the gene was

triggered by incubation with an inhibitor of the Cytochrome b_6f complex under illumination (Kiss et al. 2012). The transcript abundance increased from 1% up to 50% of the total pool of transcripts for all *psbA* genes. In *A. marina*, the addition of DBMIB could have replicated a state of prolonged dark adaptation or could indicate that Group 2 sequences may be activated under a wider spectrum of conditions. It is not known at this moment if the gene encoding a Group 2 D1 in this strain is expressed during darkness.

In conclusion, we have investigated in detail the evolution and diversification of the D1 reaction center protein of Photosystem II in cyanobacteria. We showed that within the known diversity of cyanobacteria there are at least four major groups of D1 proteins, reflecting several ancestral duplications of the reaction center gene before their last common ancestor. The phylogenetic position of these diverse groups of sequences appears to reflect evolutionary transitions from an ancestral anoxygenic Type II reaction center to the origin of water oxidation. We suggested that the ancestral homodimeric Type II reaction center might have been capable of binding of metals. Thus some form of light-driven metal-based oxidative catalysis could have occurred early during the evolution of photosynthesis, after just a few changes to the amino acid sequence in a homodimeric Type II reaction center. A rudimentary version of water oxidation could have occurred well before the evolution of current sophisticated version of the enzyme. At this moment it is difficult to predict what is the function of the atypical versions of D1 and what sorts of advantages, if any, they confer to the organism. Purification and characterization of Photosystem II in which these ancestral D1 forms are present should not only prove useful in deciphering their roles, but also provide more insights on how oxygenic photosynthesis evolved.

Materials and Methods

All nonredundant PsbA (D1) and PsbD (D2) amino acid sequences were retrieved on January 2, 2014 using Position-Specific Iterated (PSI)-BLAST. To retrieve D1 and D2 subunits deposited in the RefSeq database, the D1 (accession number: YP_001864696, NpunF1022) protein from *Nostoc punctiforme* PCC 73102 was used as a search query. The search was restricted to the phylum Cyanobacteria, excluding environmental samples. Partial sequences were removed from the data set. In total, 490 sequences were retrieved, 360 D1 and 130 D2 subunits. 20 D1 sequences representing the diversity of photosynthetic eukaryotes were added to the data set. All the sequences used in this study are available in supplementary text S1 (Supplementary Material online).

Sequences were aligned using the Clustal Omega algorithm, with a maximum of 10^4 guide tree iterations and 10^7 Hidden Markov Model iterations (Sievers et al. 2011). All alignments are available on request. To confirm that the alignment was correct, the 3D structures of the D1 and D2 proteins (PDB ID: 3WU2) were overlapped using the CEalign (Jia et al. 2004) plug-in for PyMOL (Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC) and structural homologous positions were cross-checked with the alignment. Maximum likelihood phylogenies were constructed using

PhyML 3.1 (Guindon et al. 2010). Phylogenies were constructed using the LG model of amino acid substitution. The equilibrium frequencies and the proportion of invariant sites were set to be estimated by the software. Four gamma rate categories were used with the gamma shape parameter left to be calculated by the program. The nearest neighbor interchange method was used for tree improvement. Branch support was calculated with the approximate likelihood ratio test option (Anisimova and Gascuel 2006). Branch support of 0.7 (70%) was considered to be informative. Trees were plotted using Dendroscope 3.2.8 (Huson and Scornavacca 2011).

Parsimony, distance, and Bayesian analyses were performed in a subset of sequences and the trees are shown in the supplementary information (supplementary fig. S3, Supplementary Material online). Parsimony was calculated by randomizing the sequences ten times and with 1,000 bootstrap replicates. BioNJ was computed with 10^5 bootstrap replicates and with an observed distribution. Both Parsimony and BioNJ trees were done with Seaview 4.4.3 (Gouy et al. 2010). Bayesian analysis was performed with Phylobayes 3.3 (Lartillot and Philippe 2004) using the CAT mixture model to account for compositional heterogeneity across sites and applying relative exchange rates and four gamma rate categories (CAT + GTR + Γ). Four independent chains were run until convergence (26,000 cycles, maxdiff of 0.089). The first 5,200 trees were discarded as "burn-in," and the remaining trees from each chain were used to test for convergence and compute the majority rule consensus tree.

Structural homology models were generated using the SWISS-MODEL online service (http://swissmodel.expasy.org/, last accessed February 18, 2015; Arnold et al. 2006). To produce models for the atypical D1 forms, we selected the atypical sequence from *G. kilaueensis* JS-1 (WP_023174186.1), the Group 1 and Group 2 D1 sequences from *Croococcidiopsis thermalis* PCC 7203 (WP_015153111.1 and WP_015152761.1, respectively) and *Synechococcus* sp. PCC 7335 (WP_006456314.1 and WP_006458236.1). These were modeled to the D1 subunit from the crystal structure of Photosystem II from *T. vulcanus* at 1.9 Å (Umena et al. 2011). The amino acid numbering in the crystal structures were used throughout the text for clarity. Molecular models were visualized using PyMOL.

Supplementary Material

Supplementary figures S1–S3 and text S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

References

Anisimova M, Gascuel O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Syst Biol.* 55: 539–552.

Arato A, Bondarava N, Krieger-Liszkay A. 2004. Production of reactive oxygen species in chloride- and calcium-depleted Photosystem II and their involvement in photoinhibition. *Biochim Biophys Acta.* 1608:171–180.

Arnold K, Bordoli L, Kopp J, Schwede T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* 22:195–201.

Bader KP. 1994. Physiological and evolutionary aspects of the O_2/H_2O_2-cycle in Cyanobacteria. *Biochim Biophys Acta.* 1188:213–219.

Beanland TJ. 1990. Evolutionary relationships between Q-type photosynthetic reaction centers: hypothesis-testing using parsimony. *J Theor Biol.* 145:535–545.

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

Bryant DA, Liu ZF. 2013. Green bacteria: insights into green bacterial evolution through genomic analyses. In: Beatty JT, editor. Genome evolution of photosynthetic bacteria. Advances in botanical research. Vol. 66, San Diego (CA): Academic Press. p. 99–150.

Campbell KA, Force DA, Nixon PJ, Dole F, Diner BA, Britt RD. 2000. Dual-mode EPR detects the initial intermediate in photoassembly of the Photosystem II mn cluster: the influence of amino acid residue 170 of the D1 polypeptide on Mn coordination. *J Am Chem Soc.* 122: 3754–3761.

Cardona T. 2014. A fresh look at the evolution and diversification of photochemical reaction centers. *Photosynth Res.* Advance Access published December 16, 2014, doi: 10.1007/s11120-014-0065-x.

Cardona T, Sedoud A, Cox N, Rutherford AW. 2012. Charge separation in Photosystem II: a comparative and evolutionary overview. *Biochim Biophys Acta.* 1817:26–43.

Chu HA, Nguyen AP, Debus RJ. 1995a. Amino acid residues that influence the binding of manganese or calcium to Photosystem II. 1. the lumenal interhelical domains of the D1 polypeptide. *Biochemistry* 34:5839–5858.

Chu HA, Nguyen AP, Debus RJ. 1995b. Amino acid residues that influence the binding of manganese or calcium to Photosystem II. 2. The carboxy-terminal domain of the D1 polypeptide. *Biochemistry* 34: 5859–5882.

Clausen J, Winkler S, Hays AMA, Hundelt M, Debus RJ, Junge W. 2001. Photosynthetic water oxidation in *Synechocystis* sp PCC6803: mutations D1-E189K, R and Q are without influence on electron transfer at the donor side of Photosystem II. *Biochim Biophys Acta.* 1506:224–235.

Cox N, Pantazis DA, Neese F, Lubitz W. 2013. Biological water oxidation. *Acc Chem Res.* 46:1588–1596.

Criscuolo A, Gribaldo S. 2011. Large-scale phylogenomic analyses indicate a deep origin of primary plastids within cyanobacteria. *Mol Biol Evol.* 28:3019–3032.

Crofts A, Robinson H, Andrews K, Van Doren S, Berry E, 1987. Catalytic sites for reduction and oxidation of quinones. In: Papa S, Chance B, Ernster L, editors. Cytochrome systems: molecular biology and bioenergetics. New York: Plenum. p. 617–624.

Cser K, Vass I. 2007. Radiative and non-radiative charge recombination pathways in Photosystem II studied by thermoluminescence and chlorophyll fluorescence in the cyanobacterium *Synechocystis* 6803. *Biochim Biophys Acta.* 1767:233–243.

Dagan T, Roettger M, Stucken K, Landan G, Koch R, Major P, Gould SB, Goremykin VV, Rippka R, de Marsac NT, et al. 2013. Genomes of stigonematalean cyanobacteria (Subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. *Genome Biol Evol.* 5:31–44.

Dau H, Zaharieva I, Haumann M. 2012. Recent developments in research on water oxidation by Photosystem II. *Curr Opin Chem Biol.* 16:3–10.

Debus RJ. 2001. Amino acid residues that modulate the properties of tyrosine Y_Z and the manganese cluster in the water oxidizing complex of Photosystem II. *Biochim Biophys Acta.* 1503: 164–186.

Debus RJ. 2014. Evidence from FTIR difference spectroscopy that D1-Asp61 influences the water reactions of the oxygen-evolving Mn_4CaO_5 cluster of Photosystem II. *Biochemistry* 53:2941–2955.

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.

Diner BA, Nixon PJ. 1992. The rate of reduction of oxidized redox-active Tyrosine, Z^+, by exogenous Mn^{2+} is slowed in a site-directed mutant, at aspartate 170 of polypeptide D1 of Photosystem II, inactive for photosynthetic oxygen evolution. *Biochim Biophys Acta.* 1101: 134–138.

Dismukes GC, Klimov VV, Baranov SV, Kozlov YN, DasGupta J, Tyryshkin A. 2001. The origin of atmospheric oxygen on Earth: the innovation of oxygenic photosynthesis. *Proc Natl Acad Sci U S A.* 98:2170–2175.

Erickson JM, Rahire M, Rochaix JD. 1984. *Chlamydomonas reinhardii* gene for the 32 000 mol. wt. protein of Photosystem II contains four large introns and is located entirely within the chloroplast inverted repeat. *EMBO J.* 3:2753–2762.

Ferreira KN, Iverson TM, Maghlaoui K, Barber J, Iwata S. 2004. Architecture of the photosynthetic oxygen-evolving center. *Science* 303:1831–1838.

Fine PL, Frasch WD. 1992. The oxygen-evolving complex requires chloride to prevent hydrogen peroxide formation. *Biochemistry* 31: 12204–12210.

Flannery DT, Walter MR. 2012. Archean tufted microbial mats and the Great Oxidation Event: new insights into an ancient problem. *Aust J Earth Sci.* 59:1–11.

Gan F, Shen G, Bryant DA. 2014. Occurrence of far-red light photoacclimation (FaRLiP) in diverse cyanobacteria. *Life* 5:4–24.

Gan F, Zhang S, Rockwell NC, Martin SS, Lagarias JC, Bryant DA. 2014. Extensive remodeling of a cyanobacterial photosynthetic apparatus in far-red light. *Science* 345:1312–1317.

Giorgi LB, Nixon PJ, Merry SAP, Joseph DM, Durrant JR, Rivas JD, Barber J, Porter G, Klug DR. 1996. Comparison of primary charge separation in the Photosystem II reaction center complex isolated from wild-type and D1-130 mutants of the cyanobacterium *Synechocystis* PCC 6803. *J Biol Chem.* 271:2093–2101.

Gould SB, Waller RR, McFadden GI. 2008. Plastid evolution. *Annu Rev Plant Biol.* 59:491–517.

Gouy M, Guindon S, Gascuel O. 2010. SeaView Version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol.* 27:221–224.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 59:307–321.

Gupta RS. 2010. Molecular signatures for the main phyla of photosynthetic bacteria and their subgroups. *Photosynth Res.* 104:357–372.

Guskov A, Kern J, Gabdulkhakov A, Broser M, Zouni A, Saenger W. 2009. Cyanobacterial Photosystem II at 2.9-angstrom resolution and the role of quinones, lipids, channels and chloride. *Nat Struct Mol Biol.* 16:334–342.

Hasegawa K, Noguchi T. 2014. Molecular interactions of the quinone electron acceptors Q_A, Q_B, and Q_C in Photosystem II as studied by the fragment molecular orbital method. *Photosynth Res.* 120: 113–123.

Hillier W, Wydrzynski T. 1993. Increases in peroxide formation by the Photosystem II oxygen-evolving reactions upon removal of the extrinsic 16, 22 and 33 kDa proteins are reversed by $CaCl_2$ addition. *Photosynth Res.* 38:417–423.

Hoganson CW, Babcock GT. 1997. A metalloradical mechanism for the generation of oxygen from water in photosynthesis. *Science* 277: 1953–1956.

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

Huson DH, Scornavacca C. 2011. A survey of combinatorial methods for phylogenetic networks. *Genome Biol Evol.* 3:23–35.

Ishikita H, Knapp EW. 2005. Control of quinone redox potentials in Photosystem II: electron transfer and photoprotection. *J Am Chem Soc.* 127:14714–14720.

Ishikita H, Saenger W, Loll B, Biesiadka J, Knapp EW. 2006. Energetics of a possible proton exit pathway for water oxidation in Photosystem II. *Biochemistry* 45:2063–2071.

Jia YT, Dewey G, Shindyalov IN, Bourne PE. 2004. A new scoring function and associated statistical significance for structure alignment by CE. *J Comput Biol.* 11:787–799.

Johnson JE, Webb SM, Thomas K, Ono S, Kirschvink JL, Fischer WW. 2013. Manganese-oxidizing photosynthesis before the rise of cyanobacteria. *Proc Natl Acad Sci U S A.* 110:11238–11243.

Kiss E, Kos PB, Chen M, Vass I. 2012. A unique regulation of the expression of the *psbA*, *psbD*, and *psbE* genes, encoding the D1, D2 and Cytochrome *b*559 subunits of the Photosystem II complex in the chlorophyll *d* containing cyanobacterium *Acaryochloris marina*. *Biochim Biophys Acta.* 1817:1083–1094.

Klimov V, Ananyev G, Zastryzhnaya O, Wydrzynski T, Renger G. 1993. Photoproduction of hydrogen peroxide in Photosystem II membrane fragments: a comparison of four signals. *Photosynth Res.* 38:409–416.

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

Koyama K, Suzuki H, Noguchi T, Akimoto S, Tsuchiya T, Mimuro M. 2008. Oxygen evolution in the thylakoid-lacking cyanobacterium *Gloeobacter violaceus* PCC 7421. *Biochim Biophys Acta.* 1777:369–378.

Kulkarni RD, Golden SS. 1994. Adaptation to high light intensity in *Synechococcus* sp. strain PCC 7942: regulation of three *psbA* genes and two forms of the D1 protein. *J Bacteriol.* 176:959–965.

Kuroda H, Kodama N, Sun X-Y, Ozawa S-i, Takahashi Y. 2014. Requirement of asparagine 298 on D1 protein for oxygen evolution: analyses by exhaustive amino acid substitution in a green alga *Chlamydomonas reinhardtii*. *Plant Cell Physiol.* doi:10.1093/pcp/pcu073.

Lambreva MD, Russo D, Polticelli F, Scognamiglio V, Antonacci A, Zobnina V, Campi G, Rea G. 2014. Structure/function/dynamics of Photosystem II plastoquinone binding sites. *Curr Protein Pept Sci.* 15:285–295.

Larom S, Salama F, Schuster G, Adir N. 2010. Engineering of an alternative electron transfer path in Photosystem II. *Proc Natl Acad Sci U S A.* 107:9650–9655.

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

Lartillot N, Philippe H. 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol Biol Evol.* 21:1095–1109.

Lee WT, Munoz SB, Dickie DA, Smith JM. 2014. Ligand modification transforms a catalase mimic into a water oxidation catalyst. *Angew Chem Int Ed Engl.* 53:9856–9859.

Lidholm J, Szmidt A, Gustafsson P. 1991. Duplication of the *psbA* gene in the chloroplast genome of two *Pinus* species. *Mol Gen Genet.* 226:345–352.

Linke K, Ho FM. 2013. Water in Photosystem II: structural, functional and mechanistic considerations. *Biochim Biophys Acta.* 1837:14–32.

Loll B, Kern J, Saenger W, Zouni A, Biesiadka J. 2005. Towards complete cofactor arrangement in the 3.0 angstrom resolution structure of Photosystem II. *Nature* 438:1040–1044.

Lyons TW, Reinhard CT, Planavsky NJ. 2014. The rise of oxygen in Earth's early ocean and atmosphere. *Nature* 506:307–315.

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

Merry SA, Nixon PJ, Barter LM, Schilstra M, Porter G, Barber J, Durrant JR, Klug DR. 1998. Modulation of quantum yield of primary radical pair formation in Photosystem II by site-directed mutagenesis affecting radical cations and anions. *Biochemistry* 37:17439–17447.

Metz J, Bishop NI. 1980. Identification of a chloroplast membrane polypeptide associated with the oxidizing side of Photosystem II by the use of select low-fluorescent mutants of *Scenedesmus*. *Biochem Bioph Res Commun.* 94:560–566.

Michel H, Deisenhofer J. 1988. Relevance of the photosynthetic reaction center from purple bacteria to the structure of Photosystem II. *Biochemistry* 27:1–7.

Michel H, Epp O, Deisenhofer J. 1986. Pigment protein interactions in the photosynthetic reaction center from *Rhodopseudomonas viridis*. *EMBO J.* 5:2445–2451.

Mimuro M, Tsuchiya T, Inoue H, Sakuragi Y, Itoh Y, Gotoh T, Miyashita H, Bryant DA, Kobayashi M. 2005. The secondary electron acceptor of Photosystem I in *Gloeobacter violaceus* PCC 7421 is menaquinone-4 that is synthesized by a unique but unknown pathway. *FEBS Lett.* 579:3493–3496.

Muh F, Glockner C, Hellmich J, Zouni A. 2012. Light-induced quinone reduction in Photosystem II. *Biochim Biophys Acta.* 1817:44–65.

Mulo P, Sakurai I, Aro EM. 2012. Strategies for *psbA* gene expression in cyanobacteria, green algae and higher plants: from transcription to PSII repair. *Biochim Biophys Acta.* 1817:247–257.

Mulo P, Sicora C, Aro EM. 2009. Cyanobacterial *psbA* gene family: optimization of oxygenic photosynthesis. *Cell Mol Life Sci.* 66:3697–3710.

Murray JW. 2012. Sequence variation at the oxygen-evolving centre of Photosystem II: a new class of 'rogue' cyanobacterial D1 proteins. *Photosynth Res.* 110:177–184.

Murray JW, Barber J. 2006. Identification of a calcium-binding site in the PsbO protein of Photosystem II. *Biochemistry* 45:4128–4130.

Nakamura Y, Kaneko T, Sato S, Mimuro M, Miyashita H, Tsuchiya T, Sasamoto S, Watanabe A, Kawashima K, Kishida Y, et al. 2003. Complete genome structure of *Gloeobacter violaceus* PCC 7421, a cyanobacterium that lacks thylakoids. *DNA Res.* 10:137–145.

Nitschke W, Rutherford AW. 1991. Photosynthetic reaction centers: variations on a common structural theme. *Trends Biochem Sci.* 16:241–245.

Nixon PJ, Diner BA. 1992. Aspartate-170 of the Photosystem II reaction center polypeptide-D1 is involved in the assembly of the oxygen-evolving manganese cluster. *Biochemistry* 31:942–948.

Nixon PJ, Diner BA. 1994. Analysis of water-oxidation mutants constructed in the cyanobacterium *Synechocystis* sp. PCC 6803. *Biochem Soc Trans.* 22:338–343.

Nixon PJ, Trost JT, Diner BA. 1992. Role of the carboxy terminus of polypeptide-D1 in the assembly of a functional water-oxidizing manganese cluster in Photosystem II of the cyanobacterium *Synechocystis* sp. PCC 6803: assembly requires a free carboxyl group at C-terminal position 344. *Biochemistry* 31:10859–10871.

Ochoa de Alda JA, Esteban R, Diago ML, Houmard J. 2014. The plastid ancestor originated among one of the major cyanobacterial lineages. *Nat Commun.* 5:4937.

Oettmeier W. 1999. Herbicide resistance and supersensitivity in Photosystem II. *Cell Mol Life Sci.* 55:1255–1277.

Ogami S, Boussac A, Sugiura M. 2012. Deactivation processes in PsbA1-Photosystem II and PsbA3-Photosystem II under photoinhibitory conditions in the cyanobacterium *Thermosynechococcus elongatus*. *Biochim Biophys Acta.* 1817:1322–1330.

Park JJ, Lechno-Yossef S, Wolk CP, Vieille C. 2013. Cell-specific gene expression in *Anabaena variabilis* grown phototrophically, mixotrophically, and heterotrophically. *BMC Genomics* 14:759.

Petrouleas V, Crofts A. 2005. The iron-quinone acceptor complex. In: Wydrzynski TJ, Satoh K, editors. Photosystem II: the light-driven water:plastoquinone oxidoreductase. Dordrecht: Springer. p. 177–206.

Qian M, Dao LA, Debus RJ, Burnap RL. 1999. Impact of mutations within the putative Ca^{2+}-binding lumenal interhelical a-b loop of the Photosystem II D1 protein on the kinetics of photoactivation and H_2O-oxidation in *Synechocystis* sp. PCC 6803. *Biochemistry* 38:6070–6081.

Rappaport F, Guergova-Kuras M, Nixon PJ, Diner BA, Lavergne J. 2002. Kinetics and pathways of charge recombination in Photosystem II. *Biochemistry* 41:8518–8527.

Rasmussen B, Fletcher IR, Brocks JJ, Kilburn MR. 2008. Reassessing the first appearance of eukaryotes and cyanobacteria. *Nature* 455: 1101–1109.

Raymond J, Siefert JL, Staples CR, Blankenship RE. 2004. The natural history of nitrogen fixation. *Mol Biol Evol.* 21:541–554.

Rippka R, Deruelles J, Waterbury JB, Herdman M, Stanier RY. 1979. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *J Gen Microbiol.* 111:1–61.

Robinson HH, Eatonrye JJ, Vanrensen JJS, Govindjee. 1984. The effects of bicarbonate depletion and formate incubation on the kinetics of oxidation-reduction reactions of the Photosystem II quinone acceptor complex. *Z Naturforsch C* 39:382–385.

Rodriguez-Ezpeleta N, Brinkmann H, Burey SC, Roure B, Burger G, Loffelhardt W, Bohnert HJ, Philippe H, Lang BF. 2005. Monophyly of primary photosynthetic eukaryotes: green plants, red algae, and glaucophytes. *Curr Biol.* 15:1325–1330.

Roose JL, Pakrasi HB. 2004. Evidence that D1 processing is required for manganese binding and extrinsic protein assembly into Photosystem II. *J Biol Chem.* 279:45417–45422.

Rutherford AW, Faller P. 2003. Photosystem II: evolutionary perspectives. *Philos Trans R Soc Lond B Biol Sci.* 358:245–253.

Rutherford AW, Nitschke W. 1996. Photosystem II and the quinone–iron-containing reaction centers. In: Baltscheffsky H, editor. Origin and evolution of biological energy conversion. New York: VCH. p. 143–175.

Saito K, Rutherford AW, Ishikita H. 2013. Mechanism of proton-coupled quinone reduction in Photosystem II. *Proc Natl Acad Sci U S A.* 110: 954–959.

Sauer K, Yachandra VK. 2002. A possible evolutionary origin for the Mn_4 cluster of the photosynthetic water oxidation complex from natural MnO_2 precipitates in the early ocean. *Proc Natl Acad Sci U S A.* 99: 8631–8636.

Saw JHW, Schatz M, Brown MV, Kunkel DD, Foster JS, Shick H, Christensen S, Hou SB, Wan XH, Donachie SP. 2013. Cultivation and complete genome sequencing of *Gloeobacter kilaueensis* sp nov., from a lava cave in Kilauea Caldera, Hawai'i. *PLoS One* 8:e76376.

Sedoud A, Kastner L, Cox N, El-Alaoui S, Kirilovsky D, Rutherford AW. 2011. Effects of formate binding on the quinone-iron electron acceptor complex of Photosystem II. *Biochim Biophys Acta.* 1807: 216–226.

Service RJ, Hillier W, Debus RJ. 2010. Evidence from FTIR difference spectroscopy of an extensive network of hydrogen bonds near the oxygen-evolving Mn_4Ca cluster of Photosystem II involving D1-Glu65, D2-Glu312, and D1-Glu329. *Biochemistry* 49:6655–6669.

Shevela D, Eaton-Rye JJ, Shen JR, Govindjee.. 2012. Photosystem II and the unique role of bicarbonate: a historical perspective. *Biochim Biophys Acta.* 1817:1134–1151.

Shi T, Ilikchyan I, Rabouille S, Zehr JP. 2010. Genome-wide analysis of diel gene expression in the unicellular N_2-fixing cyanobacterium *Crocosphaera watsonii* WH 8501. *ISME J.* 4:621–632.

Shih PM, Wu D, Latifi A, Axen SD, Fewer DP, Talla E, Calteau A, Cai F, Tandeau de Marsac N, Rippka R, et al. 2013. Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proc Natl Acad Sci U S A.* 110:1053–1058.

Sicora CI, Brown CM, Cheregi O, Vass I, Campbell DA. 2008. The *psbA* gene family responds differentially to light and UVB stress in *Gloeobacter violaceus* PCC 7421, a deeply divergent cyanobacterium. *Biochim Biophys Acta.* 1777:130–139.

Sicora CI, Ho FM, Salminen T, Styring S, Aro EM. 2009. Transcription of a "silent" cyanobacterial *psbA* gene is induced by microaerobic conditions. *Biochim Biophys Acta.* 1787:105–112.

Siegbahn PE, Blomberg MR. 2010. Quantum chemical studies of proton-coupled electron transfer in metalloenzymes. *Chem Rev.* 110: 7040–7061.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li WZ, Lopez R, McWilliam H, Remmert M, Soding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 7:539.

Sugiura M, Azami C, Koyama K, Rutherford AW, Rappaport F, Boussac A. 2013. Modification of the pheophytin redox potential in *Thermosynechococcus elongatus* Photosystem II with PsbA3 as D1. *Biochim Biophys Acta.* 1837:139–148.

Sugiura M, Inoue Y. 1999. Highly purified thermo-stable oxygen-evolving Photosystem II core complex from the thermophilic cyanobacterium *Synechococcus elongatus* having his-tagged CP43. *Plant Cell Physiol.* 40:1219–1231.

Sugiura M, Kato Y, Takahashi R, Suzuki H, Watanabe T, Noguchi T, Rappaport F, Boussac A. 2010. Energetics in Photosystem II from *Thermosynechococcus elongatus* with a D1 protein encoded by either the *psbA1* or *psbA3* gene. *Biochim Biophys Acta.* 1797: 1491–1499.

Sugiura M, Ogami S, Kusumi M, Un S, Rappaport F, Boussac A. 2012. Environment of TyrZ in Photosystem II from *Thermosynechococcus elongatus* in which PsbA2 is the D1 protein. *J Biol Chem.* 287: 13336–13347.

Summerfield TC, Toepel J, Sherman LA. 2008. Low-oxygen induction of normally cryptic *psbA* genes in cyanobacteria. *Biochemistry* 47: 12939–12941.

Tichy M, Lupinkova L, Sicora C, Vass I, Kuvikova S, Prasil O, Komenda J. 2003. *Synechocystis* 6803 mutants expressing distinct forms of the Photosystem II D1 protein from *Synechococcus* 7942: relationship between the *psbA* coding region and sensitivity to visible and UV-B radiation. *Biochim Biophys Acta.* 1605:55–66.

Toepel J, Welsh E, Summerfield TC, Pakrasi HB, Sherman LA. 2008. Differential transcriptional analysis of the cyanobacterium *Cyanothece* sp. strain ATCC 51142 during light-dark and continuous-light growth. *J Bacteriol.* 190:3904–3913.

Turner S, Pryer KM, Miao VPW, Palmer JD. 1999. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small submit rRNA sequence analysis. *J Eukaryot Microbiol.* 46: 327–338.

Ulas G, Brudvig GW. 2011. Redirecting electron transfer in Photosystem II from water to redox-active metal complexes. *J Am Chem Soc.* 133: 13260–13263.

Umena Y, Kawakami K, Shen JR, Kamiya N. 2011. Crystal structure of oxygen-evolving Photosystem II at a resolution of 1.9 A. *Nature* 473: 55–60.

Vass I. 2011. Role of charge recombination processes in photodamage and photoprotection of the Photosystem II complex. *Physiol Plant.* 142:6–16.

Vassiliev S, Zaraiskaya T, Bruce D. 2012. Exploring the energetics of water permeation in Photosystem II by multiple steered molecular dynamics simulations. *Biochim Biophys Acta.* 1817:1671–1678.

Vermaas W. 1993. Molecular-biological approaches to analyze Photosystem II structure and function. *Annu Rev Plant Physiol Plant Mol Biol.* 44:457–481.

Vermaas W, Charite J, Shen GZ. 1990. Q_A binding to D2 contributes to the functional and structural integrity of Photosystem II. *Z Naturforsch C* 45:359–365.

Vermaas W, Vass I, Eggers B, Styring S. 1994. Mutation of a putative ligand to the nonheme iron in Photosystem II—implications for Q_A reactivity, electron-transfer, and herbicide binding. *Biochim Biophys Acta.* 1184:263–272.

Vinyard DJ, Gimpel J, Ananyev GM, Mayfield SP, Dismukes GC. 2014. Engineered Photosystem II reaction centers optimize photochemistry versus photoprotection at different solar intensities. *J Am Chem Soc.* 136:4048–4055.

Waldbauer JR, Sherman LS, Sumner DY, Summons RE. 2009. Late Archean molecular fossils from the Transvaal Supergroup record the antiquity of microbial diversity and aerobiosis. *Precambrian Res.* 169:28–47.

Wegener KM, Nagarajan A, Pakrasi HB. 2014. An atypical *psbA* gene encodes a sentinel D1 protein to form a physiologically relevant inactive Photosystem II complex in cyanobacteria. *J Biol Chem.* doi: 10.1074/jbc.M114.604124.

Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida K, Bhattacharya D. 2009. A single origin of the photosynthetic organelle in different *Paulinella* lineages. *BMC Evol Biol.* 9:98.

Zehr JP, Bench SR, Carter BJ, Hewson I, Niazi F, Shi T, Tripp HJ, Affourtit JP. 2008. Globally distributed uncultivated oceanic N_2-fixing cyanobacteria lack oxygenic Photosystem II. *Science* 322:1110–1112.

Zhang P, Eisenhut M, Brandt AM, Carmel D, Silen HM, Vass I, Allahverdiyeva Y, Salminen TA, Aro EM. 2012. Operon *flv4-flv2*

provides cyanobacterial Photosystem II with flexibility of electron transfer. *Plant Cell* 24:1952–1971.

Zhang X, Sherman LA. 2012. Alternate copies of D1 are used by cyanobacteria under different environmental conditions. *Photosynth Res.* 114:133–135.

Zimmermann JL, Rutherford AW. 1986. Photoreductant-induced oxidation of Fe^{2+} in the electron-acceptor complex of Photosystem II. *Biochim Biophys Acta.* 851:416–423.