Biochimica et Biophysica Acta 1787 (2009) 46–59

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbabio

Genetic analysis of the Photosystem I subunits from the red alga, *Galdieria sulphuraria*

Christopher Vanselow [a,1], Andreas P.M. Weber [b], Kirsten Krause [c], Petra Fromme [a,*]

[a] *Department of Chemistry and Biochemistry, Arizona State University, Tempe 85287-1604, AZ, USA*
[b] *Institute of Plant Biochemistry, Heinrich-Heine-University of Düsseldorf, Universitätsstraße 1, D-40225 Düsseldorf, Germany*
[c] *Department of Biology, University of Tromsø, 9037 Tromsø, Norway*

ARTICLE INFO

Article history:
Received 15 July 2008
Received in revised form 28 September 2008
Accepted 2 October 2008
Available online 22 October 2008

Keywords:
Photosynthesis
Red algae
Photosystem I
Evolution
Gene identification
Light harvesting
Electron transfer

ABSTRACT

Currently, there are very little data available regarding the photosynthetic apparatus of red algae. We have analyzed the genes for Photosystem I in the recently sequenced genome of the red alga *Galdieria sulphuraria*. All subunits that are conserved between plants and cyanobacteria were unambiguously identified in the *Galdieria* genome: PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaJ, PsaK and PsaL. From the plant specific subunits, PsaN and PsaO were identified but the sequence homology was much lower than for the subunits that are present in plants and cyanobacteria. The subunit PsaX, which is specific for thermophilic cyanobacteria, is not present in the *Galdieria* genome, whereas PsaM is a plastid-encoded protein as in other red algae. The sequences of the core subunits of PSI were further analyzed by mapping of the conserved areas in the crystal structures of cyanobacterial and plant PSI. The structural comparison shows that PSI from the red alga *Galdieria* may represent a common ancestral structure at the interface between cyanobacterial and plant PSI. Some subunits have a "zwitter" structure that contains structural elements that show similarities with either plant or cyanobacterial PSI. The structure of PsaL, which is responsible for the trimerization of PSI in cyanobacteria, lacks a short helix and the Ca^{2+} binding site, which are essential for trimer formation indicating that the *Galdieria* PSI is a monomer. However the sequence homology to plant PsaL is low and lacks strong conservation of the interaction sites with PsaH. Furthermore, the sites for interaction of plant PSI with the LHCI complex are not well conserved between plants and *Galdieria*, which may indicate that *Galdieria* may contain a PSI that is evolutionarily much more ancient than PSI from green algae, plants and the current cyanobacteria.

Published by Elsevier B.V.

1. Introduction

The production of our food supply, maintenance of life and the production of oxygen are all dependent on photosynthetic processes. Photosynthesis converts light energy into chemical energy, using light-driven transmembrane charge separation that couples the oxidation of water and formation of molecular oxygen (O_2) to the fixation of carbon dioxide (CO_2). The process is driven by Photosystems I and II (PSI and II), two large protein cofactor complexes located in the thylakoid membrane. This process simultaneously generates an aerobic atmosphere and produces a readily usable carbon source, sustaining almost all life on earth. Initially, the Earth's atmosphere was anoxygenic, and oxygen would have been toxic to most organisms on Earth at that time. About 2.5 billion years ago a series of evolutionary events led to the development of oxygenic photosynthesis by the ancestors of cyanobacteria, which contained two photosystems and used water as an electron source, thereby evolving oxygen as a by-product. This led to a change in the Earth's atmosphere, transforming it from anoxygenic to an oxygenic one. Cyanobacteria, eukaryotic algae, and higher plants are all currently capable of oxygenic photosynthesis. Photosynthesis in plants and algae occurs in organelles known as chloroplasts. Cyanobacteria are prokaryotes, capable of oxygenic photosynthesis. According to the Endosymbiosis hypothesis explaining the origin of chloroplasts, an ancestor of the cyanobacteria was engulfed by a eukaryotic cell, which subsequently led to the formation of chloroplasts [1]. Green algae, red algae, glaucophyta, and plants are the product of this primary endosymbiotic event and are thus called the Archaeplastida [2,3]. *Galdieria sulphuraria* belongs to the group of rhodophytes, and is thus a member of the Archaeplastida.

Photosynthesis can be divided into two major pathways: the light reactions and the dark reactions. The light reactions include the light induced splitting of water and electron transfer across the thylakoid membrane, ultimately resulting in the formation of the high energy products ATP and NADPH. ATP and NADPH are subsequently used in the dark reactions for carbohydrate synthesis from carbon dioxide.

While the central protein complexes that are involved in the light induced electron transfer are highly conserved between plants, algae

* Corresponding author. Tel.: +1 480 965 9028; fax: +1 480 965 2747.
E-mail address: pfromme@asu.edu (P. Fromme).
[1] Present address: Mayo Clinic, 13400 East Shea Blvd., Scottsdale, AZ 85259, USA.

0005-2728/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.bbabio.2008.10.004

and cyanobacteria, these organisms differ considerably in their peripheral light-harvesting antenna complexes. These antenna complexes are required to funnel photons into the PSI and PSII reaction centers, where the charge separation that drives photosynthesis occurs.

Photosystem I is a large, multi-subunit transmembrane protein complex responsible for catalyzing the light-induced electron transfer from plastocyanin or cytochrome c_6 at the luminal side of the membrane to ferredoxin or flavodoxin at the stromal side of the membrane. The electron is ultimately used to catalyze the reduction of $NADP^+$ and H^+ to NADPH.

In cyanobacteria, PSI is a trimer and each monomer consists of 12 subunits [4], whereas plant PSI is monomeric and consists of 15 subunits [5,6].

G. sulphuraria is a thermo- and acidophilic red microalga that inhabits extreme environments [7]. It belongs to an ancient group of extremophilic red algae, called the Cyanidiales. Cyanidiales are the principle photosynthetic organisms in hot acid: but also cyanobacteria are frequently found in hot environments such as hot springs in the Yellowstone National Park [8]. Cyanobacteria never occur at pH <4.0 — these very acidic environments are dominated by eukaryotic red microalgae (i.e., the Cyanidiales) that represent the principle photosynthetic organisms in hot acid. For example, native *Galdieria* growth conditions exist on the edges of hot springs with a pH of ~1.5 and temperatures of ~50 °C. [9]. In contrast to other members of the Cyanidiales, *G. sulphuraria* has the ability to grow either photoautotrophically (sunlight as energy source with carbon dioxide as the only source of carbon), photoheterotrophically (growth on organic carbon sources in the presence of light), or heterotrophically (growth on organic carbon sources in the dark) on more than 50 different sources of carbon [10,11]. *Galdieria* thrives at the upper temperature limit of eukaryotic live [12]. It not only tolerates these extreme temperatures, it actively grows under these conditions. As a bona fide eukaryote, the *Galdieria* genome encodes all principle components of eukaryotic cells, in particular those of mitochondria, the endomembrane system, and plastids [13]. To start to explore the potential of *Galdieria* for gaining a deeper understanding of the evolution of the eukaryotic photosynthetic machinery, we initiated a comparative study of the sequence and structural properties of the photosynthetic reaction centers of cyanobacteria, primitive red algae, and higher plants and algae.

The comparison of nucleotide and/or amino acid sequences from different organisms is a very powerful tool in molecular biology and biochemistry [14,15]. Using BLAST software from NCBI, known protein or nucleotide sequences can be used to search newly constructed databases to find areas of homology. By finding these regions of homology, evolutionary relationships between organisms may be determined [14,15]. Photosystem I is the only protein complex involved in the light reactions, for which structural information exists for both cyanobacteria and a higher plant [4], [16,5]. However, no structure exists from any eukaryotic algae. This paper shines new light onto the evolution of photosynthesis as it compares the sequences of the Photosystem I subunits of the red alga *G. sulphuraria* with the sequences of plant and cyanobacterial Photosystem I. The similarities and differences are discussed on the basis of the structures of plant and cyanobacterial Photosystem I. Similarities and differences between amino acids in structural and functionally important domains of the individual subunits lead to the prediction of the structural and functional features of the red algae Photosystem I (see Fig. 1). These studies give, for the first time, a strong indication for a chimerical nature of red algae PSI, which may indicate that *G. sulphuraria* represents an important evolutionary link between plants and cyanobacteria.

2. Materials and methods

Sequencing of the *G. sulphuraria* nuclear and chloroplast genomes was recently completed at Michigan State University [17,10], {http://genomics.msu.edu/galdieria/}. To identify PSI subunits from *Galdieria*, a local Blast database was constructed from the ~15 million nucleotides that make up the *Galdieria* genome. An online version of the database is publicly accessible at http://genomics.msu.edu/cgi-bin/galdieria/blast.cgi. The nucleotide data were organized into 432 individual supercontigs (stigs), or regions of organized text, to simplify identification of the nucleotide sequence after obtaining positive sequence searches from the database. The plastid genome was contained on stig_35 and stig_158 and consists of about 163 kilobases in total. Amino acid sequences from the known PSI subunits from *Thermosynechococcus elongatus* [18]{http://bacteria.kazusa.or.jp/cyanobase/}, *Arabidopsis thaliana* [19,20]{http://www.arabidopsis.org/} and other higher plants that were used in recent structural and functional studies [5,6] were used to search the translated *Galdieria* nucleotide database for orthologous sequences in the *Galdieria*

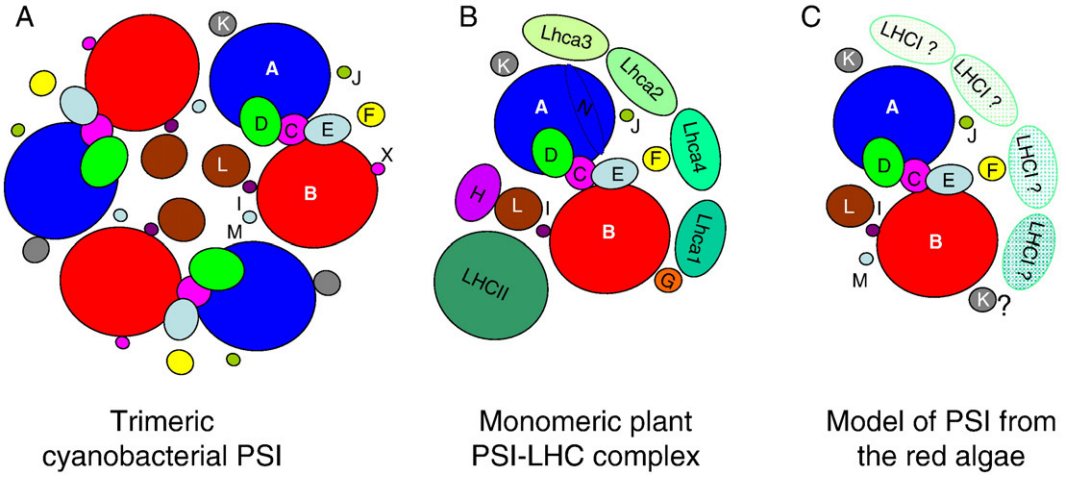

Fig. 1. Schematic representation of the structural arrangement of the Photosystem I subunits. (A) cyanobacteria and (B) plants. In (C) the proposed subunit composition and oligomeric structure of PSI from *Galdieria sulphuraria* based on the results of this work is depicted. Each subunit is represented by a circle that roughly reflects its position in the X-ray structures from the trimeric PSI from *Themosynecococcus elongatus* [4] in (A) and the PSI–LHCI complex from pea [5] in (B). In addition, the potential interaction site of PSI with the LHC II complex as derived from electron microscopy is shown in (B). The size of the circles is indicative of the size of the subunits, with PsaA and B represented by the largest circles and the smallest subunits that contain only one transmembrane helix (PsaI, J, M and X) represented by the smallest circle.

genome. This was done using the blastall and tblastn functions from the BLAST program. Alignment score and *E*-values were calculated as described previously [14,15].

Initial positive hits covering partial sequences were found for PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaG, PsaH, PsaI, PsaJ, PsaK, PsaL, PsaN and PsaO. Complete open reading frames (ORFs) for the corresponding genes including start and stop codons were obtained by subsequent manual curation of the sequence data identifying ORFs for the subunits PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaJ, PsaK and PsaL. All these subunits are well conserved between plants and cyanobacteria. In the case of the *psaI* coding sequence, only 23 amino acids belonging to the 5′end of the PsaI protein were encoded on stig_35 due to a gap in the sequence around position 141152. To close this gap and obtain the missing sequence information on PsaI, we performed a PCR on genomic DNA isolated from *Galdieria* using four different primers (forward: 5′-TCAAACAGGTAGAGTTCC-3′ and 5′-GATGGT-TAGCTGTTCATGC-3′; reverse: 5′-CCTTCAATACTAGTTCCG-3′and 5′-CAACTTGAGTGACACTTGG-3′) and sequenced the missing part.

Sequences related to the plant specific subunits PsaG and PsaH were found on stig_17 and stig_62, but the sequence similarities were much lower than for the rest of the sequences.

Amino acid sequence alignments were done with the individual PSI subunit sequences from different organisms and areas of sequence homology were color coded. The annotated amino acid alignment data were then used to identify and label the homologous amino acids in the known Protein Data Bank (PDB) structures from *T. elongatus* [4] (pdb accession code 1JB0) and angiosperm plant pea [5] (pdb accession code 2O01) using the modeling program PyMol. The individual amino acids in the 3-D structure were color coded with the colors listed in the figure legends to give a visual representation of the areas of sequence homology of individual domains of the proteins.

Theoretical masses and isoelectric points for the annotated *Galdieria* PSI subunits were determined using the compute pI/Mw program from ExPASy (Expert Protein Analysis System) (Table 2), as previously described in [21,22].

3. Results

The known PSI subunit amino acid sequences from *T. elongatus, A. thaliana,* pea and spinach were used to search the *G. sulphuraria* database to identify orthologous sequences. All subunits that are present in both cyanobacteria and plants have been identified in the genome of *G. sulphuraria*: PsaA, PsaB, PsaC, PsaD, PsaE, PsaF, PsaI, PsaJ, PsaK and PsaL. Furthermore, the cyanobacterial specific PsaM subunit as well as the plant specific subunit PsaN, were also identified. These results, complete with the location of the genes in the genome, are listed in Table 1. Most of these subunits were found on stig 35 and stig 158 which together represent the plastid genome. Only PsaN was found to be encoded by the nuclear genome and, accordingly, possesses a predicted target peptide, suggesting it might also be a functional Photosystem I subunit. Confirmation of positive hits included the verification of start and stop codons in a reasonable position when aligned with other known PSI subunit amino acid sequences and the presence of a significant sequence homology. Based on these criteria, the search for the plant specific subunits PsaG and PsaH generated no hits. Although stretches of sequence homology were found in the nuclear genome that were too high to be random, these were confined to the target peptide of the preprotein.

Table 1
PSI subunits identified in the genome of *Galdieria sulphuraria*

Subunit	Location	Start site	Stop site	# AA	Identity w. *A. thaliana*	% identity w. *T. elongatus*
PsaA	stig_35	92054	94367	770	605 AA (79%)	628 AA (82%)
PsaB	stig_35	94420	96624	734	591 AA (80%)	593 AA (81%)
PsaC	stig_35	21236	21481	81	71 AA (87%)	76 AA (93%)
PsaD	stig_35	99378	98963	141	84 AA (60%)	83 AA (59%)
PsaE	stig_35	66040	66235	65	36 AA (55%)	45 AA (69%)
PsaF	stig_35	145856	146425	189	87 AA (46%)	83 AA (44%)
PsaI	stig_35	141243	located in stig _35 sequence gap (see M+m)	36	14 AA (39%)	13 AA (36%)
PsaJ	stig_35	146474	146602	42	19 AA (45%)	29 AA (69%)
PsaK	stig_35	23729	23523	68	21 AA (31%)	22 AA (32%)
PsaL	stig_35	1399043	138618	141	66 AA (47%)	82 AA (48%)
PsaM	stig_158	5544	5455	29		14 AA (48%)
PsaN	stig_41	100512	100916	134	25 AA (19%)	
PsaO	stig_39	16421	16979	167	40 AA (24%)	

Table 2
Molecular weights (MW) and theoretical isoelectric point (pI) from PSI subunits from *A. thaliana*, *T. elongatus* and *G. sulphuraria*

Subunit	*A. thaliana*		*T. elongatus*		*G. sulphuraria*	
	MW [kDa]	pI	MW [kDa]	pI	MW [kDa]	pI
PsaA	83.2	6.60	84.8	7.32	86.0	8.86
PsaB	82.5	6.89	83.0	6.43	82.5	7.57
PsaC	9.0	6.68	8.8	5.65	8.8	5.65
PsaD	22.3 (*17.7)	9.81 (*9.33)	15.3	8.99	16.0	8.97
PsaE	15.2 (*10.5)	9.94 (*9.52)	8.4	9.52	7.5	8.96
PsaF	24.2 (*17.3)	9.58 (*9.34)	17.6	7.7	21.5	9.47
PsaI	4.1	10.0	4.3	3.79	4.05	3.67
PsaJ	5.0	5.88	4.8	6.69	4.8	5.88
PsaK	13.2 (*8.5)	10.52 (*8.21)	9.5	6.69	7.2	9,52
PsaL	23.1 (*18.0)	9.85 (*9.05)	16.3	5.16	15.4	6.54
PsaM			3.42	5.82	3.15	5.82
PsaN	18.4 (*9.71)	9.11 (*8.71)			14.9 (*12.8)	8,67 (*7.03)
PsaO	15.1 (*11.1)	9.78 (*8.05)			19.1 (*13,3)	9.06 (*5,13)

For subunits that are imported from the nucleus, the values for the precursor and in addition that of the mature protein (in brackets and with asterisk) are given.

The amino acid sequence of the *Galdieria* PSI subunits were taken and placed into the compute pI/Mw program from ExPASy (Expert Protein Analysis System) and the molecular weight and isoelectric point were determined (Table 2). The *Galdieria* PSI subunits PsaC, D and L had a predicted isoelectric point (pI) that was more similar to cyanobacteria than plants, while the *Galdieria* PSI subunits PsaB, F, J and K showed a pI more similar to angiosperm plants.

3.1. Structural mapping of conserved regions in the individual subunits of PSI from Galdieria

The PDB files resulting from the structural determination of PSI from the pea plant (PDB ENTRY 2O01) [5] and the cyanobacterium *T. elongatus* (PDB ENTRY 1JB0) [4] were used to determine the location of conserved residues in the PSI subunits of *G. sulphuraria*. The homologous amino acids were color coded depending on their sequence homology. The sequences of the individual subunits and their homologies are discussed on the basis of the structural conservation between plants, cyanobacteria and the red alga *Galdieria*.

3.2. The core of Photosystem I: the large subunits PsaA and PsaB

The large subunits PsaA and PsaB form the core of Photosystem I. Each of them consists of 11 transmembrane helices. The PsaA and PsaB sequences from *G. sulphuraria* were aligned against each other and annotated with boxes to indicate the known transmembrane helices from *T. elongatus* (Fig. 2). These helices are labeled a–k for both PsaA and PsaB. PsaA and PsaB coordinate most cofactors of the electron transport chain (P700 to F_X) and 79 of the 90 antenna chlorophylls. All binding sites of the cofactors of the electron transport chain are strictly conserved in PSI from all organisms. In addition, most of the

antenna chlorophyll binding sites are conserved between plants and cyanobacteria [23,5]. Fig. 2A and B show the sequence alignment of PsaA and PsaB. Individual amino acids that show identity between plants, cyanobacteria and *Galdieria* are color coded yellow. Identity between plants and *Galdieria* was color coded green. Identity between cyanobacteria and *Galdieria* was color coded red.

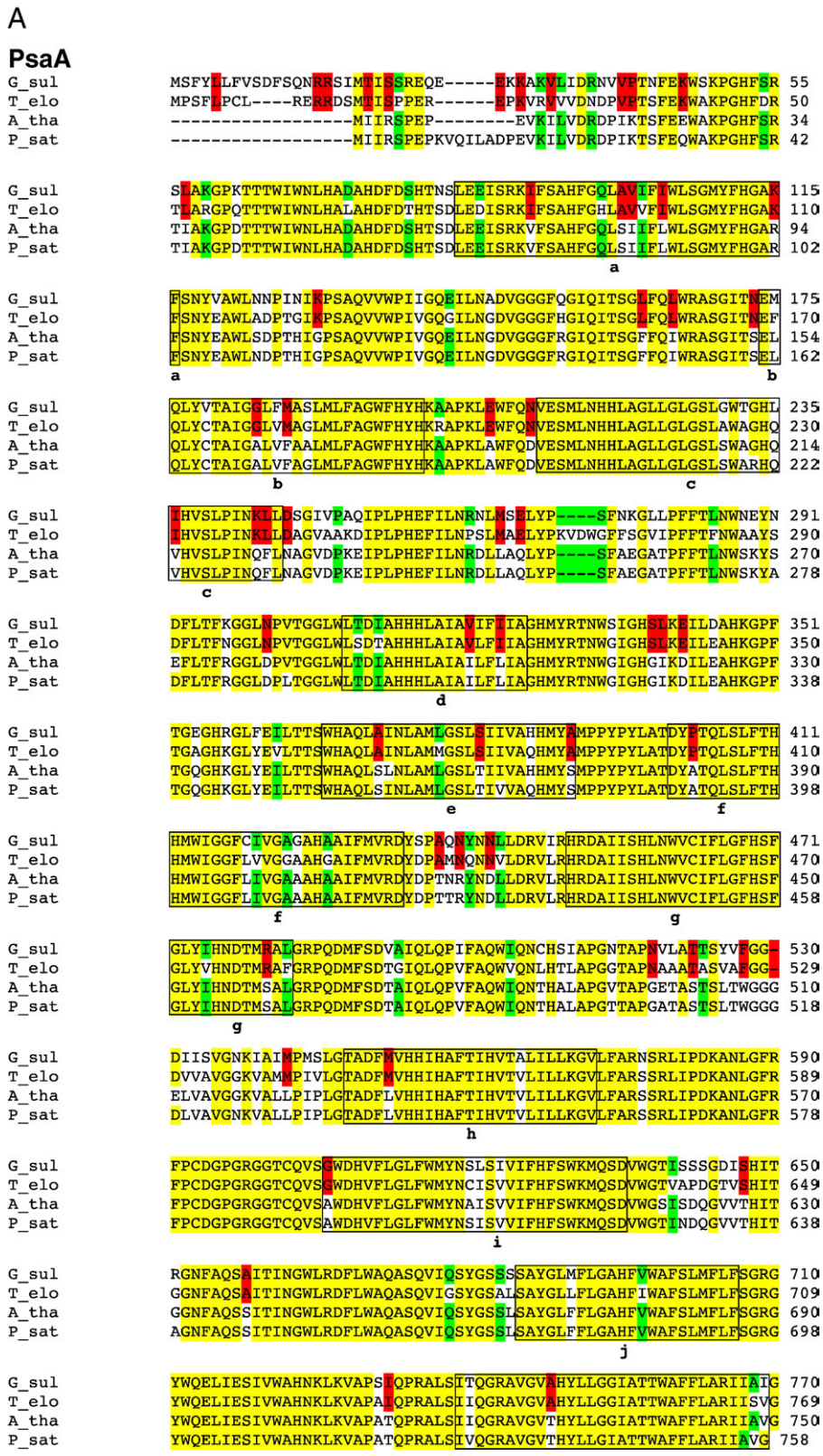

Fig. 2. Sequence alignment of the core subunits of PSI. (A) PsaA, (B) PsaB. The abbreviations stand for: G_sul: *Galdieria sulphuraria* (red alga), T_elo: *Themosynechococcus elongatus* (thermophilic cyanobacterium), A_tha, *arabidopsis* (plant), P_sat, pea (plant).Yellow: Strongly conserved amino acids; these amino acids are conserved between plants, cyanobacteria and *Galdieria*. Red: Cyanobacteria specific amino acids; amino acids that are conserved between cyanobacteria and *Galdieria* but not conserved in plants. Green: Plant specific amino acids; amino acids that are conserved between plants and *Galdieria* and not conserved in cyanobacteria. White: Amino acids with no sequence conservation are shown in white.

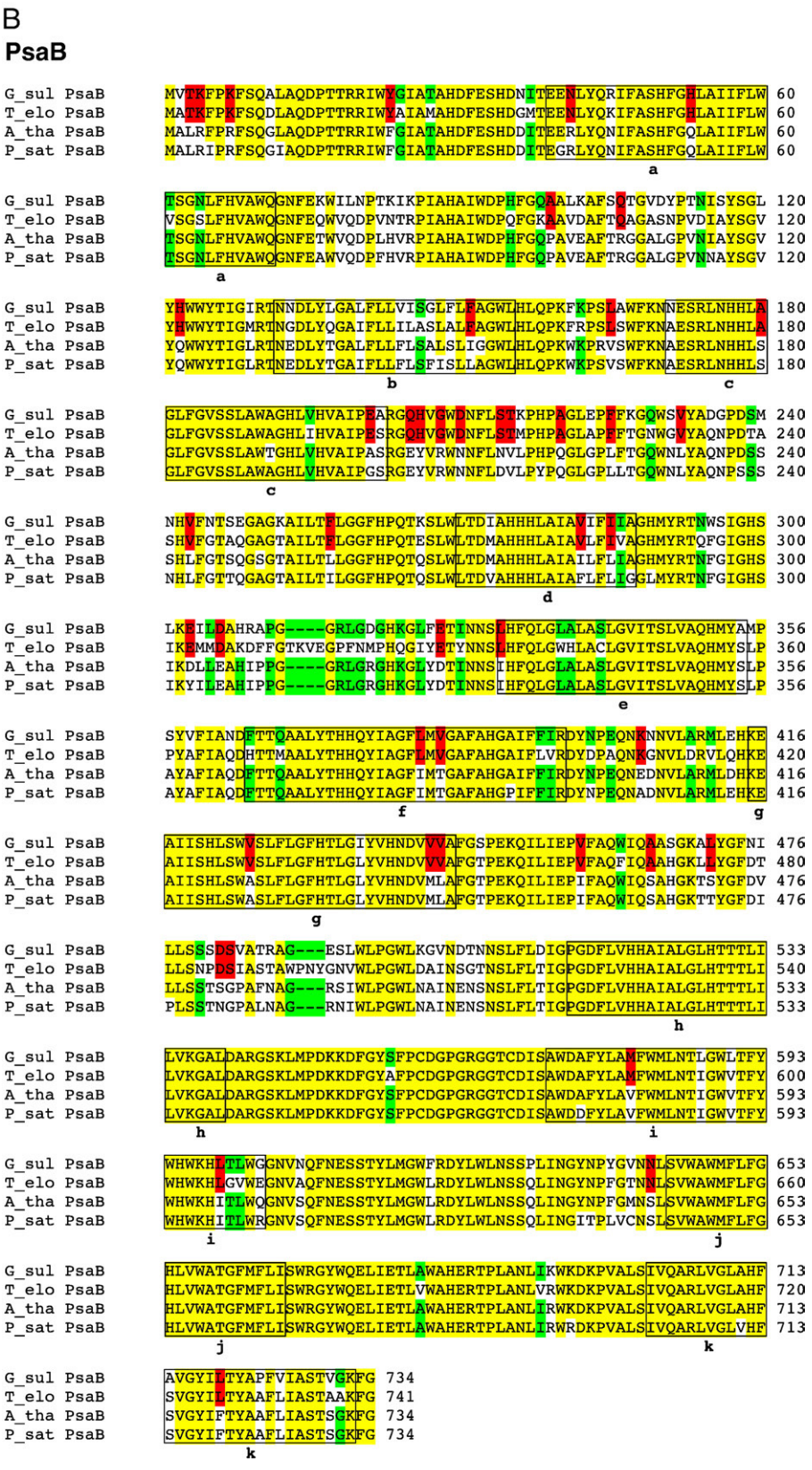

Fig. 2 (*continued*).

Fig. 3A and B show the sequence identity of PsaA/PsaB mapped on the structure of cyanobacterial and plant PSI. The structures of the core subunits are very well conserved in all three species. The only parts where relevant differences in conservation are observed are in the loop regions. The most important differences are found in the loop region between the helices d and e, the de loop. This loop

is shorter in plants than in cyanobacteria with a deletion of the short KVDWG loop in PsaA and the short loop TKVE in PsaB. Both of these deletions are also found in *Galdieria*. In addition to the shortening of the loop, the stromal loop between d and e in *Galdieria* shows strong conservation with plants (11 amino acids are exclusively conserved between plants and *Galdieria*), while only 4 amino acids in this loop share exclusive similarities with the cyanobacterial PSI. This stromal loop is located at periphery of PSI and may be shortened in plants to accommodate the binding of the LHCI complexes to the PSI core. Another deletion that is conserved between plants and *Galdieria* can be found in the luminal gh loop in PsaB. Here, the sequence WPNY is deleted both in plants and *Galdieria*. The mapping of the loop on the structure shows that this loop is also located at the interface between the PSI and the LHCI "belt" in plants. These results show that PSI from *Galdieria* contains two modifications in the loops that may interfere with the peripheral LHCI antenna in plants and leads to the conclusion that it may be able to bind LHC I proteins, similar to the plant Photosystem I.

3.3. PsaL

While PsaL is present in cyanobacteria and plants, it serves different functions in these organisms. In cyanobacteria, PsaL is involved in the formation of PSI timers [24], while it interacts with the plant specific subunit PsaH in higher plants and green algae [5] (see Fig. 1). It was suggested that the interaction with PsaH hinders the trimer formation and allows the interaction of PSI with the light harvesting complex II (LHC II) in the state transition process in plants, which balances the light absorption cross sections of Photosystem II and Photosystem I in green algae and higher plants.

The sequence comparison and the structural mapping of PsaL could therefore provide valuable answers to the question of the potential oligomeric state of PSI from *G. sulphuraria*. Fig. 4 shows the sequence alignment of PsaL. Two striking features draw immediate attention: the *Galdieria* sequence shows a deletion at the C terminus when aligned with the homologous plant and cyanobacterial sequences. *Galdieria* PsaL contains neither the short luminal C-terminal helix of the cyanobacterial protein (VVDGIMTGLFN), which is important for the trimer formation, nor does it contain the plant specific C-terminus (LPYFV(K)K), which is well conserved between *A. thaliana* and spinach.

The trimerization of cyanobacterial PSI is stabilized by Ca^{2+} ions which are coordinated by PsaL and PsaA. From the ligands identified in the structure of *T. elongatus* [4], three are not conserved in *Galdieria*. This finding, together with the lack of the luminal C-terminal helix, may indicate that *Galdieria* PSI cannot form trimers and is very likely monomeric.

Plant PSI contains an extended stromal loop between helices II and III (IAPALTLTGRKKQ) which is lacking both in cyanobacteria and in

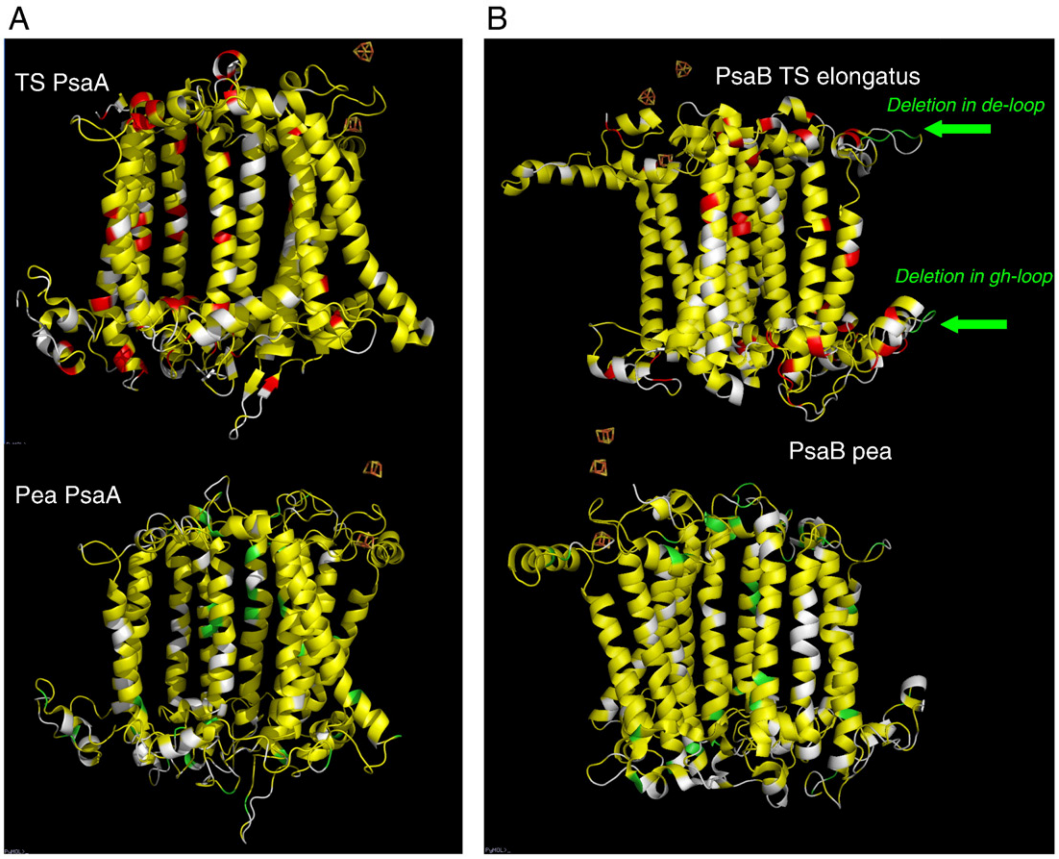

Fig. 3. Structural mapping of conserved amino acids in the cores subunits PsaA and PsaB. (A) Structural conservation of PsaA (top *TS elongatus*, bottom pea). Yellow: Strongly conserved amino acids; these amino acids are conserved between plants, cyanobacteria and *Galdieria*. Red: Cyanobacteria specific amino acids; amino acids that are conserved between cyanobacteria and *Galdieria* but not conserved in plants. Green: Plant specific amino acids; amino acids that are conserved between plants and *Galdieria* and not conserved in cyanobacteria. White: Amino acids with no sequence conservation between *Galdieria* and either plants or cyanobacteria are shown in white. (B) Structural conservation of PsaB (top *TS elongatus*. bottom pea). The color coding is as in Fig. 2A. (C) Mapping of conserved amino acids in the stromal loops of PsaA and PsaB. The view direction is onto the membrane plane from the stromal side, *TS elongatus* (top) pea (bottom). (D) Mapping of conserved amino acids in the luminal loops of PsaA and PsaB. The view direction is onto the membrane plane from the stromal side, *TS elongatus* (top) pea (bottom). (E) Mapping of PsaA and PsaB from the stromal side, *TS elongatus* (top) pea (bottom) foreground elements only. (F) Mapping of PsaA and PsaB from the luminal side, *TS elongatus* (top) pea (bottom) foreground elements only.

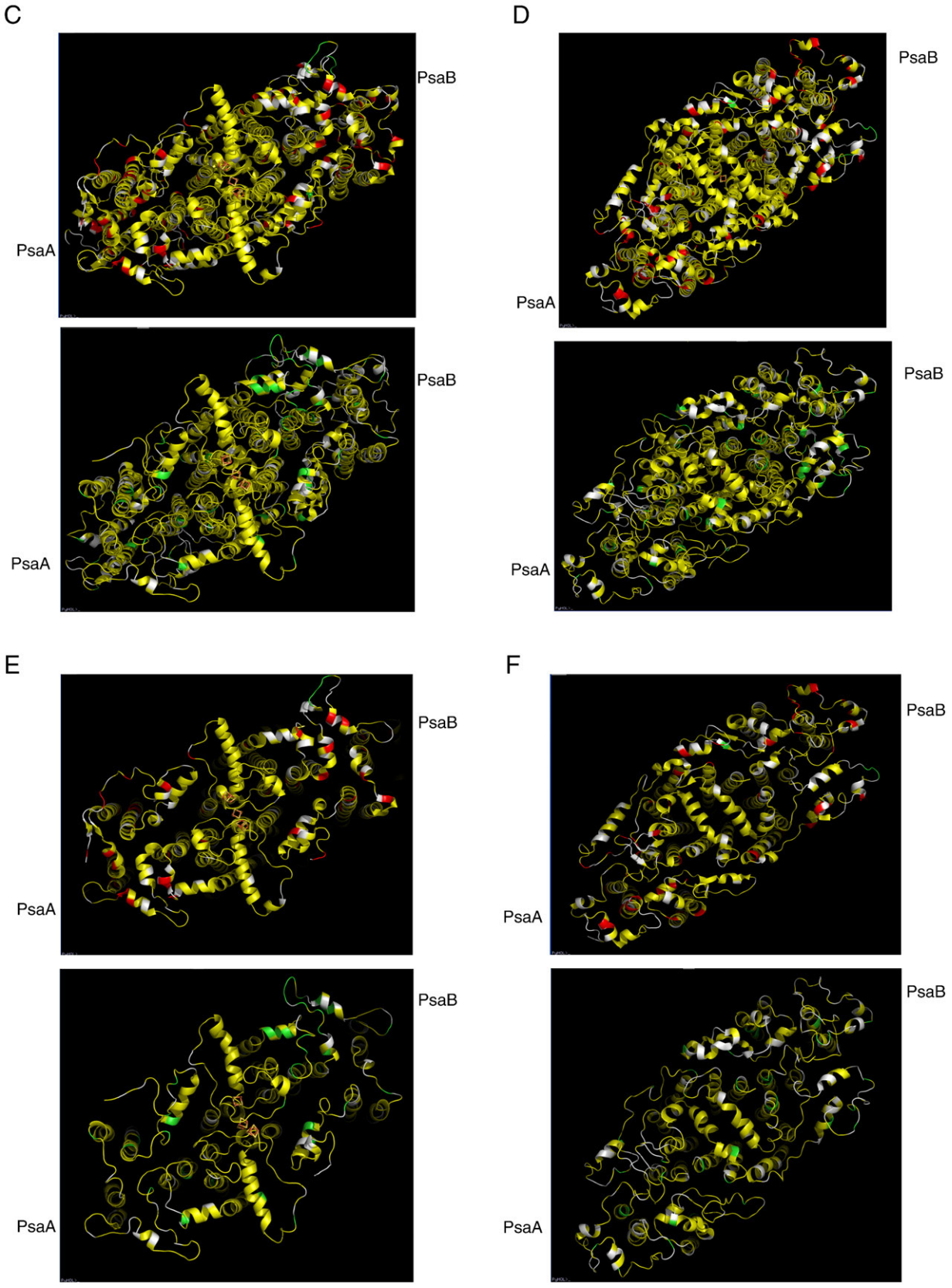

Fig. 3 (*continued*).

Galdieria. The current plant PSI structure does not show any proteins which are in contact to this loop as it sticks out of the structure. This extra-loop is very well conserved in plants, therefore, it may be important for docking of another protein (either one of the "so far not assigned" plant PSI subunits) or the LHC II complex to plant PSI. *Galdieria* has no LHC II complex, so the lack of this loop may be an

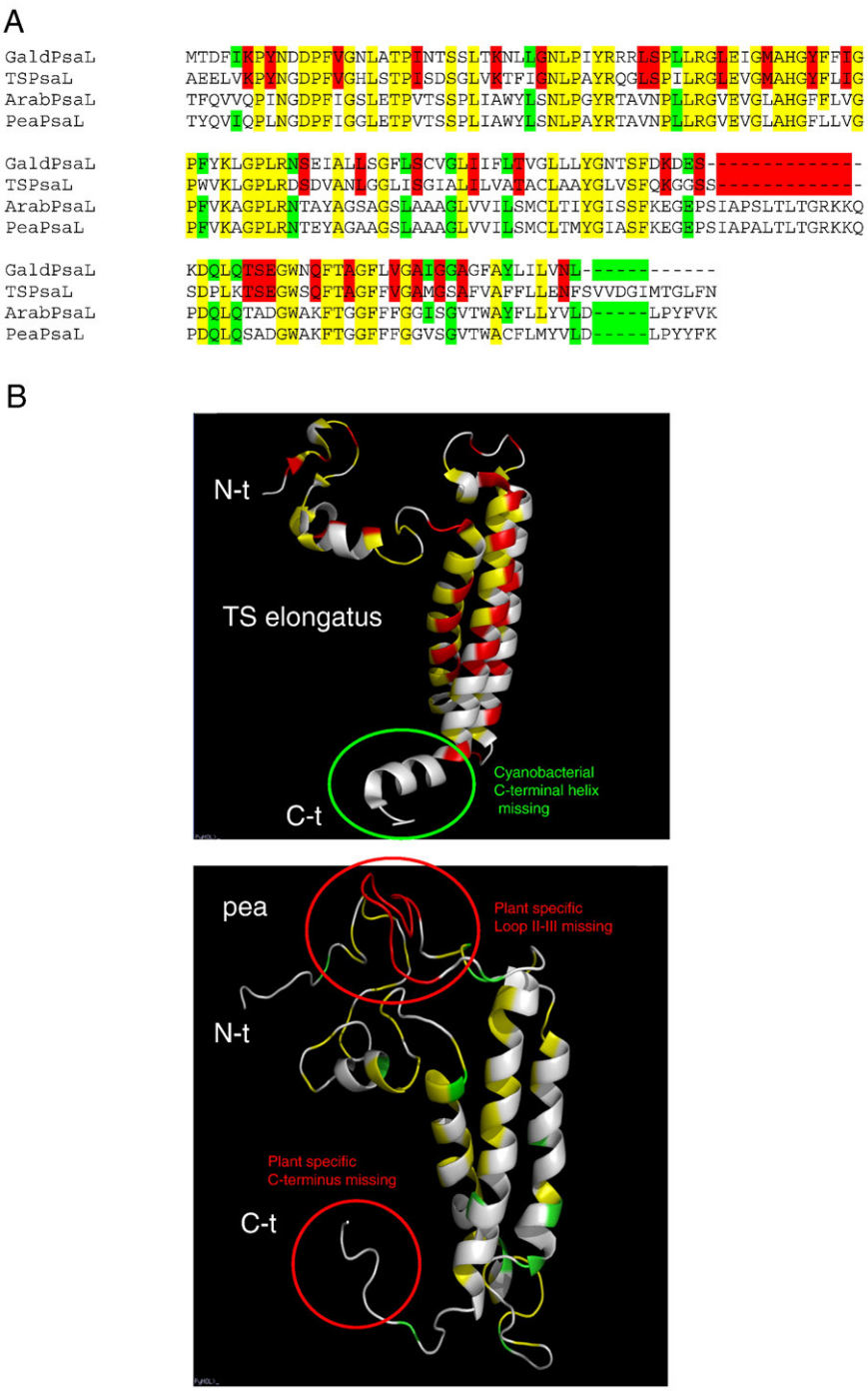

Fig. 4. PsaL. (A) Sequence alignment of PsaL. (B) Mapping of conservation of amino acids in the structure of PsaL.

indication for a potential docking site of plant LHC II to PS I. However, this remains hypothetic for now; it will be of great interest to identify the role of this loop in plants in the future.

3.4. The docking site for plastocyanin: PsaF

One of the major differences between plant and cyanobacterial PSI is the docking site of luminal electron carriers. Plants use plastocyanin as the sole luminal electron carrier, while cyanobacteria can use also cytochrome c_6 as an alternate electron carrier. Cytochrome c_6 represents the more ancient electron carrier and there is evidence that the Cu^{2+} containing plastocyanin has replaced cytochrome c_6 as a response to the iron limitation in the ocean [25,26]. Most cyanobacteria form only a transient complex with cytochrome c_6 or plastocyanin, and the rate constant for docking of PC/cyt c_6 is diffusion limited [27]. The docking site for PC/cyt c_6 in cyanobacteria is exclusively formed by loops of PsaA and PsaB. Over time, plants and green algae have evolved a mechanism to facilitate plastocyanin binding to PSI by increasing the affinity of plastocyanin binding. They form a tighter complex with plastocyanin and PsaF is directly involved in the docking process. Plant PSI contains an N-terminal extension of the luminal domain of PsaF which is rich in lysines. It has been suggested [28,29,30,31] that this positively charged domain in PsaF forms a complex with an acidic domain in plant PC.

Fig. 5 shows the sequence alignment of PsaF from spinach and *T. elongatus* with amino acid homology color coded as described above.

A

```
G_sul PsaF      MMKINRFILYVILFSSILSFSTNNDVQAEFNNLIPCKESKIFNKRLESTIKKLENKLSKY 60
T_elo PsaF      ---MRRFLALLLVLTLWLGFTP--LASADVAGLVPCKDSPAFQKR-------------- 40
A_tha mPsaF     ----------------------------DISGLTPCKDSKQFAKREKQQIKKLESSLKLY 32
S_ole mPsaF     ----------------------------DIAGLTPCKESKQFAKREKQALKKLQASLKLY 32

G_sul PsaF      EVGSSSYLAIKNTINKTNNRFHKYMESGVLCGKDGLPHLIADGFWSHAGFFVIPSLLFIY 120
T_elo PsaF      ---AAAAVNTTADPASGQKRFERYSQA--LCGEDGLPHLVVDGRLSRAGDFLIPSVLFLY 95
A_tha mPsaF     APESAPALALNAQTEKTKRRFDNYGKYGLLCGSDGLPHLIVNGDQRHWGFFITPGILFLY 92
S_ole mPsaF     ADDSAPALAIKATMEKTKKRFDNYGKYGLLCGSDGLPHLIVSGDQRHWGFFITPGILFLY 92

G_sul PsaF      ISGWIGWVGRGYLSAIKYTNKAITNEIIIDVPLALKFMSSGFIWPLSALREYTKGDLLMK 180
T_elo PsaF      IAGWIGWVGRAYLIAVRNSGEANEKEIIIDVPLAIKCMLTGFAWPLAALKELASGELTAK 155
A_tha mPsaF     IAGWIGWVGRSYLIAISGEKKPAMKEIIIDVPLASRIIFRGFIWPVAAYFEFLNGDLIAK 152
S_ole mPsaF     IAGWIGWVGRSYLIAIRDEKKPTQKEIIIDVPLASSLLFRGFSWPVAAYRELLNGELVDN 152

G_sul PsaF      DSLVTISPR 189
T_elo PsaF      DNEIIVSPR 164
A_tha mPsaF     DV------- 154
S_ole mPsaF     NF------- 154
```

B

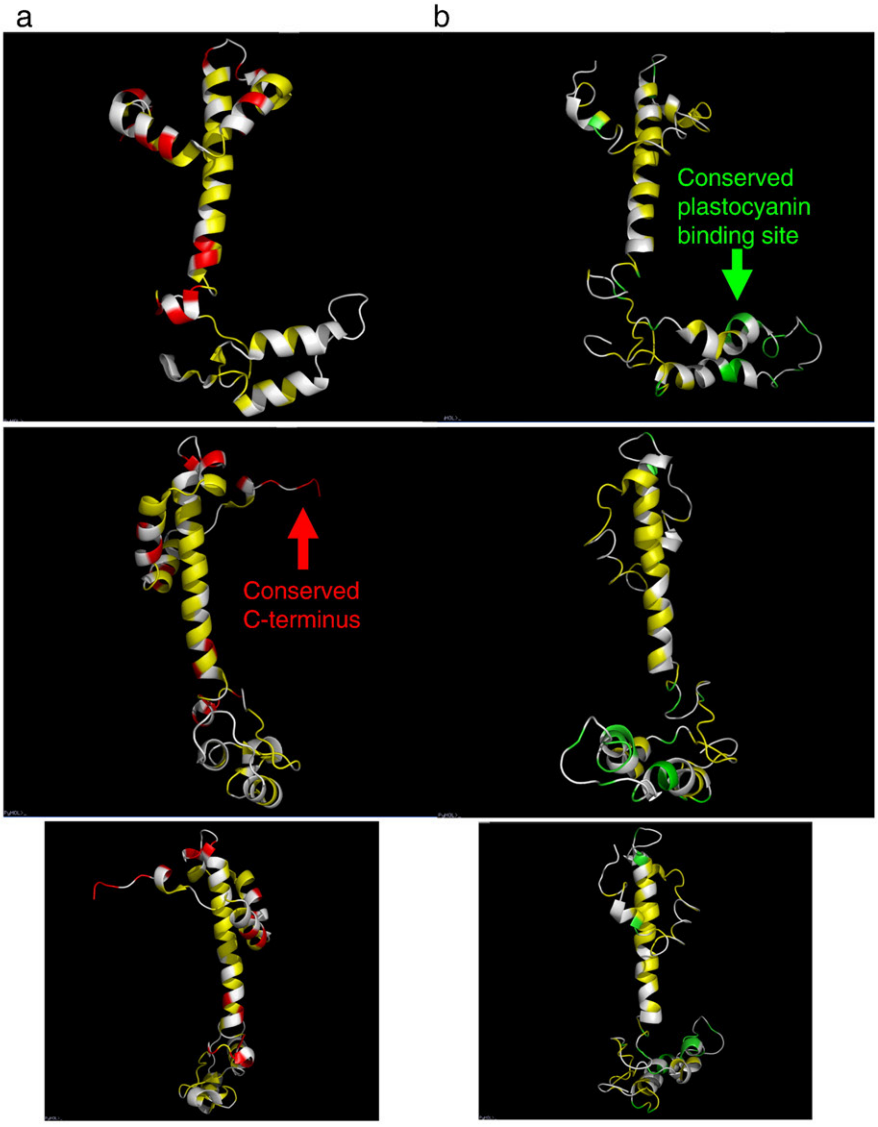

Fig. 5. PsaF. (A) Sequence alignment of PsaF. (B) Mapping of conservation of amino acids in the structure of PsaL.

The N-terminal domain of PsaF from *Galdieria* shows a remarkable sequence homology to plant PsaF. The lysine rich N-terminal domain of the plant system that is responsible for plastocyanin binding is well conserved. *Galdieria* shows a high homology in this region (green residues in plant structure) and had several conserved charged lysine residues which have been suggested to be involved in the docking of plastocyanin in green algae and plants. This result was astonishing as *Galdieria* does not contain a gene for plastocyanin, but uses the more ancient soluble electron carrier, cytochrome c_6 as the sole luminal electron carrier between the cytochrome *b6f* complex and PS I. This

finding has huge implications on the evolution of the luminal electron transfer, which is discussed in more detail in a separate publication (Vanselow et al. in preparation).

The location of the psaF gene differs between plants, green algae and *Galdieria*. In plants and green algae, PsaF is coded by the nuclear genome, while PsaF from *Galdieria* is encoded by the chloroplast. This finding has very interesting implications for the evolution of plant photosynthesis as it shows for the first time that the optimization of luminal electron transfer has taken place in the chloroplast before the PsaF gene was transferred to the nucleus.

PsaF is a chimerical subunit. While the N-terminal domain of PsaF is plant-like, the cyanobacterial C-terminal extension of PsaF consisting of 7 amino acids (EVTISPR) is present and strongly conserved in *Galdieria* with only two conservative amino acid replacements (EITVSPR). This C-terminal domain of PsaF forms strong interactions with both PsaA and PsaE and it has been suggested that it may be important for the interaction of PSI with the phycobillisomes [32]. The finding that this domain is a conserved element of *Galdieria* PSI may lead to the speculation that *Galdieria* PSI may be able to interact with both LHCI and phycobillisomes, and it would be interesting if experimental evidence could support these suggestions in the future. A gene for an LHC protein of 302 amino acids has been described by [33]. It shows sequence homology to the corresponding LHCI polypeptides LhcaR1 and LhcaR2 of *P. cruentum*. The homology is highest in the second helix, which significantly differs between LHCI and LHCII in plants and this finding led to the suggestion that these proteins are LHCI proteins, which may serve as the peripheral antenna to PSI in *Galdieria* [33].

3.5. The stromal subunits PsaC, PsaD and PsaE

Fig. 6 shows the sequence alignment of the stromal subunits that are involved in the docking of ferredoxin (see also Fig 1 for the stromal location of the subunits). PsaC, which carries the two terminal FeS clusters is the best conserved subunit in PSI. Only very few amino acids differ between the organisms. PsaC shows 6 amino acids that are only conserved between *Galdieria* and cyanobacteria and only one amino acid that is more similar to plants, so PsaC is closer related to the cyanobacterial protein; however the function and structure of PsaC is so well conserved that the minor differences may not lead to any major structural or functional differences.

PsaD shows nearly similar sequence variations when compared to cyanobacterial and plant PS I, but does not provide a specific pattern or domain structure of conserved residues.

PsaE has an overall higher sequence similarity to the cyanobacterial protein, which includes the conservation of the C and N-terminus; however, it contains a plant specific deletion of the cyanobacterial sequence (TGYSGSAS). This sequence forms the loop that connects the beta sheets 4 and 5 and interacts with PsaA in docking of cyanobacterial PsaE to PSI. The high conservation of the three extrinsic stromal subunits in all photosynthetic species suggests that the acceptor side of Photosystem I that transfers the electron to ferredoxin has not undergone major changes since the first endosymbiotic event, which is also supported by the finding that ferredoxins from very distant species (plants/cyanobacteria) show very similar electron transfer rates for cyanobacterial and plant PSI [34].

3.6. The small membrane intrinsic subunits

Fig. 7 shows the sequence alignment for the small membrane intrinsic subunits of PSI. The location of the small subunits at the periphery of the PSI complex is depicted in Fig. 1. PsaJ is located in close vicinity to PsaF. The protein forms contacts with the LHCI "belt" in higher plants (see Fig. 1B). In cyanobacteria PsaJ is located at the periphery of the trimer (see Fig. 1A) and coordinates 3 chlorophylls. The result of the sequence alignment is astonishing as it shows that the sequence of PsaJ is closer related to cyanobacteria (11 conserved amino acids that are different in plants) than to plants, where only 3

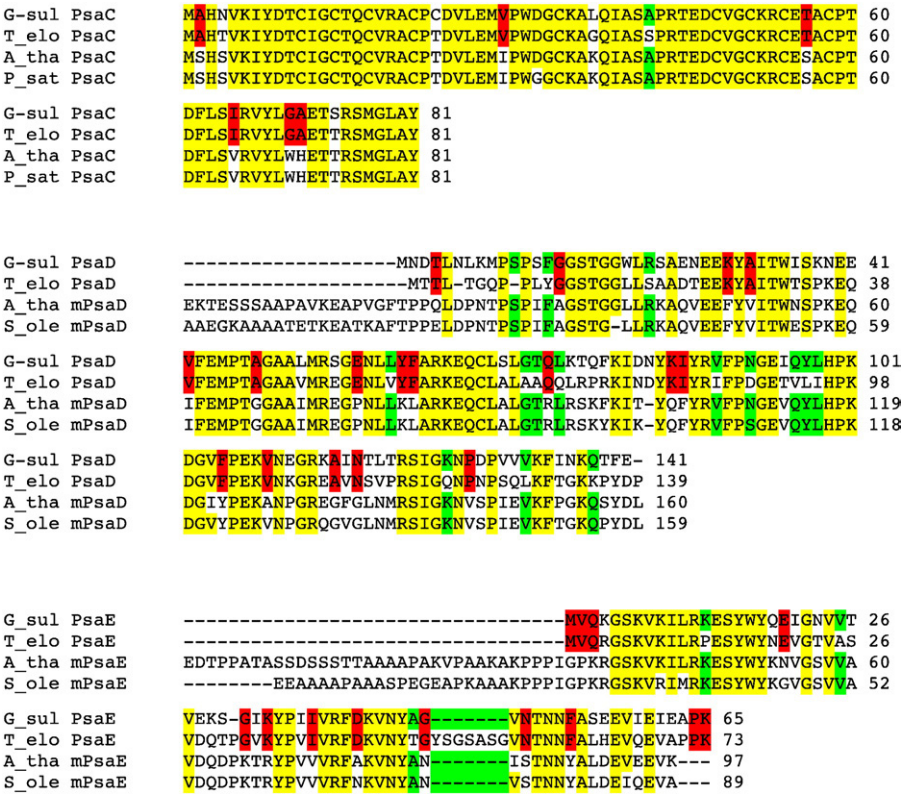

Fig. 6. Sequence alignment of the stromal subunits PsaC, PsaD and PsaE.

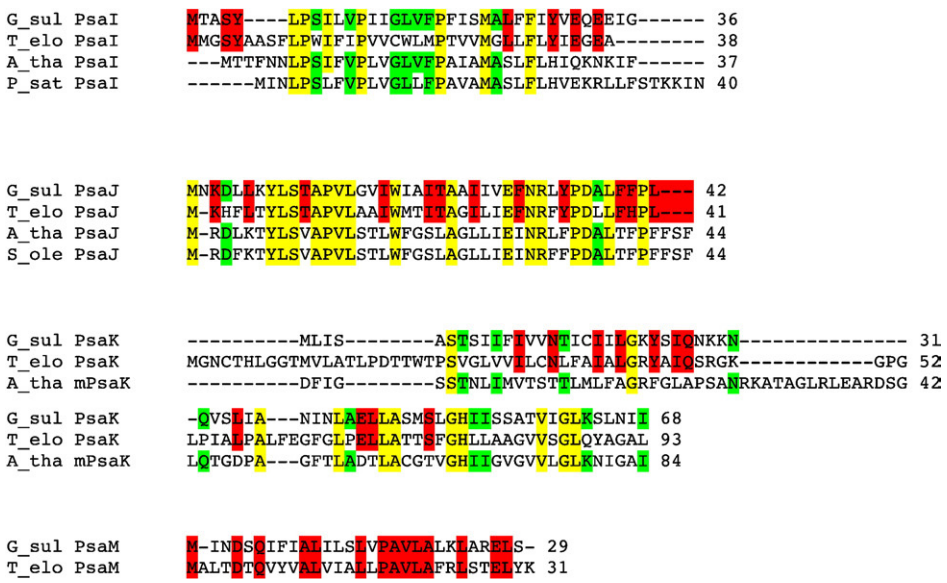

Fig. 7. Sequence alignment of the small subunits PsaI, PsaJ, PsaK and PsaM.

amino acids are plant specific. This is astonishing as it raises the question if the PSI complex in *Galdieria* interacts with the LHC I protein(s) in the same manner as in plants.

PsaK contains 2 TM helices. This protein is only very weekly conserved between plants and cyanobacteria. It is located at the periphery of the PSI monomer and interacts with PsaA (see Fig. 1). In cyanobacteria it might be involved in the interaction with the peripheral IsiA-antenna, which surrounds Photosystem I under iron deficiency [35], while it is involved in binding of the LHCI proteins in higher plants [5]. The protein is encoded by the nucleus in all plants and green algae, however the gene is coded by the chloroplast DNA in *Galdieria* which is consistent with the situation in other red algae [36]. The sequence alignment shows the chimerical nature of *Galdieria* PsaK: the N-terminal region is similar to the plant protein, lacking a large cyanobacterial specific sequence, while the C-terminus is more similar to cyanobacterial PSI with a large deletion shared with the cyanobacterial PSI.

3.7. The plant specific subunits

Plants contain four subunits that are unique to plants and green algae. Of these, two are not found in cyanobacteria: PsaG, PsaH. All plant specific subunits, except PsaO, have been identified in the crystal structure of the plant PSI–LHCI complex. Fig. 1B shows the location of three of the four plant specific subunits. The gene for PsaN is present on stig_41 and moreover contains a predicted N-terminal target peptide that would be needed for its import into the plastids. Similarly, a protein resembling PsaO could be encoded on stig_39 if one assumes that two introns interrupt the coding sequence and lead to a change in the reading frame. Therefore, we can assume that *Galdieria* PSI contains the plant specific subunits PsaN and PsaO. PsaN is located in the lumen and has been structurally identified in the structure of PSI from higher plants [5]. The function of PsaN is still under discussion; it has been discussed that it might be involved in docking of plastocyanin or that it interacts with additional LHCa proteins like Lhca5. [6,5].

Weak sequence similarities could be found for PsaG and PsaH (see Fig. 8). PsaG shows homology to PsaK and it has been suggested that it has evolved via gene duplication from PsaK. It is located at the periphery of the monomer close to PsaB and its location mirrors the binding of PsaK to the core close to PsaA. The assembly of the LHC I belt consisting of Lhca1, Lhca2, Lhca3 and Lhca4 in higher plants may be initiated by the binding of Lhca1 to the plant specific subunit PsaG [37], which shows only very weak sequence homology to PsaG in *Galdieria*. This low homology may be an indication for the evolution of PsaG from a duplicated gene of PsaK. The gene of PsaG in *Galdieria* may therefore represent a more ancient form of PsaG than the one which is now found in green algae and higher plants. In this respect the question arises if the low homology might have implications for the assembly of the LHC I complexes. In plants, PsaG might not be absolutely essential for the binding of the LHC I protein, as mutants of *Arabidopsis* that lack PsaG can still assemble the PSI–LHCI complex [38]. In this regard the differences in the sequence of PsaG in *Galdieria* might not be crucial for the binding of LHC1 protein(s).

4. Discussion

In oxygenic photosynthetic organisms, the core of Photosystem I consists of the polypeptides PsaA and PsaB. These subunits are a result of a duplication event, and both were allowed to diverge based on the selective pressures of the PSI protein. These subunits contain the major components of the electron transport chain. Both of these polypeptides have 11 alpha helices, and they form a heterodimer. Cyanobacteria contain 12 separate polypeptides that make up PSI. These include: PsaA, B, C, D, E, F, I, J, K, L, M and X. Plants contain PsaA, B, C, D, E, F, G, H, I, J, K, L and N (possibly O, but PsaO is not present in the current plant crystal structure). Taking these known subunits from plants and cyanobacteria into account, the genes for the proteins of PSI were investigated using a newly constructed database with the *G. sulphuraria* genome. All subunits that are conserved between plants and cyanobacteria (PsaA, B, C, D, E, F, I, J, K, L) have been identified in *Galdieria*. In contrast to plants and green algae, where several of the small subunits (PsaF, PsaJ, PsaI, PsaK and PsaL) are encoded by the nucleus, all these subunits are encoded by the chloroplasts in *Galdieria*. This leads to the suggestion that *Galdieria* may represent one of the most ancient eukaryotic photosynthetic organisms. The genome also contains the cyanobacterial subunit PsaM, while PsaX, which has so far only been identified in thermophilic cyanobacteria, is not present in the genome of *Galdieria*. This is in accordance with the coding capacity of other red algal genomes. Both *Porphyra* species as well as *Gracilaria tenuistipitata* possess a psaM gene but no open reading frame that shows homology to psaX [39].

One of the major differences between plant and cyanobacterial PSI is the oligomeric state. In cyanobacteria PSI is a trimer, with the

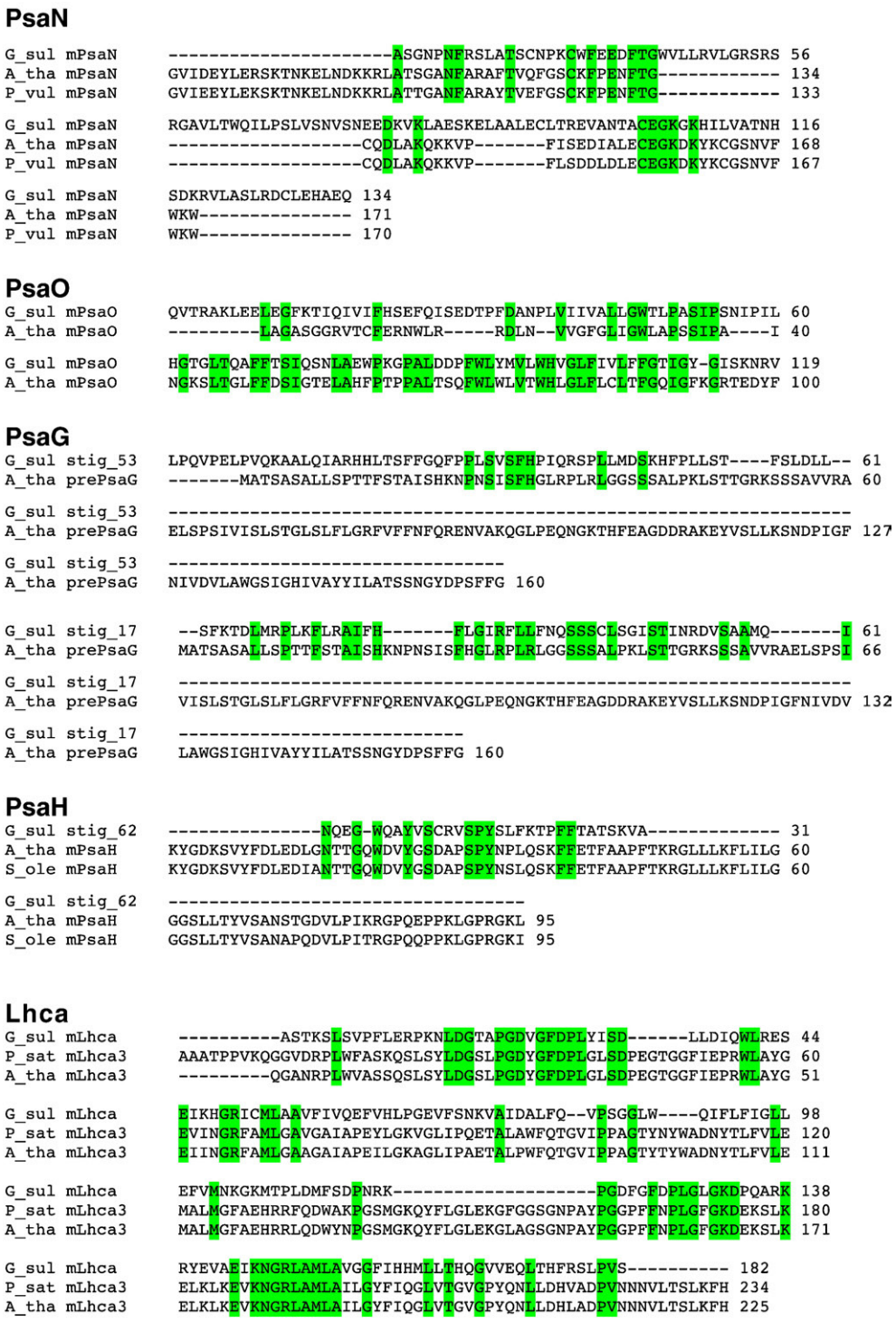

Fig. 8. Tentative sequence alignments of the plant-specific subunits PsaN, PsaO, PsaG and PsaH and the Lhca3 protein.

trimerization domain formed by PsaL [4]. In contrast, plants form PSI–LHC I supercomplexes, where a monomeric PSI contains a belt of 4 LHC I complexes that serve as peripheral antenna.

The mapping of the conserved amino acids in PsaL between *Galdieria* and cyanobacteria and plants unravels a chimerical nature of PsaL, which places *Galdieria* at a transitional state between cyanobacterial and plant PSI. PsaL lacks several elements (the C-terminal helix and the Ca binding site) that are essential for trimerization of PSI in cyanobacteria, while it also lacks similarity in plant specific sequences that intact with the plant specific subunit PsaH, which is absent in *Galdieria*. PsaH is involved in state transitions in plants and might interact with LHC II, which is not present in *Galdieria*, as phycobillisomes serve as peripheral antenna for PSII. The comparison of the PsaL sequence may indicate that PSI from *Galdieria* may represent one of the most ancient PSI, which represents a state that is close to the first endosymbiotic ancestor of all plants and green algae. This view is also supported by the chloroplast genome analysis of the red alga *Cyanidium caldarium*, which has also suggested that this red alga contains elements of cyanobacterial and higher plant concepts [40].

The chimerical nature of PSI from *Galdieria* is also found in PsaF. The lysine rich N-terminal domain of the plant system that is

responsible for plastocyanin binding in plants is well conserved in *Galdieria*, despite the fact that *Galdieria* does not contain plastocyanin but uses cytochrome c6 as the sole luminal electron carrier. This finding suggests that the "plastocyanin binding domain" of plant PSI has not been evolved after transfer of the gene into the nucleus, but is much older than previously anticipated and has been inserted into the PsaF sequence at a time when the gene was still located in the chloroplast genome. The results also indicate that the lysine rich sequence may have been originally developed for the docking of cytochrome c_6, before plants have replaced cytochrome c_6 by plastocyanin.

Four PSI reaction center polypeptides, PsaG, PsaH, PsaN and PsaO are present exclusively in plants and green algae. In the case of PsaN and PsaO, an open reading frame was found on the nuclear stig_41 and stig_39, respectively. The encoded proteins have a lower homology than the other identified subunits but target prediction programs predict localization in the plastids. The presence of a putative target peptide, thus, makes it likely that this protein plays a role in PSI. PsaG and H provided positive hits from the *Galdieria* genome, but upon further investigation, the homology also seemed very low. PsaG forms strong interactions in plants with the Lhca1 protein of the LHCI "belt". Homologous sequences to plant PsaG were identified on stig_17 and stig_62. Interestingly, the signal sequence is even better conserved than the protein itself. Overall, the sequence homology between PsaG from plants and *Galdieria* is weak.

PsaK from stig_35 of the *Galdieria* genome shows deletions homologous with both plants and cyanobacteria, showing a chimerical type structure. The 14 amino acid deletion homologous with plants is present at the N terminus of the polypeptide, while the 16 amino acid deletion homologous with cyanobacteria is present in the middle of the polypeptide. There is also a small deletion homologous with plants near the C-terminus of the polypeptide that contains a deletion of 3 amino acids.

Five LHC proteins are expressed in *G. sulphuraria* as shown by Marquart et al. [41]. N-terminal sequencing of the LHC bands revealed only one sequence XTKSPSVP which corresponds to a gene they named Lhca4. The genes were named in the order they have been isolated and did not reflect a high degree of sequence homology with plant LHC protein. The protein identified in the PSI preparation is the only one which shows significant homology to plant LHCI proteins. Fig. 8 shows a sequence alignment of this protein with the Lhc3 protein from plants. The sequences are highly conserved thereby strongly indicating that *Galdieria* contains at least one LHCI protein. Interestingly, Lhca3 is located in the plant PSI structure at the "north" end of the LHCI belt in close interaction with PsaK, a subunit that is present in both plants and cyanobacteria and has been discussed above. Interestingly, the plant specific subunit PsaG shows sequence homologies to PsaK and it has been suggested that it had been evolved by a gene duplication of PsaK [42]. Mutants that lack PsaG and PsaK in *Arabidopsis* have multiple effects ranging from decrease of Lhc2 and Lhc3 proteins to destabilization of the core to a potential role of PsaK in state transitions. However these results also suggest that PsaG and PsaK are not absolutely essential for the assembly of the LHCI belt as the complexes isolated without PsaG and PsaK still contain LHCI proteins[43]. This finding supports the idea that the PSI from *Galdieria* interacts with LHCI protein(s) in a similar fashion as plant PSI despite the absence of PsaG. It might also be possible that this protein may contain two copies of PsaK flanking PSI at both the "north" site close to PsaA and the "south" site in close vicinity to PsaB. It will be very interesting to prove this hypothesis and investigate the structure of the PSI complex in *Galdieria* in the future. The most popular model for the evolution of PSI has developed a viewpoint that puts the trimeric cyanobacterial PSI at the base of the evolutionary tree and discusses all changes in the plant system as additions and modifications of the cyanobacterial model. However, we may also have to consider the possibility that the "ur"-PSI at the time of the primary Endosymbiotic event that led to the development of green algae and plants was monomeric, with a belt of LHCI proteins loosely attached as a peripheral antenna system. In this scenario, the cyanobacteria may have then later developed the trimeric PSI and the invention of the phycobillisomes, which are much more efficient antenna as they can absorb green light, thereby dramatically increasing the use of the solar spectrum for photosynthesis. This hypothesis is supported by the fact that red algae (derived from a secondary endosymbiotic event) and cyanobacteria, which both contain phycobillisomes, are the most successful and abundant species in the ocean.

In summary, the results of this study indicate that *Galdieria* may contain a PSI that is evolutionarily much more ancient than PSI from green algae, plants and the current cyanobacteria. This places *G. sulphuraria* at the roots of the evolutionary trees of all Photosystem I proteins. These findings make this PSI a very interesting target for future structural and functional studies.

Acknowledgements

This work is supported by the National Science Foundation grant (award MCB-0417142) to PF and the NSF funding of the *Galdieria* genome project to APMW (award EF-0332882) and the NSF IGERT fellowship (interdisciplinary nanotechnology) to Christopher Vanselow.

References

[1] A. Reyes-Prieto, A.P. Weber, D. Bhattacharya, The origin and establishment of the plastid in algae and plants, Annu. Rev. Genet. 41 (2007) 147–168.
[2] S.M. Adl, A.G. Simpson, M.A. Farmer, R.A. Andersen, O.R. Anderson, J.R. Barta, S.S. Bowser, G. Brugerolle, R.A. Fensome, S. Fredericq, T.Y. James, S. Karpov, P. Kugrens, J. Krug, C.E. Lane, L.A. Lewis, J. Lodge, D.H. Lynn, D.G. Mann, R.M. McCourt, L. Mendoza, O. Moestrup, S.E. Mozley-Standridge, T.A. Nerad, C.A. Shearer, A.V. Smirnov, F.W. Spiegel, M.F. Taylor, The new higher level classification of eukaryotes with emphasis on the taxonomy of protists, J. Eukaryot. Microbiol. 52 (2005) 399–451.
[3] S.B. Gould, R.F. Waller, G.I. McFadden, Plastid evolution, Annu. Rev. Plant Biol. 59 (2008) 491–517.
[4] P. Jordan, P. Fromme, H.T. Witt, O. Klukas, W. Saenger, N. Krauss, Three-dimensional structure of cyanobacterial Photosystem I at 2.5 Å resolution, Nature 411 (2001) 909–917.
[5] A. Amunts, O. Drory, N. Nelson, The structure of a plant Photosystem I supercomplex at 3.4 Å resolution, Nature 447 (2007) 58–63.
[6] P.E. Jensen, R. Bassi, E.J. Boekema, J.P. Dekker, S. Jansson, D. Leister, C. Robinson, H.V. Scheller, Structure, function and regulation of plant Photosystem I, Biochim. Biophys. Acta 1767 (2007) 335–352.
[7] W. Gross, C. Schnarrenberger, Purification and characterization of a galactose-1-phosphate: UDP-glucose uridyltransferase from the red alga *Galdieria sulphuraria*, Eur. J. Biochem. 234 (1995) 258–263.
[8] J.A. Toplin, T.B. Norris, C.R. Lehr, T.R. McDermott, R.W. Castenholz, Biogeographic and phylogenetic diversity of thermoacidophilic cyanidiales in Yellowstone National Park, Japan, and New Zealand, Appl. Environ. Microbiol. 74 (2008) 2822–2833.
[9] J.J. Walker, J.R. Spear, N.R. Pace, Geobiology of a microbial endolithic community in the Yellowstone geothermal environment, Nature 434 (2005) 1011–1014.
[10] A.P. Weber, C. Oesterhelt, W. Gross, A. Brautigam, L.A. Imboden, I. Krassovskaya, N. Linka, J. Truchina, J. Schneidereit, H. Voll, L.M. Voll, M. Zimmermann, A. Jamai, W.R. Riekhof, B. Yu, R.M. Garavito, C. Benning, EST-analysis of the thermo-acidophilic red microalga *Galdieria sulphuraria* reveals potential for lipid A biosynthesis and unveils the pathway of carbon export from rhodoplasts, Plant Mol. Biol. 55 (2004) 17–32.
[11] C. Oesterhelt, S. Klocke, S. Holtgrefe, V. Linke, A.P. Weber, R. Scheibe, Redox regulation of chloroplast enzymes in *Galdieria sulphuraria* in view of eukaryotic evolution, Plant Cell Physiol. 48 (2007) 1359–1373.
[12] A.P. Weber, R.J. Horst, G.G. Barbier, C. Oesterhelt, Metabolism and metabolomics of eukaryotes living under extreme conditions, Int. Rev. Cytol. 256 (2007) 1–34.
[13] A. Weber, G. Barbier, R. Shrestha, R. Horst, A. Minoda, C. Oesterhelt, A genomics approach to understanding the biology of thermo-acidophilic red algae, in: J. Seckbach (Ed.), Algae and Cyanobacteria in Extreme Environments, Springer, Berlin, 2007, pp. 503–518.
[14] S.F. Altschul, T.L. Madden, A.A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[15] S.W. Schaeffer, C.S. Walthour, D.M. Toleno, A.T. Olek, E.L. Miller, Protein variation in Adh and Adh-related in *Drosophila pseu*doobscura. Linkage disequilibrium

between single nucleotide polymorphisms and protein alleles, Genetics 159 (2001) 673–687.
[16] A. Ben-Shem, N. Nelson, F. Frolow, Crystallization and initial X-ray diffraction studies of higher plant Photosystem I, Acta Crystallogr., D Biol. Crystallogr. 59 (2003) 1824–1827.
[17] G. Barbier, C. Oesterhelt, M.D. Larson, R.G. Halgren, C. Wilkerson, R.M. Garavito, C. Benning, A.P. Weber, Comparative genomics of two closely related unicellular thermo-acidophilic red algae, *Galdieria sulphuraria* and *Cyanidioschyzon merolae*, reveals the molecular basis of the metabolic flexibility of *Galdieria sulphuraria* and significant differences in carbohydrate metabolism of both algae, Plant Physiol. 137 (2005) 460–474.
[18] Y. Nakamura, T. Kaneko, S. Sato, M. Ikeuchi, H. Katoh, S. Sasamoto, A. Watanabe, M. Iriguchi, K. Kawashima, T. Kimura, Y. Kishida, C. Kiyokawa, M. Kohara, M. Matsumoto, A. Matsuno, N. Nakazaki, S. Shimpo, M. Sugimoto, C. Takeuchi, M. Yamada, S. Tabata, Complete genome structure of the thermophilic cyanobacterium *Thermosynechococcus elongatus BP-1*, DNA Res. 9 (2002) 123–130.
[19] Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*, Nature 408 (2000) 796–815.
[20] S. Parinov, V. Sundaresan, Functional genomics in *Arabidopsis*: large-scale insertional mutagenesis complements the genome sequencing project, Curr. Opin. Biotechnol. 11 (2000) 157–161.
[21] B. Bjellqvist, B. Basse, E. Olsen, J.E. Celis, Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions, Electrophoresis 15 (1994) 529–539.
[22] B. Bjellqvist, G.J. Hughes, C. Pasquali, N. Paquet, F. Ravier, J.C. Sanchez, S. Frutiger, D. Hochstrasser, The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences, Electrophoresis 14 (1993) 1023–1031.
[23] A. Ben-Shem, F. Frolow, N. Nelson, Crystal structure of plant Photosystem I, Nature 426 (2003) 630–635.
[24] V.P. Chitnis, P.R. Chitnis, PsaL subunit is required for the formation of Photosystem I trimers in the cyanobacterium *Synechocystis sp. PCC 6803*, FEBS Lett. 336 (1993) 330–334.
[25] M. Hervas, J.A. Navarro, M.A. De La Rosa, Electron transfer between membrane complexes and soluble proteins in photosynthesis, Acc. Chem. Res. 36 (2003) 798–805.
[26] A. Diaz-Quintana, J.A. Navarro, M. Hervas, F.P. Molina-Heredia, B. De la Cerda, M.A. De la Rosa, A comparative structural and functional analysis of cyanobacterial plastocyanin and cytochrome c_6 as alternative electron donors to Photosystem I, Photosynth. Res. 75 (2003) 97–110.
[27] M. Hervas, J.A. Navarro, A. Diaz, H. Bottin, M.A. De la Rosa, Laser-flash kinetic analysis of the fast electron transfer from plastocyanin and cytochrome c_6 to Photosystem I. Experimental evidence on the evolution of the reaction mechanism, Biochemistry 34 (1995) 11321–11326.
[28] M. Hippler, K. Redding, J.D. Rochaix, Chlamydomonas genetics, a tool for the study of bioenergetic pathways, Biochim. Biophys. Acta 1367 (1998) 1–62.
[29] M. Hippler, J. Reichert, M. Sutter, E. Zak, L. Altschmied, U. Schroer, R.G. Herrmann, W. Haehnel, The plastocyanin binding domain of Photosystem I, EMBO J. 15 (1996) 6374–6384.
[30] N. Nelson, A. Ben-Shem, Photosystem I reaction center: past and future, Photosynth. Res. 73 (2002) 193–206.
[31] N. Nelson, A. Ben-Shem, The structure of Photosystem I and evolution of photosynthesis, BioEssays 27 (2005) 914–922.
[32] P. Fromme, A. Melkozernov, P. Jordan, N. Krauss, Structure and function of Photosystem I: interaction with its soluble electron carriers and external antenna systems, FEBS Lett. 555 (2003) 40–44.
[33] J. Marquardt, S. Wans, E. Rhiel, A. Randolf, W.E. Krumbein, Intron–exon structure and gene copy number of a gene encoding for a membrane-intrinsic light-harvesting polypeptide of the red alga *Galdieria sulphuraria*, Gene 255 (2000) 257–265.
[34] P. Setif, Ferredoxin and flavodoxin reduction by Photosystem I, Biochim. Biophys. Acta 1507 (2001) 161–179.
[35] J. Nield, E.P. Morris, T.S. Bibby, J. Barber, Structural analysis of the Photosystem I supercomplex of cyanobacteria induced by iron deficiency, Biochemistry 42 (2003) 3180–3188.
[36] N. Ohta, M. Matsuzaki, O. Misumi, S.Y. Miyagishima, H. Nozaki, K. Tanaka, I.T. Shin, Y. Kohara, T. Kuroiwa, Complete sequence and analysis of the plastid genome of the unicellular red alga *Cyanidioschyzon merolae*, DNA Res. 10 (2003) 67–77.
[37] A. Amunts, N. Nelson, Functional organization of a plant Photosystem I: evolution of a highly efficient photochemical machine, Plant Physiol. Biochem. 46 (2008) 228–237.
[38] P.E. Jensen, L. Rosgaard, J. Knoetzel, H.V. Scheller, Photosystem I activity is increased in the absence of the PSI-G subunit, J. Biol. Chem. 277 (2002) 2798–2803.
[39] J.C. Hagopian, M. Reis, J.P. Kitajima, D. Bhattacharya, M.C. de Oliveira, Comparative analysis of the complete plastid genome sequence of the red alga *Gracilaria tenuistipitata var. liui* provides insights into the evolution of rhodoplasts and their relationship to other plastids, J. Mol. Evol. 59 (2004) 464–477.
[40] Z. Gardian, L. Bumba, A. Schrofel, M. Herbstova, J. Nebesarova, F. Vacha, Organisation of Photosystem I and Photosystem II in red alga *Cyanidium caldarium*: encounter of cyanobacterial and higher plant concepts, Biochim. Biophys. Acta 1767 (2007) 725–731.
[41] J. Marquardt, B. Lutz, S. Wans, E. Rhiel, W.E. Krumbein, The gene family coding for the light-harvesting polypeptides of Photosystem I of the red alga *Galdieria sulphuraria*, Photosynth. Res. 68 (2001) 121–130.
[42] S. Jansson, B. Andersen, H.V. Scheller, Nearest-neighbor analysis of higher-plant Photosystem I holocomplex, Plant Physiol. 112 (1996) 409–420.
[43] C. Varotto, P. Pesaresi, P. Jahns, A. Lessnick, M. Tizzano, F. Schiavon, F. Salamini, D. Leister, Single and double knockouts of the genes for Photosystem I subunits G, K, and H of *Arabidopsis*. Effects on Photosystem I composition, photosynthetic electron flow, and state transitions, Plant Physiol. 129 (2002) 616–624.