New Phytologist

Conservation of core complex subunits shaped the structure and function of photosystem I in the secondary endosymbiont alga *Nannochloropsis gaditana*

Alessandro Alboresi[1]*, Clotilde Le Quiniou[2]*, Sathish K. N. Yadav[3]*, Martin Scholz[4], Andrea Meneghesso[1], Caterina Gerotto[1], Diana Simionato[1], Michael Hippler[4], Egbert J. Boekema[3], Roberta Croce[2] and Tomas Morosinotto[1]

[1]Dipartimento di Biologia, Università di Padova, Via U. Bassi 58/B, 35121 Padova, Italy; [2]Department of Physics and Astronomy and Institute for Lasers, Life and Biophotonics, Faculty of Sciences, VU University Amsterdam, De Boelelaan 1081, 1081 HV, Amsterdam, the Netherlands; [3]Electron Microscopy Group, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Nijenborgh 7, 9747 AG Groningen, the Netherlands; [4]Institute of Plant Biology and Biotechnology, University of Münster, Münster 48143, Germany

Author for correspondence:
Tomas Morosinotto
Tel: +39 0498277484
Email: tomas.morosinotto@unipd.it

Received: *30 May 2016*
Accepted: *13 July 2016*

New Phytologist (2017) **213:** 714–726
doi: 10.1111/nph.14156

Key words: electron microscopy, Heterokonta, Light-harvesting complex, *Nannochloropsis*, photosynthesis, photosynthetic apparatus, photosystem.

Summary

- Photosystem I (PSI) is a pigment protein complex catalyzing the light-driven electron transport from plastocyanin to ferredoxin in oxygenic photosynthetic organisms. Several PSI subunits are highly conserved in cyanobacteria, algae and plants, whereas others are distributed differentially in the various organisms.
- Here we characterized the structural and functional properties of PSI purified from the heterokont alga *Nannochloropsis gaditana*, showing that it is organized as a supercomplex including a core complex and an outer antenna, as in plants and other eukaryotic algae.
- Differently from all known organisms, the *N. gaditana* PSI supercomplex contains five peripheral antenna proteins, identified by proteome analysis as type-R light-harvesting complexes (LHCr4-8). Two antenna subunits are bound in a conserved position, as in PSI in plants, whereas three additional antennae are associated with the core on the other side. This peculiar antenna association correlates with the presence of PsaF/J and the absence of PsaH, G and K in the *N. gaditana* genome and proteome.
- Excitation energy transfer in the supercomplex is highly efficient, leading to a very high trapping efficiency as observed in all other PSI eukaryotes, showing that although the supramolecular organization of PSI changed during evolution, fundamental functional properties such as trapping efficiency were maintained.

Introduction

In organisms performing oxygenic photosynthesis, Photosystem I (PSI) is responsible for the light-driven electron transport from plastocyanin/cytochrome 6 to ferredoxin (Raven *et al.*, 1999; Nelson & Yocum, 2006; Croce & van Amerongen, 2013; Bernal-Bayard *et al.*, 2015). In eukaryotes, PSI is organized as a supercomplex composed of two moieties, the core and the peripheral antenna system. The PSI core is composed of 12–14 subunits (Busch & Hippler, 2011). PsaA and PsaB bind the reaction center P700 as well as most of the cofactors involved in electron transport and *c.* 100 chlorophylls (Chls) active in light harvesting (Jordan *et al.*, 2001; Qin *et al.*, 2015). PsaA and PsaB are highly conserved in cyanobacteria, algae and plants (Allen *et al.*, 2011). Other core subunits are instead differently distributed in various phylogenetic groups. For instance PsaG and PsaH are only found in plants and green algae (Vanselow *et al.*, 2009; Busch & Hippler, 2011) where they serve as docking site for the association of the peripheral antenna (Lunde *et al.*, 2000; Ben-Shem *et al.*, 2003).

The peripheral antenna is composed of light-harvesting complexes (LHC) whose sequences vary in different organisms: LHCa1–6 are present in plants (Jansson, 1999), LHCa1–9 are present in the green alga *Chlamydomonas reinhardtii* (Mozzo *et al.*, 2010), whereas LHCR proteins are believed to comprise the peripheral antenna in red algae and diatoms (Busch *et al.*, 2010; Busch & Hippler, 2011; Thangaraj *et al.*, 2011). The number of LHC associated with PSI is also variable: four subunits are associated with the PSI core in higher plants, in the moss *Physcomitrella patens*, and in some red algae and diatoms (Ben-Shem *et al.*, 2003; Veith & Büchel, 2007; Busch *et al.*, 2010, 2013). Differently, nine antenna subunits are found associated with PSI in other red algae (Gardian *et al.*, 2007; Thangaraj *et al.*, 2011) as in the green alga *C. reinhardtii* (Drop *et al.*, 2011).

Finally, the oligomeric state of PSI also varies in different organisms. Cyanobacterial PSI often has been reported to be a

*These authors contributed equally to this work.

trimer (Boekema *et al.*, 1987; Jordan *et al.*, 2001), although PSI tetramers have been identified in a growing number of species (Watanabe *et al.*, 2011, 2014; Li *et al.*, 2014). However, PSI was found to be monomeric in all eukaryotes analyzed so far, including plants (Ben-Shem *et al.*, 2003; Kouril *et al.*, 2005a), diatoms (Veith & Büchel, 2007; Ikeda *et al.*, 2013), green and red algae (Gardian *et al.*, 2007; Drop *et al.*, 2011).

A peculiar feature of PSI is the presence of Chls absorbing at energy lower than the primary electron donor P700, called red forms (Croce & van Amerongen, 2013). Although these far red-absorbing Chls account only for a small fraction of the total absorption, they have a strong influence in excitation energy transfer and trapping, slowing down the trapping time as they introduce uphill steps in the energy transfer process (Gobets *et al.*, 2001; Jennings *et al.*, 2003; Engelmann *et al.*, 2006; Wientjes *et al.*, 2011b). The presence of low energy absorbing Chls is ubiquitous in PSI, but their energy appears to be highly species-dependent (Gobets & van Grondelle, 2001; Croce & van Amerongen, 2013). In plants, most red forms are associated with the outer antenna complexes and in particular with LHCA3 and LHCA4 (Schmid *et al.*, 1997; Castelletti *et al.*, 2003), although the core also contains low energy forms (Croce *et al.*, 1998; Gobets & van Grondelle, 2001).

In the present work, we investigated the structural and functional properties of the Photosystem I supercomplex (PSI-LHC) of the eustigmatophycea *Nannochloropsis gaditana*. This microalga belongs to the phylum Heterokonta, which also includes diatoms and brown algae (Cavalier-Smith, 2004; Riisberg *et al.*, 2009), that originated from a secondary endosymbiotic event where an eukaryotic host cell engulfed a red alga (Archibald & Keeling, 2002). In the last few years, species belonging to the *Nannochloropsis* genus have gained increased attention not only for their evolutionary position, but also for their ability to accumulate a large amount of lipids (Rodolfi *et al.*, 2009; Bondioli *et al.*, 2012; Simionato *et al.*, 2013). The *N. gaditana* photosynthetic apparatus presents distinct features with respect to other algae such as the presence of only Chl*a* and an atypical carotenoid composition with violaxanthin and vaucheriaxanthin esters as the most abundant xanthophylls (Sukenik *et al.*, 1992, 2000; Basso *et al.*, 2014). A deeper characterization of this organism thus contributes to a better understanding of the variability of photosynthetic organisms in an evolutionary context.

Materials and Methods

Cell growth and thylakoid isolation

Nannochloropsis gaditana from the Culture Collection of Algae and Protozoa (CCAP), strain 849/5, was grown in sterile F/2 medium using 32 g l^{-1} sea salts (Sigma-Aldrich), 40 mM TRIS-HCl (pH 8) and Guillard's (F/2) marine water enrichment solution (Sigma-Aldrich). Cells were grown with 100 µmol of photons m^{-2} s^{-1} of illumination using fluorescent light tubes and air enriched with 5% CO_2. The temperature was set at 22 ± 1°C. Thylakoid membranes were isolated as described previously (Basso *et al.*, 2014).

Thylakoid solubilization and PSI isolation

Thylakoid membranes (corresponding to 500 µg of Chl) were solubilized with 0.6% α-DM or 1% β-DM as described in (Basso *et al.*, 2014) and then fractionated by ultracentrifugation in a 0.1–1 M sucrose gradient containing 0.06% α-DM and 10 mM HEPES (pH 7.5) (280 000 ***g***, 18 h, 4°C). Fractions of the sucrose gradient were then harvested with a syringe. PSI samples isolated after α-DM or β-DM solubilization were named PSI-LHC$_\alpha$ and PSI-LHC$_\beta$, respectively.

Electron microscopy and image analysis

Four microliters of purified sample was absorbed onto glow discharged carbon-coated grids and subsequently stained with 2% uranyl acetate for contrast. Imaging was performed on a Tecnai T20 equipped with a LaB6 tip operating at 200 kV. The 'GRACE' system for semi-automated specimen selection and data acquisition (Oostergetel *et al.*, 1998) was used to record 2048 × 2048 pixel images at ×133 000 magnifications using a Gatan 4000 SP 4K slow-scan CCD camera with a pixel size of 0.224 nm. Three hundred thousand particles were picked from 16 000 raw images and single particles were analyzed with XMIPP software (including alignments, statistical analysis and classification, as in Scheres *et al.*, 2008) and RELION software (Scheres & Chen, 2012). The best of the class members were taken for the final class-sums.

Sequence analysis

Emiliania huxleyi, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, *Ectocarpus silicolosus*, *N. gaditana*, *Chlamydomonas reinhardtii*, *Physcomitrella patens*, *Arabidopsis thaliana*, *Oryza sativa*, *Cyanidioschyzon merolae*, *Galdieria sulphuraria* genome databases were accessed online via the National Center for Biotechnology Information (NCBI) portal using TBLASTN and BLASTP. Additional data were collected from http://bioinformatics.psb.ugent.be/genomes/view/Ectocarpus-siliculosus for *E. silicolosus* genome, Department of Energy Joint Genome Institute (JGI) for *E. huxleyi* sequences.

Nannochloropsis gaditana sequences were retrieved from www.nannochloropsis.org using TBLASTN. The multiple sequence alignment (MAS) of protein was performed using CLUSTALW in BioEdit and MUSCLE (MUltiple Sequence Comparison by Log-Expectation). Sequence similarities and secondary structure information were obtained by ESPript, 'Easy Sequencing in PostScript', using the crystal structure of spinach major light-harvesting complex as a reference (1RWT pdb code; Liu *et al.*, 2004).

MS analysis

The green bands visible on the sucrose gradient were harvested with a syringe and then loaded on SDS-PAGE (precast 12% polyacrylamide SDS gel; C.B.S. Scientific, San Diego, CA, USA). In-gel tryptic digestion was performed as described in (Shevchenko *et al.*, 2006), with the minor modification of using

acetonitrile as the organic phase. The MS measurements were performed as described by (Terashima *et al.*, 2010) using an Ultimate 3000 nanoflow HPLC system (Dionex, Sunnyvale, CA, USA) coupled with an LTQ Orbitrap XL mass spectrometer (Thermo Finnigan, Waltham, MA, USA) device for autosampling, column switching and nano-HPLC. For the identification of peptides, OMSSA (v.2.1.4) (Geer *et al.*, 2004). A new database was created by downloading *N. gaditana* sequences from www.nannochloropsis.org/page/ftp to assign the peptides resulting from MS analysis to *N. gaditana* protein.

Label-free quantification

For peptide identification and determination of protein intensities, MS raw data were processed using MAXQUANT v.1.5.1.2 (Cox & Mann, 2008). The Uniprot reference proteome of *N. gaditana* (downloaded 12 December 2014) concatenated with randomized sequences of all entries was used for database search, with default settings for high-resolution MS1 (Orbitrap) and low-resolution MS2 (ion trap). Oxidation of methionine and acetylation of the protein N-terminus were allowed as variable modifications. False discovery rate (FDR) was set to 1% at the peptide and protein level. The feature 'match between runs' was activated. Due to the fundamentally different protein composition of each fraction, intensity normalization by MaxQuant was omitted. Instead, normalization factors were calculated based on the sums of all non-normalized protein intensities of each fraction.

Pigment analysis

Chlorophylls and carotenoids were extracted using 80% acetone, and pigment content was determined by HPLC (Beckman System Gold) as described previously (Färber & Jahns, 1998). The vaucheriaxanthin retention factor was estimated from the one of violaxanthin, with a 10% correction accounting for the different absorption spectrum.

Spectroscopic steady state measurements

Absorption of isolated PSI-LHC particles was measured at 77 K and room temperature (RT) with a Varian Cary 4000 UV-Vis spectrophotometer. For 77 K measurements, samples were in 65% glycerol. Fluorescence spectra were measured at 77 K and RT on a Fluorolog spectrofluorimeter (Jobin Yvon Horiba, Kyoto, Japan). Samples were diluted to Q_y optical density (OD_{Qy}) 0.07 cm^{-1} in order to avoid self-absorption in a buffer containing 10 mM Hepes (pH 7.5) and 0.06% α-DM. The Circular-Dichroism (CD) spectra were measured using a Chirascan-Plus CD Spectrometer at RT.

Time-resolved measurements – streak camera set-up

Picosecond-time-resolved fluorescence measurements were performed with a streak camera setup as previously described (Gobets *et al.*, 2001; van Stokkum & Van Oort, 2008; Le Quiniou *et al.*, 2015b). The samples were measured upon 400 nm excitation. The repetition rate was set at 250 kHz, the pulse energy was below 0.4 nJ. The samples were stirred in all cases. A power study confirmed the absence of annihilation in the measuring conditions.

Fluorescence was detected from 590 nm to 860 nm in three different time ranges (TR): 0 to 155 ps (TR1 temporal response 4–5 ps), 0 to 400 ps (TR2 temporal response 6–7 ps) and 0 to 1500 ps (TR4 temporal response 20 ps). The streak camera data were treated in HPD-TA 8.4.0 (Hamamatsu Photonics, Hamamatsu, Japan) as described in detail in Le Quiniou *et al.* (2015b). The instrument response function (IRF) was modeled with a simple Gaussian for all the TRs. PSI-LHC particles were diluted at an OD of 0.6 cm^{-1} for PSI-LHC$_\alpha$ and 1.2 cm^{-1} for PSI-LHC$_\beta$ at the Qy maximum with a buffer containing 10 mM HEPES (pH 7.5) and 0.06% α-DM. The chosen OD enabled time-resolved measurements without self-absorption.

Data analysis of time-resolved measurements

The streak camera measurements were analyzed globally in Glotaran (Snellenburg *et al.*, 2012) with a sequential model. The average decay time $\tau_{av\ CS}$ (Eqn 1) characterizes the time until charge separation (CS) occurs (for open RCs) with A_n the area under the DAS of the n^{th} component (i.e. its total amplitude) and $A_n / \sum_n A_n$ the relative amplitude.

$$\tau_{av\ CS} = \sum_n (\tau_n \cdot A_n) / \sum_n A_n \qquad \text{Eqn 1}$$

See Le Quiniou *et al.* (2015b) for more details.

Results

Structural analysis reveals unusual organization of the peripheral antenna in the PSI-LHC complexes from *Nannochloropsis gaditana*

PSI-LHC can be purified from *N. gaditana* thylakoids in supercomplexes containing the core complex with its associated antenna (Basso *et al.*, 2014). The structure of PSI-LHC from *N. gaditana* purified after solubilization with either α- or β-DM (PSI-LHC$_\alpha$ and PSI-LHC$_\beta$) was investigated using electron microscopy (EM). Projection maps of negatively stained PSI-LHC$_\alpha$ and PSI-LHC$_\beta$ were obtained by single particle analysis (Scheres *et al.*, 2008; Boekema *et al.*, 2009; Scheres & Chen, 2012). In PSI-LHC$_\alpha$ we found two types of particles: a monomeric complex (Fig. 1a–d) and a dimeric complex (Fig. 1e), in a ratio 8 : 2 (*c.* 230 000 and 58 000 particles, respectively). The monomers have a size comparable to that of the PSI-LHCI isolated from *A. thaliana* (Fig. 1f,g). β-DM gives harsher solubilization than α-DM (Basso *et al.*, 2014) and PSI-LHC$_\beta$ samples showed the presence of monomers with < 1% of the particles that could be classified as dimers. An equivalent P700 : Chl ratio was measured in PSI-LHC$_\alpha$ and PSI-LHC$_\beta$ supporting the hypothesis that monomers and dimers have similar composition (Supporting Information Fig. S1).

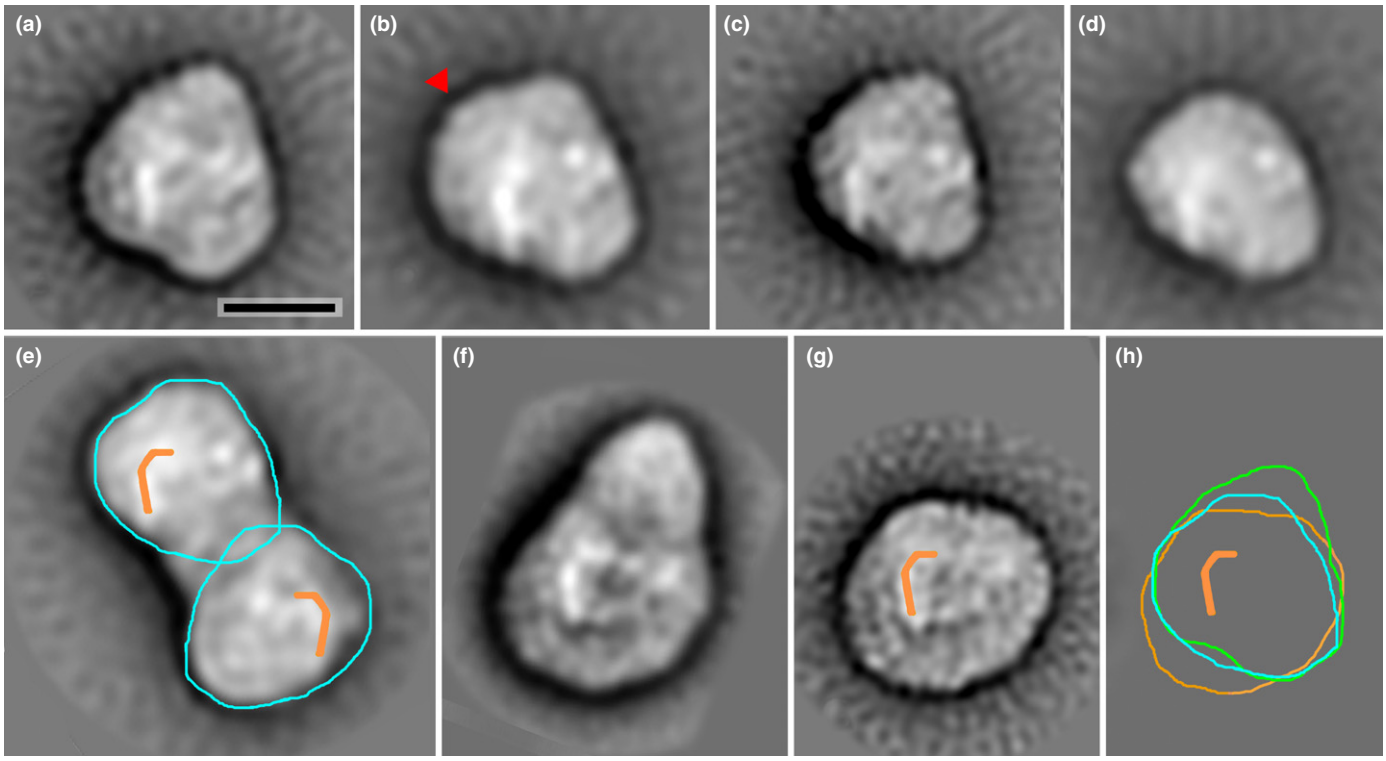

Fig. 1 Electron microscopy analysis of Photosystem I light-harvesting super complex (PSI-LHC) from *Nannochloropsis gaditana*. (a–d) Projection maps of the *N. gaditana* monomer, sums of four best classes after processing 200 000 particles, see text. The red arrowhead in (b) shows extra density, absent in (a); scale bar, 10 nm. (e) Projection map of the *N. gaditana* dimer, after processing 50 000 projections. (f, g) Maps of *Arabidopsis thaliana* PSI-LHC and PSI, reanalyzed from Kouril *et al.* (2005b). (h) Aligned contours of *N. gaditana* (a, green, upper monomer of (e) in blue) and *A. thaliana* (g, orange). The 'reversed J' motif (orange) and other densities were helpful to align the core parts and also to establish that the dimers of *N. gaditana* (e) consist of up and down oriented monomers, as indicated.

In PSI-LHC$_\alpha$ four main types of monomers were found upon classification. The largest ones have a triangular shape (Fig. 1a) and some of them show an extra density on the upper left side (indicted with a red arrow in Fig. 1b). In other particles, this density is less pronounced, which can partly be caused by a tilt of the particles on the carbon support film (Fig. 1c,d). All projection maps show clear densities that in the *A. thaliana* map (Fig. 1g) correspond to the extrinsic subunits of PSI core. These densities can then be used for the alignment for *N. gaditana* particles (green contour in Fig. 1h) with that one of *A. thaliana* (orange contour, Fig. 1h). This comparison shows that the *N. gaditana* particles have a different shape than those of *A. thaliana*, being substantially smaller at the bottom but larger at the top, according to the orientation shown in the figure.

In the dimeric complex the upper monomer has exactly the same outline as the monomer of Fig. 1(d), indicating a small loss of density at the tip (blue contour in Fig. 1e,h). The lower monomer of the dimer is, however, not related by two-fold symmetry to the upper one, as the blue contours plus the orange-marked density clearly demonstrate. This is only compatible with a flipping plus a rotation, indicating that monomers are oriented upside-up and upside-down. Such an opposite orientation of the two monomers suggests that the dimers are artificial, as earlier observed in plants (Kouril *et al.*, 2005a).

The best two maps of *N. gaditana* (Fig. 2a,b) were overlaid with the atomic model of the plant PSI structure (Ben-Shem *et al.*, 2003; Qin *et al.*, 2015). In plants such as pea and *A. thaliana* the LHCA1–4 monomers constitute a belt at one side of the core, whereas in *N. gaditana* there appears to be space only for two LHCs, in a position corresponding to the plant LHCA2/3 dimer, at the lower right tip (Fig. 2a,b). The model further shows space for at least three other LHCs at the other side of the core complex, especially in the more bulky map of Fig. 2(b). This modeling clearly shows that *N. gaditana* PSI has a unique antenna structure, different from plants and green algae, which likely is composed of five LHC subunits associated to two opposite sides of the core complex.

PSI-LHC composition in *Nannochloropsis gaditana*

PSI core complex subunits were identified in the *N. gaditana* genome using similarity searches with the blastp tool at NCBI (using query sequences from *A. thaliana*, *Cyanidioschizon merolae* and *Pheodactylum tricornutum*, Table 1). Sequences of PSI subunits from various additional species from *Viridiplantae* (*A. thaliana thaliana*, *O. sativa*, *P. patens*, *C. reinhardtii*), red algae (*G. sulphuraria*), heterokonts (diatoms and *E. silicosus*) and haptophyta (*E. huxleyi*) also were searched to build a comprehensive view of the subunit distribution in different photosynthetic eukaryotes (Tables 1, S1). Some subunits (PsaA-E), known to be involved in charge separation and electron transport, are highly conserved in all photosynthetic organisms including

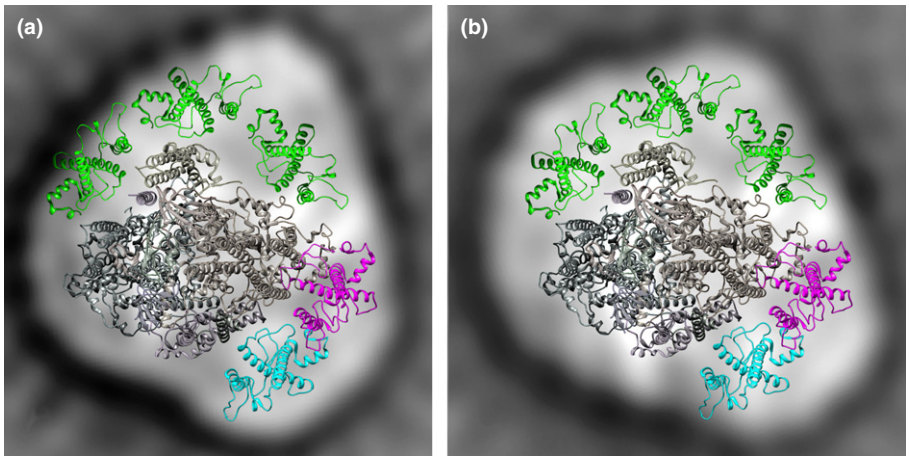

Fig. 2 Model for subunit positions in *Nannochloropsis gaditana* Photosystem I (PSI). (a, b) Projection maps of the *N. gaditana* monomer from Fig. 1(a, b) (respectively) overlaid with wire models of the components of the plant high-resolution protein data bank structure 4XK8 from Qin *et al.* (2015). From the plant model subunits G, H, K and LHCA1/4 were removed and all other subunits were overlaid as a fixed structure and shown in gray. LHCA2/3 are shown in cyan/purple. There is no space left for the two additional light harvesting complexes (LHCs) at the lower left positions next to LHCA2. Instead, additional LHCs (green) have been positioned at the top, to indicate ample space at this site for three extra LHCs, especially (b).

cyanobacteria. Noteworthy, their conservation is not only limited to the presence/absence of the subunits but also extends to their amino-acid sequence (Table S2). Other PSI core subunits (PsaF, PsaI, PsaJ and PsaL) are also conserved in all the species analyzed, suggesting that they play a relevant role in PSI structure and function.

Several other core complex subunits are instead differently distributed among the species analyzed. The *Nannochloropsis gaditana* genome lacks several PSI subunits identified in other organisms, namely PsaG, PsaH, PsaK, PsaM, PsaN and PsaO. As shown in Table 1, PsaG, PsaH and PsaN are present exclusively in some plants and green algae (Khrouchtchova *et al.*, 2005). PsaO is instead present only in primary endosymbiotic groups (plants and green/red algae). PsaM, typical of cyanobacteria is also found in some algae species but is missing in others including *N. gaditana* (Jensen *et al.*, 2007). PsaK showed instead a peculiar distribution because it is conserved from cyanobacteria to red algae and plants but is absent from all heterokonts and haptophytes analyzed, including diatoms and *N. gaditana*. These organisms are all secondary endosymbionts originated from a red alga, thus suggesting that PsaK was present in these organisms' ancestors but was later lost.

A similarity search for antenna sequences allowed identification of three subgroups of LHC in *N. gaditana*, which were classified as LHCf, LHCr and LHCx (Table 2), as described previously (Vieler *et al.*, 2012; Basso *et al.*, 2014). Within the three subgroups, proteins were named according to their RNA expression levels, starting from the most actively transcribed (Alboresi *et al.*, 2016).

Nannochloropsis gaditana thylakoids were solubilized with α-DM and pigment-binding protein complexes were first separated in sucrose gradients by centrifugation and then analyzed by tandem mass spectrometry (MS/MS) (Fig. 3). All PSI core and LHC subunits reported in Tables 1 and 2 with the exception of PsaC were detected by mass spectrometry confirming their presence in *N. gaditana* thylakoids. Although it does not provide an absolute quantification, MS analysis allowed determination of the distribution of each polypeptide between the different bands of the sucrose gradient (i.e. LHC monomers, LHC trimers, PSII core and PSI-LHC). All putative PSI core complex proteins listed in Table 1 showed a strong enrichment (75–100% of the total amount of each peptide) in the PSI-LHC fraction, confirming that they are indeed components of this supercomplex (Fig. 3b). Five antenna subunits, LHCr4-8, showed a distribution along the sucrose gradient bands similar to that of PSI core subunits with a strong enrichment in the PSI-LHC fraction, thus suggesting that they are also specific components of this supercomplex (Fig. 3c). One additional LHC-like subunit (Naga_100030g5), recently attributed to the group of red lineage chlorophyll binding-like proteins (RedCAP1) (Litvín *et al.*, 2016), was also enriched in the PSI-LHC fraction. All other LHCs, instead, presented a different distribution and no strong enrichment in PSI fraction (Fig. S2). It is interesting to point out that three other LHCr proteins (LHCr1-3) were detected in this experiment but not found to be specifically enriched in the PSI-LHC fraction (Fig. S2). In particular, LHCr1 was present mainly as a monomer whereas LHCr3 was accumulated mainly as a monomer/oligomer (*c.* 80%). LHCr2 showed a significant presence in PSI-LHC complexes (*c.* 35%) but also in the band of LHC mono/oligomers (*c.* 35%) and PSII core (*c.* 40%). MS analysis thus shows that these subunits are not associated specifically with PSI or at least not as strongly as LHCr4-8.

Protein sequences of LHCr4-8 from *N. gaditana* were compared with the ones from LHCa1-4 of *A. thaliana* (Fig. 3d). The α-helices A and B are well conserved including the key residues involved in the binding of Chl *a* molecules (Chl 602, 603, 610, 612 and 613). The α-helix C is more variable and a clear ligand for Chl 606 is missing, whereas a glutamic acid likely coordinating Chl 609 can be identified. With the exception of LHCr7, α-helix D is not identifiable in *N. gaditana* antennas and thus the

Table 1 Photosystem I (PSI) core complex subunits

	Viridiplantae				Rhodophyta		Chromoalveolata				
	Plants (Angio-sperm)		Plants (Bryo-phyta)	Green algae (Chloro-phyta)	Red algae		Heterokontophyta				Hapto-phyta
Protein name	*A. thaliana*	*O. sativa*	*P. patens*	*C. reinhardtii*	*G. sulphuraria*	*C. merolae*	*P. tricornutum*	*T. pseudonana*	*E. siliculosus*	*N. gaditana*	*E. huxleyi*
Present in cyanobacteria											
PsaA	+	+	+	+	+	+	+	+	+	**+**	+
PsaB	+	+	+	+	+	+	+	+	+	**+**	+
PsaC	+	+	+	+	+	+	+	+	+	**+**	+
PsaD	+	+	+	+	+	+	+	+	+	**+**	+
PsaE	+	+	+	+	+	+	+	+	+	**+**	+
PsaF	+	+	+	+	+	+	+	+	+	**+**	+
PsaI	+	+	+	+	+	+	+	+	+	**+**	+
PsaJ	+	+	+	+	+	+	+	+	+	**+**	+
PsaK	+	+	+	+	+	+	nd	nd	nd	**nd**	nd
PsaL	+	+	+	+	+	+	+	+	+	**+**	+
PsaM	nd	nd	+	nd	+	+	+	+	+	**nd**	+
Absent in cyanobacteria											
PsaG	+	+	+	+	nd	nd	nd	nd	nd	**nd**	nd
PsaH	+	+	+	+	nd	nd	nd	nd	nd	**nd**	nd
PsaN	+	+	nd	+	nd	nd	nd	nd	nd	**nd**	nd
PsaO	+	+	+	+	+	+	nd	nd	nd	**nd**	nd

The table shows the identification of PSI core subunits in model species from different taxonomic groups: *Arabidopsis thaliana, Oryza sativa*, *Physcomitrella patens*, *Chlamydomonas reinhardtii*, *Galdieria sulphuraria, Cyanidioschyzon merolae*, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*, *Ectocarpus silicolosus*, *Nannochloropsis gaditana*, *Emiliania huxleyi*. Presence/absence in cyanobacteria was retrieved from the literature (Jensen *et al.*, 2007), whereas for the other species sequences were identified (+) or not (nd) with blastp tool at NCBI (using as query sequences from *A. thaliana* or *C. merolae* and *P. tricornutum* for the subunits absent in *A. thaliana*). The *N. gaditana* column, based on sequences identified in this work, is highlighted. Sequence identification numbers are all reported in Supporting Information Table S1.
Bold highlights the *Nannochloropsis gaditana* that was analysed in this study.

Table 2 Light-harvesting complex (LHC) proteins of *Nannochloropsis gaditana*

Protein name	Gene ID	UniProt ID
LHCf		
NgLHCf1	Naga_100012g50	W7T4V5
NgLHCf2	Naga_100005g99	W7TY83
NgLHCf3	Naga_100157g5	K8Z8N4
NgLHCf4	Naga_100168g14	W7TFG9
NgLHCf5	Naga_100017g59	W7TRI0
NgLHCf6	Naga_100004g86	W7TXE6
NgLHCf7	Naga_100013g28	W7U405
NgLHCf8	Naga_100027g19	W7TCK1
LHCr		
NgLHCr1	Naga_100002g18	K8YRV9
NgLHCr2	Naga_100168g13	W7TZB5
NgLHCr3	Naga_100018g45	K8ZB38
NgLHCr4*	Naga_100056g15	W7TJ16
NgLHCr5*	Naga_100092g17	W7TX20
NgLHCr6*	Naga_100434g4	W7TJ16
NgLHCr7*	Naga_100641g3	W7T8I0
NgLHCr8*	Naga_100017g83	W7UAI7
LHCx		
NgLHCx1	Naga_100173g12	K8YWB4
NgLHCx2	Naga_100056g42	K8YZX9
NgLHCX3	Naga_101036g3	W7TIB0
LHC-like		
LHC-like–LIL1*	Naga_100030g5	W7TTN9
LHC-like–LIL2	Naga_101227g1	W7TI84

LHC are regrouped according to different subgroups as in diatoms (Vieler *et al.*, 2012). Sequences of the same subgroup were numbered based on their gene expression levels, starting from the most abundantly transcribed (Alboresi *et al.*, 2016). Gene ID is taken from http://www.nannochloropsis.org/ (Corteggiani Carpinelli *et al.*, 2014). Proteins (UniProt ID) were identified by mass spectrometry using the UniProt reference proteome of *N. gaditana* (UP000019335) for database search. *Enrichment in Photosystem I supercomplex (PSI-LHC). The correspondence with proteins from *N. oceanica* is shown in Supporting Information Table S3.

ligand for Chl a614 is not conserved as previously observed for LHCb6 of *A. thaliana* (Passarini *et al.*, 2009). In LHCa complexes Chl 603 is coordinated either to an asparagine or a histidine in the helix B. The presence of an asparagine, as in the case of LHCa3 and LHCa4, was shown to be responsible of a red-shift in the fluorescence spectrum (Morosinotto *et al.*, 2003). LHCr5-8 have a histidine in that position, whereas LHCr4 has an asparagine suggesting the possible existence of red-shifted forms in the antenna system of *N. gaditana* PSI.

Functional properties of PSI-LHC from *Nannochloropsis gaditana*

As expected Chl*a* is the only Chl species present in PSI-LHC (Table 3). The Chl : Car ratio of PSI-LHC$_\alpha$ is lower than that of PSI-LHC$_\beta$ suggesting that some of the xanthophylls (mainly violaxanthin) are loosely bound to the complex and lost with the stronger solubilization (Table 3). The Chl : Car ratio of PSI-LHC is lower (3.2 ± 0.8) than that of *A. thaliana* (4.8 ± 0.1, Wientjes *et al.*, 2009) and *C. reinhardtii* PSI (5.0 ± 0.2, Drop *et al.*, 2011). When normalized to the same number of Chls, however, the amount of β-carotene (10.2 ± 2.6 and 12.4 ± 2.1 mol per 100 Chls) is comparable to that of *A. thaliana* (13.1 ± 0.3 mol per 100 Chls), whereas the value of xanthophylls, especially violaxanthin is higher. Because the latter are preferentially bound to the antenna proteins, this suggests a relatively higher carotenoid content in this moiety, in agreement with the low Chl : Car ratio of the LHCs of *N. gaditana* (Basso *et al.*, 2014) compared with higher plants (Wientjes *et al.*, 2009) and *C. reinhardtii* (Drop *et al.*, 2011).

The absorption spectra of the PSI particles at 77 K and RT as well as their second derivatives are presented in Figs. 4, S3 and Table S4. The second derivative of the absorption spectra in the Chl region shows the presence of absorption forms *c.* 669 and 680 nm at RT (Fig. S3; Table S4) and at 669, 679.5, 685 and 697.5 nm at 77 K (Fig. 4b,c; Table S4) for both PSI particles. In the carotenoid region, minima of the second derivative are visible at *c.* 486 nm and *c.* 503 nm at 77 K (Fig. 4b; Table S4). The wavelength of the second minimum and the fact that it is similar in the two preparations suggest that it is the signature of β-carotene in the PSI core. However, the large difference in the 486 nm signal between PSI-LHC$_\alpha$ and PSI-LHC$_\beta$ suggests that it originates from xanthophylls (mainly violaxanthin), that are strongly reduced in PSI-LHC$_\beta$. The Circular Dichroism spectra of the two preparations are shown in Fig. 4(d) together with the spectrum of *A. thaliana*. Although small differences can be observed, the main components of the spectra are present in all complexes indicating that the overall pigment organization is conserved, confirming that the two preparations contain the same complex.

The fluorescence emission spectra at RT have maxima at 681.5 and 683 nm for PSI-LHC$_\alpha$ and PSI-LHC$_\beta$, respectively (Fig. 4e). The emission spectra at 77 K (Fig. 4f) of both preparations show an intense band with maximum at 722 nm, typical of the red forms. A second emission band with maximum at 675–680 nm is visible at 77 K, indicating the presence of Chls that do not transfer energy to PSI (see the shift in emission upon 400 nm and 500 nm excitation in Fig. S3c,d).

In order to study excitation energy transfer and trapping, the fluorescence decay kinetics of PSI-LHC$_\alpha$ and PSI-LHC$_\beta$ were measured with a streak camera set-up. The fluorescence was imaged along wavelengths (from 640 and 800 nm) and time (for three different time ranges up to 1500 ps). An example of a streak image is shown in Fig. 5(a). After sequential analysis, the global decay of the two samples is described as a sum of decay components (see the Materials and Methods section) and the decay associated spectra (DAS) are shown in Fig. 5(b). A minimum of four components was needed for a good description of the kinetics of both samples. The first component (with a lifetime of 10.5 ps in PSI-LHC$_\alpha$ and 13.0 ps in PSI-LHC$_\beta$) is mainly a decay component although it still contains some energy transfer features as it can be inferred by the partial absence of the expected positive vibrational band. The second component (with a lifetime of 45.5 ps in PSI-LHC$_\alpha$ and 45.1 ps in PSI-LHC$_\beta$) is a pure decay component and represents the time when most of the trapping occurs. The two slowest components show blue-shifted DAS and long lifetimes (1.7–1.6 ns and 6 ns), and correspond to

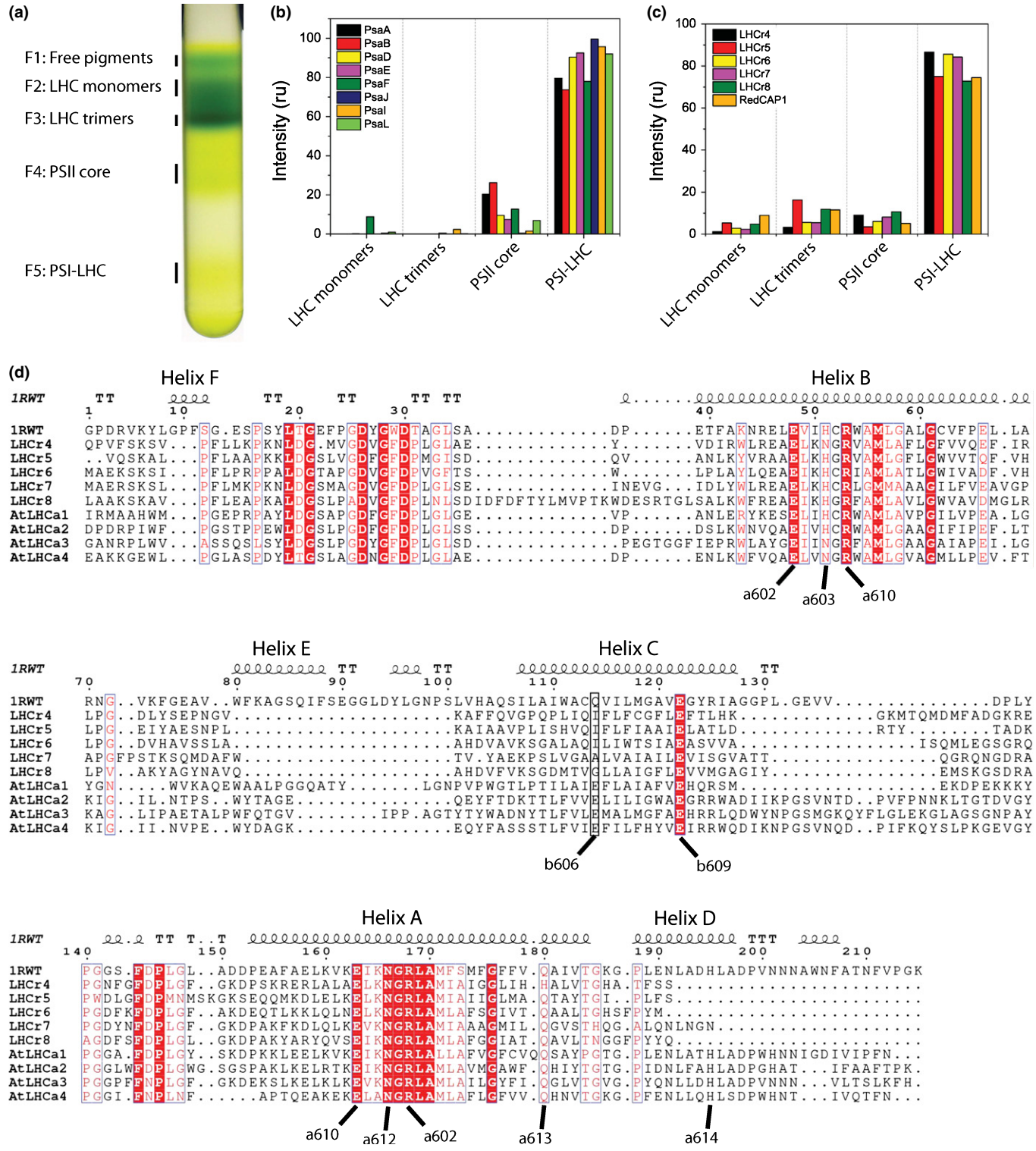

Fig. 3 Protein distribution in the different fractions harvested after sucrose gradient sedimentation of solubilized thylakoids. *Nannochloropsis gaditana* thylakoid membranes were first solubilized with α-DM and then separated by sucrose density gradient ultracentrifugation and subsequently characterized by MS/MS analysis. (a) Example of sucrose density gradient after centrifugation of thylakoid membranes of *N. gaditana*. (b) Distribution of subunits of the Photosystem I (PSI) core complex in the gradient fractions. (c) Light harvesting complexes (LHCs) showing a relative enrichment in PSI fractions, others are shown in Supporting Information Fig. S2. ru, relative units. All *N. gaditana* PSI core subunits and LHCs are listed in Tables 1 and 2. (d) Structure-based sequence alignment of the crystallized spinach LHCII (Liu *et al.*, 2004; pdb code 1RWT) with PSI associated LHCs of *N. gaditana* and *Arabidopsis thaliana*. The secondary structure of the spinach LHCII trimer is shown above the alignment together with the names of the helices. Conserved amino acids highlighted by a red background are identical and those in red letters are similar. Alpha helices are represented as helices. Conserved residues involved in the binding of Chla molecules are labeled under the alignment.

	Viola-xanthin	Vaucheria-xanthin	Antera-xanthin	Zeaxanthin	β-carotene	Chl/Car
PSI-LHC$_\alpha$	14.5 ± 2.0	2.3 ± 0.6	2.1 ± 0.8	2.5 ± 2.2	10.2 ± 2.6	3.2 ± 0.8
PSI-LHC$_\beta$	7.8 ± 2.0	1.1 ± 0.6	< 1	1.2 ± 0.8	12.4 ± 2.1	4.3 ± 0.7

Pigment content of PSI-LHC fractions purified after different thylakoids solubilization is reported, expressed in mol per 100 Chl. Values are reported as mean ± SD ($n > 3$).

Table 3 Pigment content of *Nannochloropsis gaditana* Photosystem I supercomplex (PSI-LHC)

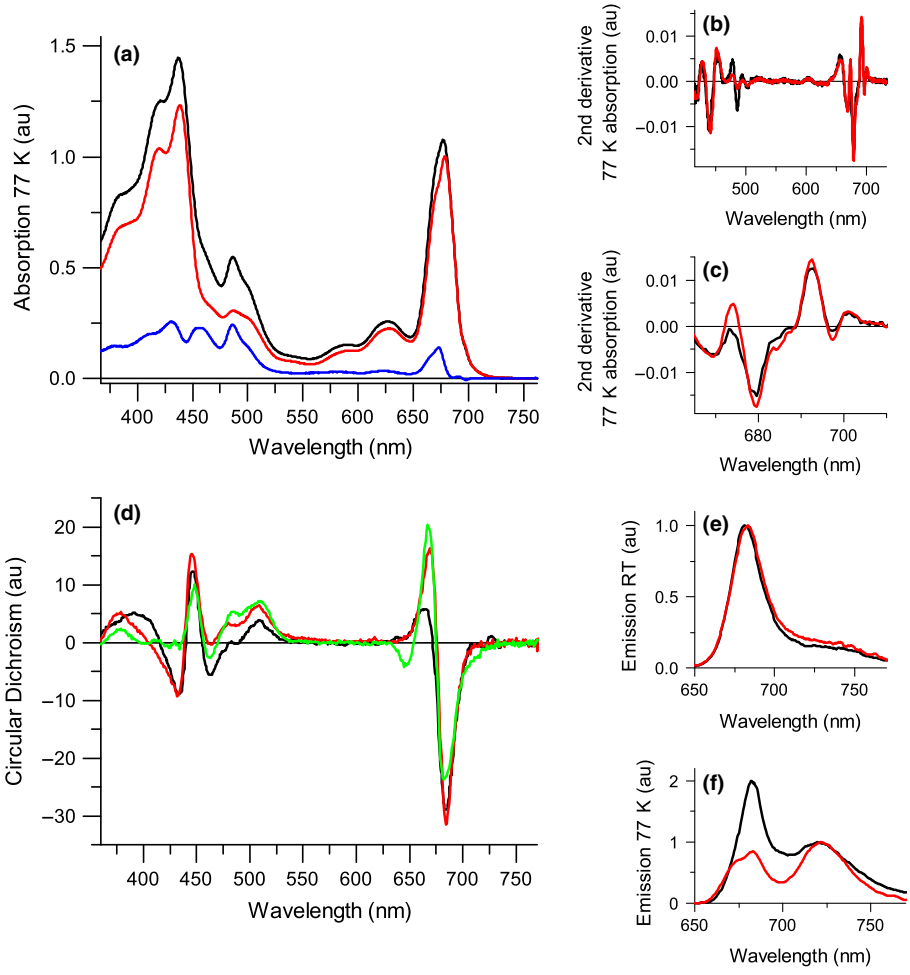

Fig. 4 Spectroscopic characterization of the *Nannochloropsis gaditana* Photosystem I supercomplex (PSI-LHC). (a) Absorption spectra of PSI-LHC$_\alpha$ (black line) and PSI-LHC$_\beta$ (red line) at 77 K and their difference (blue line). The absorption spectra are normalized to the red absorption from 705 nm and above. (b) Second derivative of the 77 K absorption spectra shown in (a). (c) Enlarged view in the Q_y region of the 2nd derivative shown in (b). See Supporting Information Fig. S3(a, b) for absorption and second derivatives at room temperature (RT). (d) Circular dichroism (CD) compared with *Arabidopsis thaliana* PSI-LHCI (green, normalized to the maximum absorption in the Q_y. (e) RT fluorescence emission upon 500 nm excitation, normalized to the maximum emission. See Fig. S3(c, d) for RT emission upon 400 nm. (f) 77 K fluorescence emission upon 400 nm excitation, normalized to the maximum in the red region. au, arbitrary units.

disconnected species. Indeed *c.* 1.7 ns is close to the lifetime of LHCAs of higher plants (Passarini *et al.*, 2010; Wientjes *et al.*, 2011a). The average decay times are calculated by using Eqn 1 (see the Materials and Methods section, Table 4) considering only the components attributed to the PSI-LHC kinetics (thus 1 and 2) and are 31 ps for PSI-LHC$_\alpha$ and 32 ps for PSI-LHC$_\beta$ (± 3 ps). The overall trapping time is thus much shorter than that of higher plants (48 ps; Wientjes *et al.*, 2011b) or *C. reinhardtii* PSI (50 ps; Le Quiniou *et al.*, 2015b).

Discussion

Structural organization and function of PSI-LHC in the heterokont *Nannochloropsis gaditana*

In this work the combination of structural, proteomic and functional analysis provides a comprehensive picture of the structure and composition of Photosytem I light-harvesting complexes (PSI-LHC) from *Nannochloropsis gaditana*. EM analysis of PSI purified from *N. gaditana* shows that this is a supercomplex composed of a core complex and a peripheral antenna system, as in all eukaryotes analyzed thus far (Fig. 1). At variance with plants and other algae (Drop *et al.*, 2011; Thangaraj *et al.*, 2011; Qin *et al.*, 2015), however, five LHCs, identified by MS analysis as LHCr4-8, are found to be associated with *N. gaditana* PSI-LHC (Table 2; Fig. 3c). MS analysis detected three more LHCr-type LHCs, eight LHCf, three LHCx, thus covering the entire LHC superfamily identified in the genome of *N. gaditana* (Table 2). These additional LHCs, however, were not specifically enriched in PSI-LHC suggesting that they are not strongly associated with this supercomplex. It is, however, not possible to exclude the possibility that some additional antenna are loosely associated with PSI *in vivo* but lost during purification.

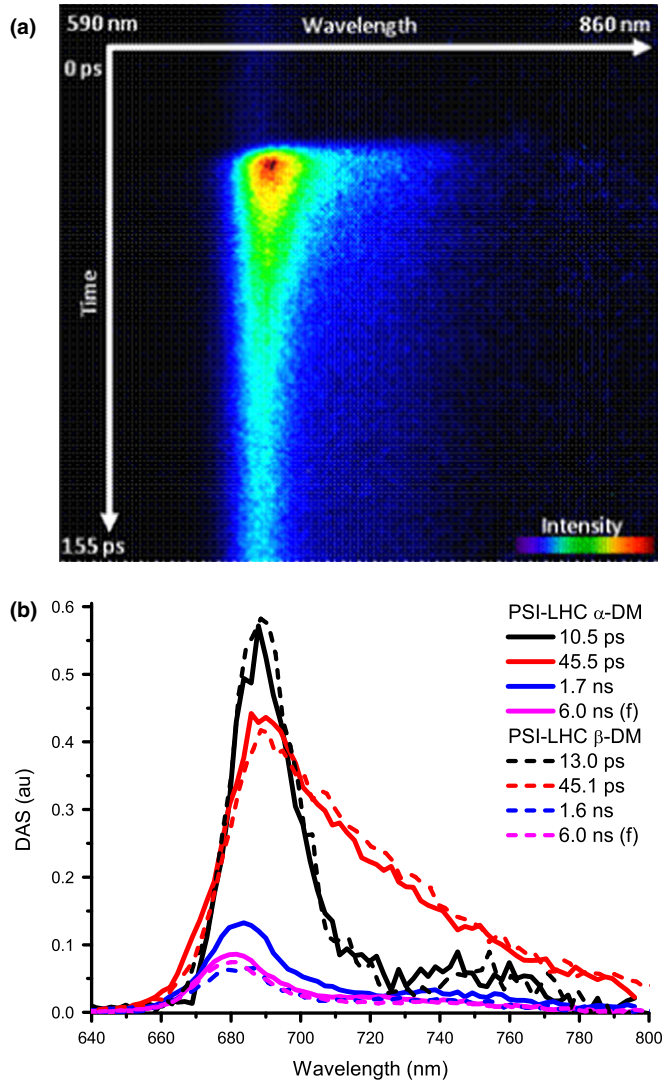

Fig. 5 Time-resolved fluorescence analysis of the Photosystem I supercomplex (PSI-LHC) for *Nannochloropsis gaditana*. (a) Fluorescence decay of the PSI-LHC$_\alpha$ upon 400 nm measured with the streak camera at the shortest time range (0–160 ps; TR1). (b) Decay-associated spectra (DAS) of the kinetics components necessary to describe the fluorescence decay from PSI-LHCI$_\alpha$ (solid lines) and PSI-LHCI$_\beta$ (dashed lines) measured upon 400 nm excitation (normalized to the initial population of excited states in the PSI-LHC particles).

Table 4 Photosystem I supercomplex (PSI-LHC) fluorescence lifetimes

	PSI-LHC α-DM	PSI-LHC β-DM
τ1 (ps)	10.5	13.0
Relative amplitude	33.1%	35.5%
τ2 (ps)	45.5	45.1
Relative amplitude	48.9%	52.0%
τ3 (ns)	1.7	1.6
Relative amplitude	11%	6.3%
τ4 (ns)	6.0 (f)	6.0 (f)
Relative amplitude	7%	6.2%
Average decay time $\tau_{av\ CS}$ (ps)	31	32

Lifetimes were obtained from the sequential analysis of the fluorescence decays of PSI-LHC$_\alpha$ and PSI-LHC$_\beta$ of *Nannochloropsis gaditana* measured upon 400 nm excitation with their relative amplitude (see the Materials and Methods section) and average decay time. The longest lifetime of both samples is attributed to free Chl*a* visible in the steady state emission spectra (Supporting Information Fig. S3c,d) and fixed (f) to 6 ns which corresponds to the free Chl*a* lifetime in acetone (Palacios *et al.*, 2002).

The superimposition of the *N. gaditana* PSI supercomplex EM map with the high-resolution structures of plant PSI evidenced a peculiar arrangement of antenna complexes with two LHCs bound in a highly conserved position, the same occupied by LHCa2/3 in plants (Fig. 2). Three additional LHCs are instead found at the other side of the core complex, where PsaL is located.

As schematized in Fig. 6, this peculiar structural organization correlates with differences in the composition of the PSI core complex derived from genome and proteome analysis (Tables 1, 2; Fig. 3). In plants the association of LHCa1/4 with the core complex is partly mediated by PsaG (Ben-Shem *et al.*, 2003; Qin *et al.*, 2015). PsaG is absent in *N. gaditana* and indeed no LHC was observed in the position corresponding to plant LHCa1/4. PsaF and J have been suggested to mediate interactions with LHCa2/3 in plants (Qin *et al.*, 2015). Consistent with the conservation of PsaF and PsaJ in *N. gaditana*, two LHCs are found in the same position also in this species. Considering that the core complex subunits are well conserved in different organisms (Table 1) this observation also suggests that two LHCs are likely to be bound in this position in PSI from all photosynthetic eukaryotes, including diatoms.

In plants, *psak* knockdown plants showed destabilized LHCa2/3 association (Jensen *et al.*, 2000) suggesting that PsaK contributes to their binding to PSI. The recent structure of plant PSI supercomplex, however, showed that PsaK does not interact directly with the peripheral antennae (Mazor *et al.*, 2015; Qin *et al.*, 2015). PsaK is not found in *N. gaditana* or in any other heterokont analyzed thus far (Table 1) (Le Corguillé *et al.*, 2009; Starkenburg *et al.*, 2014) likely because of a loss that occurred during or after the secondary endosymbiosis. In these organisms the absence of PsaK does not prevent the association of two LHCs on this side of the core, which confirms that PsaK is not strictly necessary for the association of peripheral antenna.

Additional PSI core subunits identified in plants are not conserved in heterokonts. PsaN in plants is found to be associated with PsaF in the docking site for plastocyanin (Haldrup *et al.*, 1999), the soluble PSI electron donor. In *N. gaditana* and other heterokonts only PsaF is present, suggesting that PsaN is dispensable for efficient electron transfer to PSI, consistent with its absence in cyanobacteria. This is likely correlated with a difference in electron transport chain, because *N. gaditana* genome lacks a plastocyanin encoding gene (Corteggiani Carpinelli *et al.*, 2014) and the PSI lumenal electron donor is likely cytochrome *c6*, as previously suggested for the red alga *Galdieria sulphuraria* (Vanselow *et al.*, 2009) and diatoms (Grouneva *et al.*, 2011).

It is also worth underlining the absence of PsaH, a subunit essential for the association of LHCII during state transitions in plants (Lunde *et al.*, 2000). In *N. gaditana* other LHCs associate with the core complex in the region generally occupied by PsaH (Fig. 6). This picture suggests that state transitions, if present in

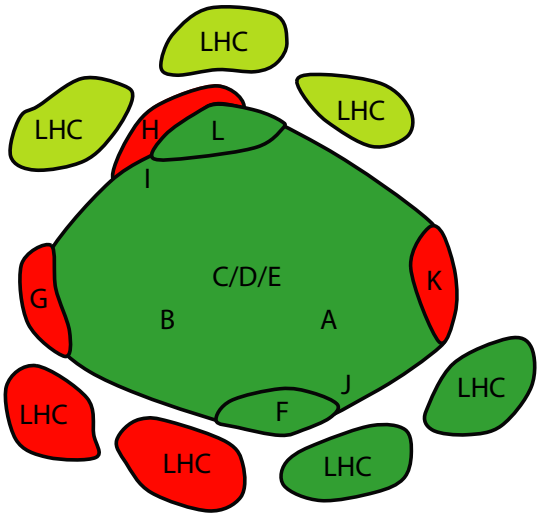

Fig. 6 Schematic comparison of Photosystem I supercomplex (PSI-LHC) structural organization in *Nannochloropsis gaditana* with plants and other algae. Letters from A to K indicate the predicted position of the corresponding Psa subunits. Dark green subunits are those conserved in all eukaryotes (see Tables 1 and 2), red subunits are present in plants, and light green subunits are only in *N. gaditana*.

N. gaditana, most likely involve different structural interactions between antenna and PSI complexes than those described in plants and green algae (Kouril *et al.*, 2005b; Drop *et al.*, 2014).

The functional data show that, despite this different organization, energy transfer and trapping in the PSI complex of *N. gaditana* is very fast. Indeed, the average decay time of the *N. gaditana* PSI-LHC monomer is even shorter (32 ± 3 ps) than that of both *Arabidopsis thaliana* and *Chlamydomonas reinhardtii* PSI-LHCI (*c.* 50 ps; Wientjes *et al.*, 2011b; Le Quiniou *et al.*, 2015b). This difference can be due to a smaller number of pigments associated with the PSI of *N. gaditana* and/or to a difference in the red forms. It is well documented that the number and the energy of these forms influence the excitation energy migration towards the reaction center in PSI (i.e. more red forms, slower transfer) (Gobets *et al.*, 2001; Wientjes *et al.*, 2011b; Le Quiniou *et al.*, 2015b). Although the number of Chls associated with the LHC in *N. gaditana* is not known and we can thus not exclude that the antenna size of PSI in *N. gaditana* is smaller, it is very likely that part of the observed difference in trapping time is due to the diversity in red forms. Indeed, the red forms of PSI-LHC of *N. gaditana* are at a higher energy than those of *A. thaliana* as indicated by their emission maximum (722 nm for *N. gaditana* vs 735 nm for *A. thaliana*) and absorption maximum (697 nm in *N. gaditana* vs 705–710 nm in *A. thaliana*; Morosinotto *et al.*, 2003; Croce *et al.*, 2007) and are then expected to have a smaller influence on the trapping time than the red forms of plants. Independently from the exact origin of this difference, the very fast trapping time observed for PSI-LHC of *N. gaditana* also indicates that all of the LHC subunits are functionally well connected with the core, allowing for fast excitation energy transfer and high quantum efficiency of energy conversion. This high efficiency is common to all PSI complexes analyzed so far (Wientjes *et al.*, 2011b; Le Quiniou *et al.*, 2015a,b) and thus appears to be independent of the organization of the antenna around the core because, for example, the position of the additional LHC in *N. gaditana* differs from that of both LHCI and LHCII in plants and *C. reinhardtii* (Kouril *et al.*, 2005b; Drop *et al.*, 2011, 2014; Qin *et al.*, 2015). This suggests that the design of the PSI core allows the functional association of additional subunits to different part of the complex such that even a PSI core completely surrounded by antennae can maintain a very high quantum efficiency.

Acknowledgements

We thank Dr Gert Oostergetel for critical discussion and Stefania Basso for the preparation of thylakoids from *N. gaditana*. K.N.S.Y. and E.J.B. acknowledge a grant from NWO-ALW. T.M. acknowledges support from ERC starting grant BIOLEAP no. 309485. R.C. acknowledges support from the ERC consolidator grant ASAP no. 281341. M.H. acknowledges support from the German Science foundation (DFG).

Author contributions

T.M. and R.C. planned and designed the research; A.A., C.L.Q., S.Y., M.S., A.M., C.G. and D.S. performed experiments; A.A., C.L.Q., K.N.S.Y., M.S., M.H., E.J.B., R.C. and T.M. analyzed the data; A.A., C.L.Q., S.Y., E.J.B., R.C. and T.M. wrote the manuscript and all authors revised and approved it.

References

Alboresi A, Perin G, Vitulo N, Diretto G, Block MA, Jouhet J, Meneghesso A, Valle G, Giuliano G, Maréchal E *et al.* 2016. Light remodels lipid biosynthesis in nannochloropsis gaditana by modulating carbon partitioning between organelles. *Plant Physiology* **171**: 2468–2482.

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

Archibald JM, Keeling PJ. 2002. Recycled plastids: a 'green movement' in eukaryotic evolution. *Trends in Genetics* **18**: 577–584.

Basso S, Simionato D, Gerotto C, Segalla A, Giacometti GM, Morosinotto T. 2014. Characterization of the photosynthetic apparatus of the Eustigmatophycean *Nannochloropsis gaditana*: evidence of convergent evolution in the supramolecular organization of photosystem I. *Biochimica et Biophysica Acta* **1837**: 306–314.

Ben-Shem A, Frolow F, Nelson N. 2003. Crystal structure of plant photosystem I. *Nature* **426**: 630–635.

Bernal-Bayard P, Pallara C, Carmen Castell M, Molina-Heredia FP, Fernández-Recio J, Hervás M, Navarro JA. 2015. Interaction of photosystem I from *Phaeodactylum tricornutum* with plastocyanins as compared with its native cytochrome c6: reunion with a lost donor. *Biochimica et Biophysica Acta* **1847**: 1549–1559.

Boekema EJ, Dekker JP, van Heel MG, Rögner M, Saenger W, Witt I, Witt HT. 1987. Evidence for a trimeric organization of the photosystem I complex from the thermophilic cyanobacterium *Synechococcus* sp. *FEBS Letters* **217**: 283–286.

Boekema EJ, Folea M, Kouřil R. 2009. Single particle electron microscopy. *Photosynthesis Research* **102**: 189–196.

Bondioli P, Della Bella L, Rivolta G, Chini Zittelli G, Bassi N, Rodolfi L, Casini D, Prussi M, Chiaramonti D, Tredici MR. 2012. Oil production by the marine microalgae *Nannochloropsis* sp. F&M-M24 and *Tetraselmis suecica* F&M-M33. *Bioresource Technology* **114**: 567–572.

Busch A, Hippler M. 2011. The structure and function of eukaryotic photosystem I. *Biochimica et Biophysica Acta* **1807**: 864–877.

Busch A, Nield J, Hippler M. 2010. The composition and structure of photosystem I-associated antenna from *Cyanidioschyzon merolae*. *Plant journal* **62**: 886–897.

Busch A, Petersen J, Webber-Birungi MT, Powikrowska M, Lassen LMM, Naumann-Busch B, Nielsen AZ, Ye J, Boekema EJ, Jensen ON *et al.* 2013. Composition and structure of photosystem I in the moss *Physcomitrella patens*. *Journal of Experimental Botany* **64**: 2689–2699.

Castelletti S, Morosinotto T, Robert B, Caffarri S, Bassi R, Croce R. 2003. Recombinant Lhca2 and Lhca3 subunits of the photosystem I antenna system. *Biochemistry* **42**: 4226–4234.

Cavalier-Smith T. 2004. Only six kingdoms of life. *Proceedings of the Royal Society B* **271**: 1251–1262.

Corteggiani Carpinelli E, Telatin A, Vitulo N, Forcato C, D'Angelo M, Schiavon R, Vezzi A, Giacometti GM, Morosinotto T, Valle G. 2014. Chromosome scale genome assembly and transcriptome profiling of *Nannochloropsis gaditana* in nitrogen depletion. *Molecular Plant* **7**: 323–335.

Cox J, Mann M. 2008. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology* **26**: 1367–1372.

Croce R, Chojnicka A, Morosinotto T, Ihalainen JA, van Mourik F, Dekker JP, Bassi R, van Grondelle R. 2007. The low-energy forms of photosystem I light-harvesting complexes: spectroscopic properties and pigment-pigment interaction characteristics. *Biophysical Journal* **93**: 2418–2428.

Croce R, van Amerongen H. 2013. Light-harvesting in photosystem I. *Photosynthesis Research* **116**: 153–166.

Croce R, Zucchelli G, Garlaschi FM, Jennings RC. 1998. A thermal broadening study of the antenna chlorophylls in PSI-200, LHCI, and PSI core. *Biochemistry* **37**: 17355–17360.

Drop B, Webber-Birungi M, Fusetti F, Kouřil R, Redding KE, Boekema EJ, Croce R. 2011. Photosystem I of *Chlamydomonas reinhardtii* contains nine light-harvesting complexes (Lhca) located on one side of the core. *Journal of Biological Chemistry* **286**: 44878–44887.

Drop B, Yadav KNS, Boekema EJ, Croce R. 2014. Consequences of state transitions on the structural and functional organization of photosystem I in the green alga *Chlamydomonas reinhardtii*. *Plant Journal* **78**: 181–191.

Engelmann E, Zucchelli G, Casazza AP, Brogioli D, Garlaschi FM, Jennings RC. 2006. Influence of the photosystem I-light harvesting complex I antenna domains on fluorescence decay. *Biochemistry* **45**: 6947–6955.

Färber A, Jahns P. 1998. The xanthophyll cycle of higher plants: influence of antenna size and membrane organization. *Biochimica et Biophysica Acta (BBA) – Bioenergetics* **1363**: 47–58.

Gardian Z, Bumba L, Schrofel A, Herbstova M, Nebesarova J, Vacha F. 2007. Organisation of Photosystem I and Photosystem II in red alga *Cyanidium caldarium*: encounter of cyanobacterial and higher plant concepts. *Biochimica et Biophysica Acta* **1767**: 725–731.

Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH. 2004. Open mass spectrometry search algorithm. *Journal of Proteome Research* **3**: 958–964.

Gobets B, van Grondelle R. 2001. Energy transfer and trapping in photosystem I. *Biochimica et Biophysica Acta* **1507**: 80–99.

Gobets B, van Stokkum IHM, Rögner M, Kruip J, Schlodder E, Karapetyan NV, Dekker JP, Van Grondelle R. 2001. Time-resolved fluorescence emission measurements of photosystem I particles of various cyanobacteria: a unified compartmental model. *Biophysical Journal* **81**: 407–424.

Grouneva I, Rokka A, Aro E-M. 2011. The thylakoid membrane proteome of two marine diatoms outlines both diatom-specific and species-specific features of the photosynthetic machinery. *Journal of Proteome Research* **10**: 5338–5353.

Haldrup A, Naver H, Scheller HV. 1999. The interaction between plastocyanin and photosystem I is inefficient in transgenic Arabidopsis plants lacking the PSI-N subunit of photosystem. *Plant Journal* **17**: 689–698.

Ikeda Y, Yamagishi A, Komura M, Suzuki T, Dohmae N, Shibata Y, Itoh S, Koike H, Satoh K. 2013. Two types of fucoxanthin-chlorophyll-binding proteins I tightly bound to the photosystem I core complex in marine centric diatoms. *Biochimica et Biophysica Acta* **1827**: 529–539.

Jansson S. 1999. A guide to the Lhc genes and their relatives in Arabidopsis. *Trends in Plant Science* **4**: 236–240.

Jennings RC, Garlaschi FM, Morosinotto T, Engelmann E, Zucchelli G. 2003. Corrigendum to: the room temperature emission band shape of the lowest energy chlorophyll spectral form of LHCI (FEBS 27430). *FEBS Letters* **549**: 181.

Jensen PE, Bassi R, Boekema EJ, Dekker JP, Jansson S, Leister D, Robinson C, Scheller HV. 2007. Structure, function and regulation of plant photosystem I. *Biochimica et biophysica acta* **1767**: 335–352.

Jensen PE, Gilpin M, Knoetzel J, Scheller HV. 2000. The PSI-K subunit of photosystem I is involved in the interaction between light-harvesting complex I and the photosystem I reaction center core. *Journal of Biological Chemistry* **275**: 24701–24708.

Jordan P, Fromme P, Witt HT, Klukas O, Saenger W, Krauss N. 2001. Three-dimensional structure of cyanobacterial photosystem I at 2.5 A resolution. *Nature* **411**: 909–917.

Khrouchtchova A, Hansson M, Paakkarinen V, Vainonen JP, Zhang S, Jensen PE, Scheller HV, Vener AV, Aro E-M, Haldrup A. 2005. A previously found thylakoid membrane protein of 14 kDa (TMP14) is a novel subunit of plant photosystem I and is designated PSI-P. *FEBS Letters* **579**: 4808–4812.

Kouril R, van Oosterwijk N, Yakushevska AE, Boekema EJ. 2005a. Photosystem I: a search for green plant trimers. *Photochemical & Photobiological Sciences* **4**: 1091–1094.

Kouril R, Zygadlo A, Arteni AA, de Wit CD, Dekker JP, Jensen PE, Scheller HV, Boekema EJ. 2005b. Structural characterization of a complex of photosystem I and light-harvesting complex II of *Arabidopsis thaliana*. *Biochemistry* **44**: 10935–10940.

Le Corguillé G, Pearson G, Valente M, Viegas C, Gschloessl B, Corre E, Bailly X, Peters AF, Jubin C, Vacherie B *et al.* 2009. Plastid genomes of two brown algae, *Ectocarpus siliculosus* and *Fucus vesiculosus*: further insights on the evolution of red-algal derived plastids. *BMC Evolutionary Biology* **9**: 253.

Le Quiniou C, van Oort B, Drop B, van Stokkum IHM, Croce R. 2015a. The high efficiency of Photosystem I in the green alga *Chlamydomonas reinhardtii* is maintained after the antenna size is substantially increased by the association of Light-harvesting Complexes II. *Journal of Biological Chemistry* **290**: 30587–30595.

Le Quiniou C, Tian L, Drop B, Wientjes E, van Stokkum IHM, van Oort B, Croce R. 2015b. PSI-LHCI of *Chlamydomonas reinhardtii*: increasing the absorption cross section without losing efficiency. *Biochimica et Biophysica Acta* **1847**: 458–467.

Li M, Semchonok DA, Boekema EJ, Bruce BD. 2014. Characterization and evolution of tetrameric photosystem I from the thermophilic cyanobacterium *Chroococcidiopsis* sp TS-821. *Plant Cell* **26**: 1230–1245.

Litvín R, Bína D, Herbstová M, Gardian Z. 2016. Architecture of the light-harvesting apparatus of the eustigmatophyte alga *Nannochloropsis oceanica*. *Photosynthesis Research* doi: 10.1007/s11120-016-0234-1.

Liu Z, Yan H, Wang K, Kuang T, Zhang J, Gui L, An X, Chang W. 2004. Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428**: 287–292.

Lunde C, Jensen PE, Haldrup A, Knoetzel J, Scheller HV. 2000. The PSI-H subunit of photosystem I is essential for state transitions in plant photosynthesis. *Nature* **408**: 613–615.

Mazor Y, Borovikova A, Nelson N. 2015. The structure of plant photosystem I super-complex at 2.8 Å resolution. *eLife* **4**: e07433.

Morosinotto T, Breton J, Bassi R, Croce R. 2003. The nature of a chlorophyll ligand in Lhca proteins determines the far red fluorescence emission typical of photosystem I. *Journal of Biological Chemistry* **278**: 49223–49229.

Mozzo M, Mantelli M, Passarini F, Caffarri S, Croce R, Bassi R. 2010. Functional analysis of Photosystem I light-harvesting complexes (Lhca) gene products of *Chlamydomonas reinhardtii*. *Biochimica et Biophysica Acta* **1797**: 212–221.

Nelson N, Yocum CF. 2006. Structure and function of photosystems I and II. *Annual Review of Plant Biology* **57**: 521–565.

Oostergetel GT, Keegstra W, Brisson A. 1998. Automation of specimen selection and data acquisition for protein electron crystallography. *Ultramicroscopy* **74**: 47–59.

Palacios MA, de Weerd FL, Ihalainen JA, van Grondelle R, van Amerongen H. 2002. Superradiance and exciton (de)localization in Light-Harvesting Complex II from green plants? *Journal of Physical Chemistry B* **106**: 5782–5787.

Passarini F, Wientjes E, Hienerwadel R, Croce R. 2009. Molecular basis of light harvesting and photoprotection in CP24: unique features of the most recent antenna complex. *Journal of Biological Chemistry* **284**: 29536–29546.

Passarini F, Wientjes E, van Amerongen H, Croce R. 2010. Photosystem I light-harvesting complex Lhca4 adopts multiple conformations: red forms and excited-state quenching are mutually exclusive. *Biochimica et Biophysica Acta* **1797**: 501–508.

Qin X, Suga M, Kuang T, Shen J-R. 2015. Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. *Science* **348**: 989–995.

Raven JA, Evans MCW, Korb RE. 1999. The role of trace metals in photosynthetic electron transport in O_2-evolving organisms. *Photosynthesis Research* **60**: 111–150.

Riisberg I, Orr RJS, Kluge R, Shalchian-Tabrizi K, Bowers HA, Patil V, Edvardsen B, Jakobsen KS. 2009. Seven gene phylogeny of heterokonts. *Protist* **160**: 191–204.

Rodolfi L, Chini Zittelli G, Bassi N, Padovani G, Biondi N, Bonini G, Tredici MR. 2009. Microalgae for oil: strain selection, induction of lipid synthesis and outdoor mass cultivation in a low-cost photobioreactor. *Biotechnology and Bioengineering* **102**: 100–112.

Scheres SHW, Chen S. 2012. Prevention of overfitting in cryo-EM structure determination. *Nature Methods* **9**: 853–854.

Scheres SHW, Núñez-Ramírez R, Sorzano COS, Carazo JM, Marabini R. 2008. Image processing for electron microscopy single-particle analysis using XMIPP. *Nature Protocols* **3**: 977–990.

Schmid VHR, Cammarata KV, Bruns BU, Schmidt GW. 1997. In vitro reconstitution of the photosystem I light-harvesting complex LHCI-730: heterodimerization is required for antenna pigment organization. *Proceedings of the National Academy of Sciences, USA* **94**: 7667–7672.

Shevchenko A, Tomas H, Havlis J, Olsen JV, Mann M. 2006. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nature Protocols* **1**: 2856–2860.

Simionato D, Basso S, Giacometti GM, Morosinotto T. 2013. Optimization of light use efficiency for biofuels production in algae. *Biophysical Chemistry* **182**: 71–78.

Snellenburg JJ, Laptenok S, Seger R, Mullen KM, van Stokkum IHM. 2012. Glotaran: a Java-based graphical user interface for the R package TIMP. *Journal of Statistical Software* **49**: 1–22.

Starkenburg SR, Kwon KJ, Jha RK, McKay C, Jacobs M, Chertkov O, Twary S, Rocap G, Cattolico RA. 2014. A pangenomic analysis of the *Nannochloropsis* organellar genomes reveals novel genetic variations in key metabolic genes. *BMC Genomics* **15**: 212.

van Stokkum IHM, Van Oort B. 2008. (Sub)-picosecond spectral evolution of fluorescence studied with a synchroscan streak-camera system and target analysis. *Biophysical Techniques in Photosynthesis* **12**: 223–240.

Sukenik A, Livne A, Apt KE, Grossman AR. 2000. Characterization of a gene encoding the light-harvesting violaxanthin-chlorophyll protein of *Nannochloropsis* sp. (eustigmatophyceae). *Journal of Phycology* **36**: 563–570.

Sukenik A, Livne A, Neori A, Yacobi YZ, Katcoff D. 1992. Purification and characterization of a light-harvesting chlorophyll-protein complex from the marine Eustigmatophyte *Nannochloropsis* sp. *Plant and Cell Physiology* **33**: 1041–1048.

Terashima M, Specht M, Naumann B, Hippler M. 2010. Characterizing the anaerobic response of Chlamydomonas reinhardtii by quantitative proteomics. *Molecular & Cellular Proteomics* **9**: 1514–1532.

Thangaraj B, Jolley CC, Sarrou I, Bultema JB, Greyslak J, Whitelegge JP, Lin S, Kouřil R, Subramanyam R, Boekema EJ *et al.* 2011. Efficient light harvesting in a dark, hot, acidic environment: the structure and function of PSI-LHCI from *Galdieria sulphuraria*. *Biophysical Journal* **100**: 135–143.

Vanselow C, Weber APM, Krause K, Fromme P. 2009. Genetic analysis of the Photosystem I subunits from the red alga, *Galdieria sulphuraria*. *Biochimica et Biophysica Acta* **1787**: 46–59.

Veith T, Büchel C. 2007. The monomeric photosystem I-complex of the diatom *Phaeodactylum tricornutum* binds specific fucoxanthin chlorophyll proteins (FCPs) as light-harvesting complexes. *Biochimica et Biophysica Acta* **1767**: 1428–1435.

Vieler A, Wu G, Tsai C-HH, Bullard B, Cornish AJ, Harvey C, Reca I-BB, Thornburg C, Achawanantakun R, Buehl CJ *et al.* 2012. Genome, functional gene annotation, and nuclear transformation of the heterokont oleaginous alga *Nannochloropsis oceanica* CCMP1779. *PLoS Genetics* **8**: e1003064.

Watanabe M, Kubota H, Wada H, Narikawa R, Ikeuchi M. 2011. Novel supercomplex organization of photosystem I in *Anabaena* and *Cyanophora paradoxa*. *Plant & Cell Physiology* **52**: 162–168.

Watanabe M, Semchonok DA, Webber-Birungi MT, Ehira S, Kondo K, Narikawa R, Ohmori M, Boekema EJ, Ikeuchi M. 2014. Attachment of phycobilisomes in an antenna–photosystem I supercomplex of cyanobacteria. *Proceedings of the National Academy of Sciences, USA* **111**: 2512–2517.

Wientjes E, Oostergetel GT, Jansson S, Boekema EJ, Croce R. 2009. The role of Lhca complexes in the supramolecular organization of higher plant photosystem I. *Journal of Biological Chemistry* **284**: 7803–7810.

Wientjes E, van Stokkum IHM, van Amerongen H, Croce R. 2011a. Excitation-energy transfer dynamics of higher plant photosystem I light-harvesting complexes. *Biophysical Journal* **100**: 1372–1380.

Wientjes E, van Stokkum IHM, van Amerongen H, Croce R. 2011b. The role of the individual Lhcas in photosystem I excitation energy trapping. *Biophysical Journal* **101**: 745–754.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information tab for this article:

Fig. S1 Quantification of P700 content in PSI-LHC isolated from *Nannochloropsis gaditana*.

Fig. S2 Protein distribution in the different fractions harvested after sucrose gradient sedimentation.

Fig. S3 Spectra at room temperature of PSI-LHC$_\alpha$ and PSI-LHC$_\beta$.

Table S1 Sequence identification numbers

Table S2 Identity matrix for PsaA proteins

Table S3 List of LHC protein orthologs in *Nannochloropsis gaditana* and *N. oceanica*

Table S4 Spectroscopic data analysis of PSI-LHC$_\alpha$ and PSI-LHC$_\beta$

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.