The singular properties of photosynthetic cytochrome c_{550} from the diatom *Phaeodactylum tricornutum* suggest new alternative functions

Pilar BERNAL-BAYARD[1†], Consolación ÁLVAREZ[1†], Purificación CALVO[2], Carmen CASTELL[1], Mercedes RONCEL[1], Manuel HERVÁS[1], José A. NAVARRO[1*]

[1]Instituto de Bioquímica Vegetal y Fotosíntesis, cicCartuja, Universidad de Sevilla and CSIC, Sevilla, Spain.

[2]Departamento de Microbiología, Facultad de Biología, Universidad de Sevilla, Sevilla, Spain.

*Corresponding author: José A. Navarro, Instituto de Bioquímica Vegetal y Fotosíntesis, Centro de Investigaciones Científicas Isla de la Cartuja, Universidad de Sevilla & CSIC, Américo Vespucio 49, 41092-Sevilla, Spain. Fax: +34 954460165; E-mail: jnavarro@ibvf.csic.es

†These authors contributed equally to this work

ABSTRACT

Cytochrome c_{550} is an extrinsic component in the luminal side of photosystem II in cyanobacteria, as well as in eukaryotic algae from the red photosynthetic lineage including, among others, diatoms. We have established that cytochrome c_{550} from the diatom *Phaeodactylum tricornutum* can be obtained as a complete protein from the membrane fraction of the alga, although a C-terminal truncated form is purified from the soluble fractions of this diatom as well as from other eukaryotic algae. Eukaryotic cytochromes c_{550} show distinctive electrostatic features as compared with cyanobacterial cytochrome c_{550}. In addition, co-immunoseparation and mass spectrometry experiments, as well as immunoelectron microscopy analyses, indicate that although cytochrome c_{550} from *P. tricornutum* is mainly located in the thylakoid domain of the chloroplast –where it interacts with photosystem II–, it can also be found in the chloroplast pyrenoid, related with proteins linked to the CO_2 concentrating mechanism and assimilation. These results thus suggest new alternative functions of this heme protein in eukaryotes.

Abbreviations:

ß-DM, ß-dodecyl-maltoside; CA, carbonic anhydrase; Cc_{550}, cytochrome c_{550}; Cc_6, cytochrome c_6; CCM, CO_2-concentrating mechanism; DSS, disuccinimidyl suberate; FBA, fructose-bisphosphate aldolase; LC-MS/MS, liquid chromatography tandem-mass spectrometry; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; MS, mass spectrometry; MW, molecular weight; PSI, photosystem I; PSII, photosystem II; RubisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase.

INTRODUCTION

Photosynthetic cytochrome c_{550} (Cc_{550}) is a *c*-type heme protein (molecular mass around 15 kDa, encoded by the *psbV* gene) with an unusual bis-histidinyl axial coordination (Frazão et al. 2001) that acts as an extrinsic subunit of photosystem II (PSII) by binding stoichiometrically to the luminal surface of this photosynthetic complex (Shen 2015, Ago et al. 2016). In spite of its redox character, the role of Cc_{550} in PSII appears to be just structural, by stabilizing the Mn_4CaO_5 cluster and binding of Cl^- and Ca^{2+} ions, as a redox function of the Cc_{550} heme cofactor in PSII has not yet been established (Shen and Inoue 1993, Shen et al. 1998, Enami et al. 2003, Nagao et al. 2010). Cc_{550} is present in cyanobacteria and in eukaryotic algae from the red photosynthetic lineage, which comprises, among others, red algae and diatoms, but is absent in the green lineage, which includes green algae and plants, that have replaced Cc_{550} for the Ca^{2+} binding PsbP subunit, which lacks any cofactor (revised in Roncel et al. 2012).

In many organisms Cc_{550} can be easily purified as a soluble protein (Navarro et al. 1995, Kerfeld and Krogmann 1998, Bernal-Bayard et al. 2017), and the existence of two different populations of Cc_{550} (bound to the PSII or free in the lumen) has been postulated (Kirilovsky et al. 2004). Soluble Cc_{550} has a very low midpoint redox potential (ranging from –190 to –314 mV) (Alam et al. 1984, Navarro et al. 1995, Roncel et al. 2003, Bernal-Bayard et al. 2017) incompatible with a redox function in PSII. However, much more positive potential values were determined when bound to PSII (from –80 to +200 mV) (Roncel et al. 2003, Guerrero et al. 2011). Several alternative roles have been proposed for the soluble Cc_{550} in cyanobacteria, as in anaerobic carbon and hydrogen metabolism (Krogmann 1991, Morand et al. 1994), cyclic photophosphorylation (Kienzel and Peschek 1983) or in the reduction of nitrate to ammonium (Alam et al. 1984), but none of these possible functions have been clearly established.

Diatoms belong to the red lineage of algae that diverged from the green lineage that finally evolved to higher plants. In addition to Cc_{550} (or PsbV), the assembly of extrinsic proteins at the lumenal side of diatom PSII includes the cyanobacterial-like subunits PsbO and PsbU, as well as the PsbQ' subunit, also present in red algae (Enami et al. 2003, Nagao et al. 2010), and an additional extrinsic protein, named as Psb31 (Okumura et al. 2008, Nagao et al. 2017). However, although the structure of diatom PSII has not yet been solved, the crystal structure of PSII from the red alga *Cyanidium caldarium* has shown an overall structure similar to the cyanobacterial complex, including the position of Cc_{550} (Ago et al. 2016).

In order to optimize photosynthetic efficiency, the chloroplasts of diatoms display a refined thylakoid architecture (Flori et al. 2017), which includes the presence of the pyrenoid, an electron-dense semi-crystalline protein aggregate present in the chloroplast of most unicellular eukaryotic algae and considered the site for CO_2 fixation, as the enzyme Ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) is preferentially localized in this protein body,

accounting for over 90% of the pyrenoid's protein content (reviewed in Badger et al. 1998, Meyer et al. 2017). In addition, the presence of carbonic anhydrase (CA) enzymes in the pyrenoid has been reported (Tachibana et al. 2011, Sinetova et al. 2012, Hopkinson et al. 2016, Kikutani et al. 2016). Thus, the role of the pyrenoid is to sustain efficient carbon fixation by increasing the CO_2 concentration (CO_2 concentrating mechanism, CCM) in the vicinity of RubisCO (Giordano et al. 2005, Matsuda et al. 2017). Moreover, in most unicellular eukaryotic algae, including diatoms, the pyrenoid is penetrated by one or more thylakoid lamellae (Bedoshvili et al. 2009, Engel et al. 2015, Meyer et al. 2017). In the marine diatom *Phaeodactylum tricornutum*, in particular, the pyrenoid matrix is crossed by a single thylakoid lamella that contains a luminal θ-type CA, that adds to the two ß-type CAs, PtCA1 and PtCA2, located inside the pyrenoid (Tachibana et al. 2011, Kikutani et al. 2016). Components of the PSI complex have also been found in pyrenoids, indicating that intrapyrenoid thylakoids could supply ATP to this compartment via cyclic electron transport, thus covering the energetic demands of carbon fixation at the same time as avoiding the evolution of inhibitory O_2 (McKay and Gibbs 1991, Meyer and Griffiths 2013). Consequently, it was generally accepted that intrapyrenoid thylakoids are enriched in PSI and lack PSII, although in the green alga *Chlamydomonas reinhardtii* the presence of PSII components has been recently reported (Mackinder et al. 2017). However, the presence in this space of either the soluble photosynthetic cytochrome c_6 (Cc_6) or the Cc_{550}, has not been yet addressed.

Recently, the photosynthetic Cc_{550} from *P. tricornutum* has been purified from algal cells and characterized (Bernal-Bayard et al. 2017). The purified protein was described as being truncated in the last hydrophobic residues of the C-terminus. Moreover, a weaker affinity of Cc_{550} for the diatom PSII complex was also described (Bernal-Bayard et al. 2017). Here we provide additional insights into the peculiar properties of Cc_{550} from the diatom *P. tricornutum*. First, we have determined that it is possible to purify Cc_{550} as the complete non-truncated form. In addition, *P. tricornutum* Cc_{550} has been immunolocalized in the pyrenoid of the chloroplast, in close contact with proteins previously described as located in this microcompartment. Thus, in *P. tricornutum* the Cc_{550} could play a yet to be defined role related to the pyrenoid function.

MATERIALS AND METHODS

Cell cultures

Cells of the coastal pennate diatom *Phaeodactylum tricornutum* CCAP 1055/1 were grown in Artificial Seawater (ASW) medium (12 µM Fe, iron-replete culture) (McLachlan 1964, Goldman and McCarthy 1978) in a rotatory shaker (50 rpm) at 20ºC, with regular transfer of the cells into fresh media. The cultures were illuminated by white led lamps giving an intensity of 4.35 W m^{-2} (T8-150MWBL led lamps, Wellmax, China) under a light/dark cycle of 16/8 h. *Chaetoceros muelleri*, *Nannochloropsis gaditana* and *Isochrysis galbana* cells were obtained from indoor cultures from the *C.P.I.F.P. Marítimo Zaporito* (San Fernando, Spain).

Protein purification

Purification of truncated Cc_{550} from *P. tricornutum*, *C. muelleri*, *N. gaditana* and *I. galbana* cells was carried out essentially as recently described in Bernal-Bayard et al. (2017), in the presence of the protease inhibitors PMSF (1 mM), benzamidine (1 mM), aminocaproic acid (100 mM), D-phenylalanine (10 mM) and hydrocinnamic acid (1 mM). Purification of the complete Cc_{550} from *P. tricornutum* membrane fractions was performed essentially following the protocol previously described to obtain PSII-enriched samples (Bernal-Bayard et al. 2017) with some modifications. Basically, *Phaeodactylum* cells were disrupted three times in a French press at 15,000 psi, the ß-DM incubation was done for 45 min, and the solubilized samples were loaded onto a 0.17 to 0.3 M sucrose continuous density gradient. Finally, the complete Cc_{550} was collected from the top gradient band, corresponding to the free protein (Bernal-Bayard et al. 2017). For the MS analysis, samples were exhaustively washed and concentrated by ultrafiltration with 5 mM Tris-HCl, pH 8.5, buffer.

Protein samples for in-gel peptide fingerprint analysis of *P. tricornutum* Cc_{550} were obtained from crude cell extracts resolved on polyacrylamide gel electrophoresis. Fresh *P. tricornutum* cells were washed twice in ice-cold PBS 20 mM, pH 8.0, buffer, and resuspended in 25 mM Tris-HCl, pH 8.5, 50 mM NaCl buffer, supplemented with 0.5% Triton X-100, DNase and the protease inhibitors PMSF, benzamidine and aminocaproic acid. Samples were incubated for 5 min at 70ºC followed by a cycle of French-press disruption (20,000 psi). Samples were then centrifuged at 170,000x*g* for 15 min, and the resultant supernatant was resolved on 20% (w/v) SDS-PAGE and visualized by Coomassie blue-staining.

For the co-immunoseparation and immunodetection of Cc_{550}, polyclonal antibodies raised against this cytochrome were purified by their affinity to pure Cc_{550}, by using 4 mg of Cc_{550} (Bernal-Bayard et al. 2017) and the AminoLink Plus Immobilization Kit, 2 mL (Thermo Scientific, USA), at pH 7.2, according to the instructions of the supplier. Antibodies against Cc_6 (Roncel et al. 2016) and the large and small plant RubisCO subunits (Agrisera, Sweden) were

also used in control experiments. For immunodetection of *P. tricornutum* RubisCO subunits, initial crude extracts were in some cases enriched in RubisCO by precipitation with PEG 4000 (20% w/v), as previously described (Haslam et al. 2005).

DNA analysis

Genomic DNA extraction was carried out using a simplified CTAB-extraction procedure (Lukowitz et al. 2000). Total extracted DNA from *C. muelleri* and *I. galbana* was used as template to amplify by PCR the corresponding *psbV* genes, by using adequate oligonucleotide pairs (Fig. S1, supporting material). Thus, whereas *C. muelleri* gene amplification was completed by using oligonucleotide primers (C1-F and C2-R) designed from the *C. gracilis* sequence, *I. galbana psbV* gene amplification and sequencing was done in two steps. First, using degenerate oligonucleotide primers designed from a consensus sequence from the *C. gracilis*, *P. tricornutum* and the five closest *psbV* genes (according to a BLAST search) allowed the amplification and sequencing of the 3'-end (242 pb), that showed a 83% identity with a similar sequence of the *psbV* gene from the Isochrysidal alga *Emiliania huxleyi*. Thus, a second PCR-amplification was carried out to obtain the *psbV* 5'-end by using oligonucleotide primers designed from *E. huxleyi.* In all cases, the PCR fragments were cloned for subsequent sequencing in the pGEM-T vector (Promega, USA) according to the protocol of the supplier. DNA sequencing was carried out in the sequencing service of the company Secugen (Spain).

Co-immunoseparation analysis

Three different samples were used in the co-immunoseparation experiments, as follows: First, fresh *P. tricornutum* cells, resuspended in 25 mM Tris-HCl, pH 7.5, 50 mM NaCl and 0.1% Triton X-100 buffer, supplemented with DNase and the protease inhibitors PMSF, benzamidine and aminocaproic acid, were disrupted in a French press at 20,000 psi (three cycles). Samples were then centrifuged at 170,000x*g* for 30 min, and the resultant supernatant was considered as the soluble fraction. Second, *P. tricornutum* cells resuspended in 50 mM Tris-HCl, pH 7.5, 50 mM NaCl and 5 mM EDTA buffer (buffer A) supplemented with proteases inhibitors, DNase and 1 M betaine, were disrupted at a lower pressure (7,500 psi, two cycles). Unbroken cells were separated by centrifugation at 5,000x*g* for 5 min and the supernatant was centrifuged at 170,000x*g* for 30 min. The resultant pellet was resuspended in buffer A at 1 mg Chl mL^{-1} and later diluted to 0.5 mg Chl mL^{-1} with the same volume of ß-dodecyl-maltoside (ß-DM) 3% (w/v) prepared in buffer A. The solution was then incubated 30 min in the dark at 4ºC under gentle stirring followed by centrifugation at 170,000x*g* for 30 min. The resulting detergent-solubilized supernatant was diluted with buffer A to a ß-DM concentration of 0.1% and concentrated in an Amicon pressure cell (10 K cutoff filter). The final concentrated solution was considered as the membrane-extracted fraction. Last, crosslinked samples, with the crosslinking reagent

disuccinimidyl suberate (DSS; Thermo Scientific), were obtained from fresh *P. tricornutum* cells by following the instructions of the supplier. Cells were resuspended in PBS 20 mM (pH 8.0) buffer, and washed three times with ice-cold PBS (pH 8.0) to remove amine-containing agents from the culture media. Cells were then treated with DSS to a 5 mM final concentration, followed by 30 min incubation at room temperature. Crosslinking reactions were stopped by the addition of Quench Solution (1M Tris-HCl, pH 7.5) to a final concentration of 20 mM Tris, followed by 15 min incubation at room temperature. Samples were then centrifuged at 5,000x*g* for 5 min and pellets were resuspended in 20 mM Tris-HCl, pH 7.5, 50 mM NaCl and 5 mM EDTA buffer, supplemented with protease inhibitors and DNase, and disrupted in a French press at 20,000 psi (three cycles). Unbroken cells were discarded by centrifugation, and cell lysates were diluted to 1 mg Chl mL^{-1} and treated with ß-DM as described previously for the membrane-extracted fraction. The resulting final concentrated supernatant was considered as the crosslinked fraction. In all cases the protein content of the samples was quantified using the Lowry method, whereas chlorophyll concentration was determined according to Arnon (1949).

Immunopurifications using µ columns and µMACS^{TM} separator (Miltenyi Biotec, Germany) were carried out following the instructions of the supplier. An amount of 5 mg of protein mixed with 5 µL of purified polyclonal antibodies against Cc_{550} were used, and the eluted fraction was analysed by LC-MS/MS., and the eluted fraction was analysed by LC-MS/MS. Control immunoseparation experiments, using the preimmune serum, were also carried out and the identified proteins were subtracted from those obtained when using the immune serum.

Mass spectrometry analysis

MALDI-TOF MS and LC-MS/MS analyses were performed at the IBVF *Proteomic Service* (Sevilla, Spain). Tryptic digestion of in-gel *P. tricornutum* proteins for peptide fingerprint analysis and MALDI-TOF MS were carried out as previously described (Bernal-Bayard et al. 2017). Tryptic digestion in solution of immunoseparated samples, as well as peptide analysis by liquid chromatography and MS, were performed basically as previously described (Vowinckel et al. 2013) on a Tandem Quadrupole Time-of-Flight mass spectrometer (AB Sciex TripleTOF5600) coupled to a Nanospray III Ion Source (AB Sciex) and nano-HPLC (Eksigent Ekspert nanoLC 425). Peptide separation was first carried out on a pre-column (Acclaim PepMap 100 C18, 5 µm, 100 Å, 100 µm id x 20 mm, Thermo Fisher Scientific) and then eluted onto the analytical column (New Objective PicoFrit column, 75 µm id x 250 mm, emitter included and packed with Reprosil-PUR 3 µm). Data acquisition was achieved as previously described (Vowinckel et al. 2013). The ion accumulation time was set to 250 ms (MS) and to 65 ms (MS/MS), resulting in a total duty cycle of 2.89 s.

LC-MS/MS data acquired in DDA mode were analyzed and processed basically as described in Vowinckel et al. (2013), by using the Paragon algorithm (Protein Pilot software, AB

Sciex, v. 5.0.1) and the reference proteome of *P. tricornutum* in the UniProt database of protein sequences. AB Sciex contaminants and rabbit proteins from the UniProt database were discarded. Only peptides with a confidence score > 0.05 were considered for further analysis.

Immunoelectron microscopy

P. tricornutum cells pellets were fixed in 2% glutaraldehyde in 0.1 M cacodylate buffer, pH 7.3, for 2 h at 4ºC, washed several times in the same buffer at 4 ºC and dehydrated through a graded ethanol series. Finally, cells were embedded in LR White medium grade resin (Sigma-Aldrich) as described (Lichtlé et al. 1992).

For immunolocalization, paled gold sections were cut with an ultramicrotome (Reichert-Jung Ultracut, Austria) with a diamond knife and mounted on nickel grids. Ultrathin sections were treated with the rabbit purified antiserum anti-Cc_{550} for 1 h, and the secondary antibody (goat anti-rabbit 5-nm gold particles, GAR G5 EM, Janssen Life Sciences, Belgium) was used at a dilution of 1:100 for 1 h. Finally, the grids were contrasted with 2% aqueous uranyl acetate for 8 min and observed with a Zeiss Libra 120 Electron Microscope at 80 kV. Two controls were performed: i) replacing the primary antibody for a drop of PBS buffer; and ii) incubation with the primary antibody anti-Cc_6 (luminal soluble carrier) at a dilution of 1:1000 for 1 h.

Structural models

Structures of Cc_{550} from *C. muelleri*, *N. gaditana* and *I. galbana* were modeled using the SWISS MODEL Workspace platform (https://swissmodel.expasy.org/interactive) (Guex and Peitsch, 1997), using as templates the crystal structures of Cc_{550} from the cyanobacterium *Synechocystis* sp. PCC 6803 (pdb 1E29) and the red alga *Cyanidium caldarium* (pdb 4YUU). The representation of protein surface electrostatic potentials was performed by using the Swiss-Pdb Viewer program (https://spdbv.vital-it.ch/). Although interactions of the amino acid side chains –that can change their pKas and therefore their charge in solution– are not taken into account, this program is a useful tool for comparative purposes.

RESULTS AND DISCUSSION

In a previous paper it was shown that *Phaeodactylum tricornutum* Cc_{550} is purified from the soluble fraction as a truncated protein in the two last tyrosine residues of the C-terminus (Bernal-Bayard et al. 2017). Similar results have been here observed (not shown) even though the purification was carried out in the presence of broad-spectrum proteases inhibitors (5 mM EDTA; 1 mM PMSF), as well as specific inhibitors of either serine proteases (1 mM benzamidine; 100 mM aminocaproic acid) or carboxypeptidases competitive inhibitors (10 mM D-phenylalanine; 1 mM hydrocinnamic acid) (Elkins-Kaufman and Neurath, 1949, Hartsuck and Lipscomb, 1971). In this work we have, however, developed a protocol for the purification of the complete protein by detergent-solubilization from the membrane fraction and the further isolation of the free protein from the top part of sucrose gradients (Fig. 1A). Thus, a MALDI-TOF MW analysis of the solubilized Cc_{550} showed a MW (ca. 15,433 Da) that fits with the expected MW based in the *psbV* gene sequence (15,438 Da) (Fig. 1B). The occurrence of the complete holoprotein was also confirmed by tryptic digestion and MS peptide fingerprint analysis (not shown).

A procedure was designed to check if the soluble or the membrane-associated Cc_{550} correspond to different physiological forms (a soluble truncated protein and a complete protein bound to PSII), and not to the exposition to cell proteases in the soluble fraction during the purification course. *P. tricornutum* crude cell extracts were heated at 70^{o}C and directly resolved on polyacrylamide gel electrophoresis, in order to inactivate possible proteolytic enzymes and shorten the process of proteins separation. Extracted gel spots in the MW range corresponding to Cc_{550} were analyzed by tryptic digestion and MS peptide fingerprint (Fig. S2, supporting material). The fingerprint peptide analysis accurately covered the 100% of the amino acid sequence corresponding to the complete non-truncated protein, whereas the theoretical peptide fingerprint for the truncated protein could not be correctly fitted to the obtained data (Fig. S2, supporting material), which finally indicates that truncation is a non-physiological process. Remarkably, the enzymatic activity responsible of Cc_{550} truncation seems to be widespread among several lines of the red lineage of eukaryotic algae, as truncated Cc_{550} was purified from *Chaetoceros muelleri* (a marine centric diatom), *Nannochloropsis gaditana* (Eustigmatophyte) and *Isochrysis galbana* (Haptophyte, Isochrysidales). However, this activity seems to be absent in cyanobacteria, as Cc_{550} from *Synechocystis* sp. PCC 6803, used as a control, showed the expected MW for the complete protein (not shown).

In the case of *C. muelleri* and *I. galbana*, previous sequencing of their *psbV* genes was required in order to establish the MW of both complete proteins. The *C. muelleri psbV* gene showed only 11 nucleotide variations as compared with the gene of *C. gracilis*, but resulting only in two aminoacid changes in the protein transit peptide (not shown). The *I. galbana psbV* gene, by its turn, showed 83% identity with the equivalent gene of the also Isochrysidal alga *Emiliania*

huxleyi, and the protein alignment (94% identity) is shown in Fig. S1 (supporting material). Modelled structures of *C. muelleri*, *I. galbana* and *N. gaditana* Cc_{550} were obtained from the available sequences (Fig. 2). As previously described in the case of the *P. tricornutum* protein, the three Cc_{550} show a common folding similar to that described in cyanobacteria and red algae (Fig. 2) and conserve the hydrophobic northern finger (according to the orientation presented in Fig. 2) previously described (Frazão et al. 2001). However, the surface of eukaryotic Cc_{550} also shows exclusive electrostatic features as compared with cyanobacterial Cc_{550}. Thus, whereas in the prokaryotic protein the cofactor exposed area holds a negatively charged electrostatic character (Frazão et al. 2001, Bernal-Bayard et al. 2017), in the eukaryotic Cc_{550} the area around the heme group is mainly hydrophobic, and the negative electrostatic potential is restricted to the southern area, opposite to the hydrophobic northern protuberance (Fig. 2). According to the PSII known structures, the facing surface of Cc_{550} displayed in Fig. 2, that includes de cofactor exposed area, is involved in the binding to the photosystem, maintaining close contacts with the PSII surface (Ago et al. 2016). Consequently, the distinctive surface charge distribution of eukaryotic Cc_{550} could be significant when establishing the binding affinity to PSII, as previously suggested for the *P. tricornutum* Cc_{550} (Bernal-Bayard et al. 2017).

Although Cc_{550} can be obtained from soluble cell extracts in different organisms (Evans and Krogmann 1983, Navarro et al. 1995, Kerfeld and Krogmann, 1998, Bernal-Bayard et al., 2017), this fact is particularly significant in *P. tricornutum,* where about 60-85% of total Cc_{550} is solubilized during the process of cell disruption in the absence of added detergents (Bernal-Bayard et al., 2017). A similar result has been here observed during the purification of this protein from *C. muelleri*, *N. gaditana* and *I. galbana* cells (not shown). These results could be justified by a combination of a weaker affinity for PSII and an enhanced PSII turnover, as previously suggested in diatoms (Lavaud et al. 2016), both resulting in a higher fraction of unbound Cc_{550}. In this sense, most of the protein still attached to membrane fractions in *P. tricornutum* has been shown to be released by relatively weak detergent extraction procedures, thus suggesting a comparatively weaker affinity of Cc_{550} for PSII (Bernal-Bayard et al., 2017). A less intense affinity for PSII may open the possibility of new functions for an increased fraction of unbound Cc_{550}. We have explored this possibility by performing a co-immunoseparation (from the endogenous Cc_{550} already present in the different samples) and mass spectrometry analysis, in order to identify possible novel protein interactors of *P. tricornutum* Cc_{550}. Three different cellular fractions were immunoseparated with purified Cc_{550}-specific antibodies and analysed by LC-MS/MS: i) the supernatant obtained after cell disruption at high pressure in a non-osmotically stabilized medium followed by sample ultracentrifugation (soluble fraction); ii) the detergent-extracted fraction obtained from membrane pellets after cell breaking at low pressure in an osmotically stabilized medium (membrane fraction); and iii) the supernatant obtained after DSS

treatment of whole cells followed by cell disruption and ultracentrifugation in the presence of detergent (crosslinked fraction).

A relatively low number of proteins were detected by LC-MS/MS in the three co-immunoseparated samples, more particularly in the crosslinked fraction (data available via ProteomeXchange identifier PXD008763). Moreover, in order to minimize the occurrence of false positive or redundant interactions, LC-MS/MS data were filtered according to the following criteria: i) possible protein partners have to be described as chloroplastic proteins and/or predicted to have a transit-peptide targeting to this organelle; and ii) undefined predicted proteins, ribosomal proteins and fucoxanthin-chlorophyll light-harvesting antenna proteins were discarded. From these criteria, a preliminary list of potential targets of Cc_{550} was initially obtained (Fig. 3 *left*, and see Fig. S3, supporting data), which includes RubisCO. Although this enzyme is a highly abundant chloroplast protein complex in plants, it has been shown that in microalgae RubisCO has a much lower abundance, representing less than 6% of total protein (Losh et al. 2013). In addition, the use of the purified Cc_{550}-specific antibodies in western-blot experiments with cell extracts showed a major recognition of Cc_{550}, and no bands corresponding to the predominant components of the pyrenoid (*i.e.*, RubisCO subunits or carbonic anhidrase enzymes) were detected (Fig. S4, supporting data). Thus, we consider RubisCO as a reliable co-immunoseparated target of *P. tricornutum* Cc_{550}.

Coimmunoseparation in the soluble or membrane fractions does not firmly demonstrate a specific physiological interaction, since the method involves putting together proteins that are actually in different cellular compartments, which can lead to describing artefactual interactions. Thus, protein cross-linking in whole cells has been used to demonstrate close contacts between Cc_{550} and other proteins as well as co-localization in the same subcellular compartment. Ultimately, validated protein targets where selected as appearing both in the crosslinked fraction and at least in one of the other two (soluble or membrane) fractions (Fig. 3 *left*). This restrictive criterion discarded two PSII-associated proteins that appeared both in the soluble and membrane fractions, but not in the crosslinked one (Fig. S3, supporting data), as they are the PSII extrinsic PsbQ' subunit and a FKBP peptidylprolyl isomerase protein, that in plants is related to the PSII assembly (Gollan et al. 2012). The final list of co-immunoseparated proteins fitting the restrictive stablished criteria is shown in Fig. 3 (*right*). It is important to note that most proteins that co-immunoseparated in the crosslinked fraction have been already annotated as being located in the chloroplast (Fig. 3, and see Fig. S3, supporting data), so validating the reliability of the method used. In addition to several PSII subunits (Psb31, CP43, D1 and D2), expected from the Cc_{550} location in the red algal PSII structure (Ago et al. 2016), only a very limited number of proteins co-immunoseparated in the three types of samples analyzed (Fig. 3), including RubisCO (large and small subunits) and PtCA1 (a ß-CA), both located in the pyrenoid compartment (Satoh et al. 2001, Tachibana et al. 2011, Kikutani et al. 2012). Only another protein of unknown function,

classified as a CreA-like protein, also co-immunoseparated in the three samples. The CreA family is a group of carbon metabolism transcription regulators described in other organisms (Ries et al. 2016). However, although distantly related to CreA proteins, this CreA-like protein of *P. tricornutum* lacks the typical sequence signatures corresponding to zinc fingers for DNA-binding, typically found in canonical CreA proteins, and its transit peptide indicates a chloroplast targeting. Moreover, a sequence-based structural analysis predicted the existence of a transmembrane domain in this protein (not shown). On the other hand, another pyrenoid protein, the fructose-bisphosphate aldolase (FBA; Allen et al. 2012), was detected both in the crosslinked and soluble fractions (Fig. 3, *right*). Finally, two proteins were detected both in the crosslinked and membrane fractions: the photosystem I PsaD subunit and the ATP-synthase α subunit, and these two results can be attributed to non-specific interactions of luminal Cc_{550} with the two thylakoid membrane complexes (ATP-synthase and PSI) to which these subunits belong. Actually, other ATP-synthase and PSI subunits were detected, but only in the immunoseparated membrane fractions (Fig. S3, supporting data). Nevertheless, all the potential new targets of Cc_{550} here identified are located in the pyrenoid (RubisCO, PtCA1, FBA) or may be related to carbon metabolism (CreA-like protein).

To confirm the possible presence of Cc_{550} in the pyrenoid we have carried out an immunoelectron microscopy analysis of the location of Cc_{550} in *P. tricornutum* cells, by using purified polyclonal antibodies specific against this protein. As a control, the location of the luminal soluble Cc_6 carrier has been also studied. Fig. 4 shows the electron microscopy results for the immunodetection experiments. As expected, Cc_6 seems to be located both in the stromal and the intrapyrenoid thylakoids but not in the pyrenoid matrix (Fig. 4B), in agreement with its functional association with PSI. However, although Cc_{550} showed the expected chloroplastic localization, it appears not only in the stromal and intrapyrenoid thylakoids, but also in the pyrenoid matrix (Fig. 4A). Thus, the association of Cc_{550} with the pyrenoid in *P. tricornutum* is supported both by co-immunoseparation and immunoelectron microscopy analysis.

It is important to note that our results do not imply that Cc_{550} functionally interacts with all the pyrenoid targets included in Fig. 3 (RubisCO, PtCA1 or FBA), as possible cross interactions between these proteins can result in indirect coimmunoseparations with Cc_{550}. However, our results strongly support the pyrenoid localization of Cc_{550}, confirmed by both the in vivo crosslinking with pyrenoid proteins and the pyrenoid immunolocalization (Figs. 3 and 4). The alternative could be the existence of cross-reactions between the purified antibodies against Cc_{550} with any of the observed pyrenoid targets, which can result in parallel immunoseparations and immunodetections. However, we consider that this possibility can be ruled out, as western blot experiments have not shown signs of the occurrence of such cross-reactions (Fig. S4, supporting data).

The association of Cc_{550} with the pyrenoid has not been previously reported and undoubtedly represents an intriguing result. Eukaryotic Cc_{550} is a chloroplast-encoded protein that has a transit peptide, similar to those from cyanobacteria, for targeting into the thylakoid lumen, and the interaction with RubisCO (or other pyrenoid proteins), located in the stromal compartment is thus difficult to justify. In addition, PSII activity has been initially described to be restricted to stromal thylakoids (McKay and Gibbs 1991). However, very recently the presence of both intrinsic and extrinsic PSII subunits in the pyrenoid of the green alga *Chlamydomonas reinhardtii* has been reported (Mackinder et al. 2017, Zhan et al. 2018), including the localization of the PsbQ extrinsic component of PSII in intrapyrenoid thylakoids (Mackinder et al. 2017). Thus, the presence of the PSII extrinsic PsbV subunit (Cc_{550}) in this compartment could not be totally unexpected.

We only can speculate in order to explain a location of Cc_{550} in the pyrenoid of *P. tricornutum*. In *C. reinhardtii* the described presence of PSII components in the pyrenoid could be related with the existence of a network of pyrenoid-penetrating tubules from the intrapyrenoid thylakoids. However, although an equivalent network has been described in the pyrenoid of red algae it has not been found in diatoms (Engel et al. 2015, Meyer et al. 2017). On the other hand, in unicellular algal species, it has been proposed that the pyrenoid plays an important role as defining the starting point of thylakoidal maturation and structural nucleation (reviewed in Rast et al. 2015). In particular, in *C. reinharttii* the pyrenoid is connected to biogenesis centers involved in the assembly of PSII (but not PSI) as ribosomes and mRNAs encoding PSII subunits are localized in the pyrenoid periphery (Uniacke and Zerges, 2007). In addition, PSII mutants impaired in PSII assembly showed an accumulation of early PSII intermediates at the pyrenoid (Uniacke and Zerges, 2007). A similar process in diatoms could maybe explain the presence of PSII subunits into or in close contact with the pyrenoid of *P. tricornutum*.

On the other hand, although the role (if any) that Cc_{550} could play in the pyrenoid matrix may be a matter of discussion, it should be related to carbon fixation or the CCM located in this microcompartment. Interestingly, the existence of two PSII putative Ca^{2+} binding PsbP-type proteins, PSBP3 and PSBP4, has been reported in the pyrenoid of *C. reinhardtii* (Mackinder et al. 2017). However, there is no evidence of the presence of equivalent PsbP-like proteins in diatoms. Actually, PsbP is considered to be evolutionarily recruited as a replacement of Cc_{550} in the photosynthetic green lineage, playing a similar role stabilizing the binding of Ca^{2+}, and thus increasing PSII affinity for this ion (Frazão et al. 2001, Roncel et al. 2012). Moreover, a Ca^{2+}-binding protein, CAS, has been also recently shown to specifically localize into the intrapyrenoid thylakoids of *C. reinhardtii* (Wang et al., 2016). Thus, the presence of Cc_{550} in the pyrenoid of *P. tricornutum* could be related to its features as a calcium-binding stabilizing protein.

Author contributions.

JAN and MH conceived the project and carried out the protein purification experiments. PB-B, CA and CC carried out the DNA analysis and the co-immunoseparation and immunodetection experiments. PC carried out the immunoelectron microscopy experiments with inputs from MR. JAN, MH and MR discussed the results and wrote the manuscript, which was corrected, revised and approved by all authors.

Acknowledgements.
This work was supported by the Spanish Ministry of Economy, Industry and Competitiveness (BIO2015-64169-P) and the Andalusian Government (PAIDI BIO-022). These grants were partially financed by the EU FEDER Program.
The authors thank Rocío Rodríguez (Proteomic Service, IBVF) for technical assistance and Dr. José M. Ortega for critically reading the manuscript. Authors also thank Ignacio Algarín (*C.P.I.F.P. Marítimo Zaporito*, San Fernando, Spain) for supplying us with algae fresh cultures of *C. muelleri*, *N. gaditana* and *I galbana*, and Dr. Ignacio Luque for his help in the co-immunoseparation experiments.

REFERENCES

Ago H, Adachi H, Umena Y, Tashiro T, Kawakami K, Kamiya N, Tian L, Han G, Kuang T, Liu Z, Wang F, Zou H, Enami I, Miyano M, Shen J-R (2016) Novel features of eukaryotic photosystem II revealed by its crystal structure analysis from a red alga. J Biol Chem 291:5676–5687

Alam J, Sprinkle MA, Hermodson MA, Krogmann DW (1984) Characterization of cytochrome c-550 from cyanobacteria. Biochim Biophys Acta 766:317–321

Allen AE, Moustafa A, Montsant A, Eckert A, Kroth PG, Bowler C (2012) Evolution and functional diversification of fructose bisphosphate aldolase genes in photosynthetic marine diatoms. Mol Biol Evol 29:367–379

Arnon DI (1949) Copper enzymes in isolated chloroplasts. Plant Physiol 24:1–15

Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO_2-concentrating mechanisms in algae. Can J Bot 76: 1052–1071

Bedoshvili YD, Popkova TP, Likhoshway YV (2009) Chloroplast structure of diatoms of different classes. Cell Tissue Biol 3:297–310

Bernal-Bayard P, Puerto-Galán L, Yruela I, García-Rubio I, Castell C, Ortega JM, Alonso PJ, Roncel M, Martínez JI, Hervás M, Navarro JA (2017) The photosynthetic cytochrome c_{550} from the diatom *Phaeodactylum tricornutum*. Photosynth Res 133:273–287

Elkins-Kaufman E, Neurath H (1949) Structural requirements for specific inhibitors of carboxypeptidase. J Biol Chem 178:645–654

Enami I, Iwai M, Akiyama A, Suzuki T, Okumura A, Katoh T, Tada O, Ohta H, Shen J-R (2003) Comparison of binding and functional properties of two extrinsic components, Cyt c_{550} and a 12 kDa protein, in cyanobacterial PSII with those in red algal PSII. Plant Cell Physiol 44:820–827

Engel BD, Schaffer M, Cuellar LK, Villa E, Plitzko JM, Baumeister W (2015) Native architecture of the *Chlamydomonas* chloroplast revealed by in situ cryo-electron tomography. eLife 4:1–29

Evans PK, Krogmann DW (1983) Three c-type cytochromes from the red alga *Porphyridium cruentum*. Arch Biochem Biophys 227:494–510

Flori S, Jouneau P-H, Bailleul B, Gallet B, Estrozi LF, Moriscot C, Bastien O, Eicke S, Schober A, Bártulos CR, Maréchal E, Kroth PG, Petroutsos D, Zeeman S, Breyton C, Schoehn G, Falconet D, Finazzi G (2017) Plastid thylakoid architecture optimizes photosynthesis in diatoms. Nat Commun 8, 15885:1–9

Frazão C, Enguita FJ, Coelho R, Sheldrick GM, Navarro JA, Hervás M, De la Rosa MA, Carrondo MA (2001) Crystal structure of low-potential cytochrome c_{549} from *Synechocystis* sp. PCC 6803 at 1.21 Å resolution. J Biol Inorg Chem 6:324–332

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

Goldman JC, McCarthy JJ (1978) Steady state growth and ammonium uptake of a fast-growing marine diatom 1. Limnol Oceanogr 23:695–703

Gollan PJ, Bhave M, Aro E-M (2012) The FKBP families of higher plants: Exploring the structures and functions of protein interactions specialists. FEBS Lett 586:3539–3547

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modelling. Electrophoresis 18:2714–2723

Guerrero F, Sedoud A, Kirilovsky D, Rutherford AW, Ortega JM, Roncel M (2011) A high redox potential form of cytochrome c_{550} in Photosystem II from *Thermosynechococcus elongatus*. J Biol Chem 286:5985–5894

Hartsuck JA, Lipscomb WN. Carboxypeptidase A. In The Enzymes, Vol. 3: Hydrolisis: Peptide Bonds (PD Boyer, ed.) Third Edition. Academic Press, New York & London, 1971, pp. 1–56

Haslam RP, Keys AJ, Andralojc PJ, Madgwick PJ, Andersson I, Grimsrud A, Eilertsen HC, Parry MAJ. Specificity of diatom Rubisco. In Plant Responses to Air Pollution and Global Change (K Omasa, I Nouchi, LJ De Kok, eds.) Springer, Tokyo, 2005, pp. 157–164

Hopkinson BM, Dupont CL, Matsuda Y (2016) The physiology and genetics of CO_2 concentrating mechanisms in model diatoms. Curr Opi Plant Biol 31:51–57

Kerfeld CA, Krogmann DW (1998) Photosynthetic cytochromes c in cyanobacteria, algae and plants. Annu Rev Plant Physiol Plant Mol Biol 49:397–425

Kienzel PF, Peschek GA (1983) Cytochrome c-549: an endogenous cofactor of cyclic photophosphorylation in the cyanobacterium *Anacystis nidulans*. FEBS Lett 162:76–80

Kikutani S, Nakajima K, Nagasato C, Tsuji Y, Miyatake A, Matsuda Y (2016) Thylakoid luminal θ-carbonic anhydrase critical for growth and photosynthesis in the marine diatom *Phaeodactylum tricornutum*. Proc Nat Acad Sci USA 113:9828–9833

Kirilovsky D, Roncel M, Boussac A, Wilson A, Zurita JL, Ducruet JM, Bottin H, Sugiura M, Ortega JM, Rutherford AW (2004) Cytochrome c_{550} in the cyanobacterium *Thermosynechococcus elongatus*: study of redox mutants. J Biol Chem 279:52869–52880

Krogmann DW (1991) The low-potential cytochrome c of cyanobacteria and algae. Biochim Biophys Acta 1058:35–37

Lavaud J, Six C, Campbell DA (2016) Photosystem II repair in marine diatoms with contrasting photophysiologies. Photosynth Res 127:189–199

Lichtlé C, McKay RML, Gibbs SP (1992) Immunogold localization of photosystem I and photosystem II light-harvesting complexes in cryptomonad thylakoids. Biol Cell 74:187–194

Losh JL, Young JN, Morel FMM (2013) Rubisco is a small fraction of total protein in marine phytoplankton. New Phytol 198:52–58

Lukowitz W, Gillmor CS, Scheible W-R (2000) Positional cloning in *Arabidopsis*. Why it feels good to have a genome initiative working for you. Plant Physiol 123:795–805

Mackinder LCM, Chen C, Leib RD, Patena W, Blum SR, Rodman M, Ramundo S, Adams CM, Jonikas MC (2017) A spatial interactome reveals the protein organization of the algal CO_2-concentrating mechanism. Cell 171:133–147

Matsuda Y, Hopkinson BM, Nakajima K, Dupont CL, Tsuji Y (2017) Mechanisms of carbon dioxide acquisition and CO_2 sensing in marine diatoms: a gateway to carbon metabolism. Phil Trans R Soc B 372(1728):20160403

McKay RML, Gibbs SP (1991) Composition and function of pyrenoids: cytochemical and immunocytochemical approaches. Canadian J Botany 69:1040–1052

McLachlan J (1964) Some considerations of the growth of marine algae in artificial media. Can J Microbiol 10:769–782

Meyer M, Griffiths H (2013) Origins and diversity of eukaryotic CO_2-concentrating mechanisms: lessons for the future. J Exp Botany 64:769–786

Meyer MT, Whittaker C, Griffiths H (2017) The algal pyrenoid: key unanswered questions. J Exp Botany 68:3739–3749

Morand LZ, Cheng RH, Krogmann DW (1994) Soluble electron transfer catalysts of cyanobacteria. In: The Molecular Biology of Cyanobacteria (Bryant DA, ed.). Kluwer Academic Publishers, Dordrecht, pp. 381–407

Nagao R, Moriguchi A, Tomo T, Niikura A, Nakajima S, Suzuki T, Okumura A, Iwai M, Shen J-R, Ikeuchi M, Enami I (2010) Binding and functional properties of five extrinsic proteins in oxygen-evolving photosystem II from a marine centric diatom, *Chaetoceros gracilis.* J Biol Chem 285:29191–29199

Nagao R, Suzuki T, Okumura A, Kihira T, Toda A, Dohmae N, Nakazato K, Tomo T (2017) Electrostatic interaction of positive charges on the surface of Psb31 with photosystem II in the diatom *Chaetoceros gracilis*. Biochim Biophys Acta Bioenerg 1858:779–785

Navarro JA, Hervás M, De la Cerda B, De la Rosa MA (1995) Purification and physicochemical properties of the low potential cytochrome c_{549} from the cyanobacterium *Synechocystis* sp. PCC 6803. Arch Biochem Biophys 3186:46–52

Okumura A, Nagao R, Suzuki T, Yamagoe S, Iwai M, Nakazato K, Enami I (2008) A novel protein in Photosystem II of a diatom *Chaetoceros gracilis* is one of the extrinsic proteins located on lumenal side and directly associates with PSII core components. Biochim Biophys Acta 1777:1545–1551

Rast A, Heinz S, Nickelsen J (2015) Biogenesis of thylakoid membranes. Biochim Biophys Acta 1847:821–830

Ries LNA, Beattie SR, Espeso EA, Cramer RA, Goldman GH (2016) Diverse regulation of the CreA carbon catabolite repressor in *Aspergillus nidulans*. Genetics 203:335–352

Roncel M, Boussac A, Zurita JL, Bottin H, Sugiura M, Kirilovsky D, Ortega JM (2003) Redox properties of the photosystem II cytochromes b_{559} and c_{550} in the cyanobacterium *Thermosynechococcus elongatus*. J Biol Inorg Chem 8:206–216

Roncel M, González-Rodríguez AA, Naranjo B, Bernal-Bayard P, Lindahl AM, Hervás M, Navarro JA, Ortega JM (2016) Iron deficiency induces a partial inhibition of the photosynthetic electron transport and a high sensitivity to light in the diatom *Phaeodactylum tricornutum*. Front Plant Sci 7:1050

Roncel M, Kirilovsky D, Guerrero F, Serrano A, Ortega JM (2012) Photosynthetic cytochrome c_{550}. Biochim Biophys Acta 1817:1152–1163

Satoh D, Hiraoka Y, Colman B, Matsuda Y (2001) Physiological and molecular biological characterization of intracellular carbonic anhydrase from the marine diatom *Phaeodactylum tricornutum.* Plant Physiol 126:1459–1470

Shen J-R (2015) The structure of photosystem II and the mechanism of water oxidation in photosynthesis. Annu Rev Plant Biol 66:23–48

Shen J-R, Inoue Y (1993) Binding and functional properties of two new extrinsic components, cytochrome c-550 and a 12-kDa protein, in cyanobacterial photosystem II. Biochemistry 32:1825–1832

Shen J-R, Qian M, Inoue Y, Burnap RL (1998) Functional characterization of *Synechocystis* sp. 6803 ΔpsbU and ΔpsbV mutants reveals important roles of cytochrome c-550 in cyanobacterial oxygen evolution. Biochemistry 37:1551–558

Sinetova MA, Kupriyanova EV, Markelova AG, Allakhverdiev SI, Pronina NA (2012) Identification and functional role of the carbonic anhydrase Cah3 in thylakoid membranes of pyrenoid of *Chlamydomonas reinhardtii*. Biochim Biophys Acta 1817:1248–1255

Tachibana M, Allen AE, Kikutani S, Endo Y, Bowler C, Matsuda Y (2011) Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. Photosynth Res 109:205–221

Uniacke J, Zerges W (2007) Photosystem II assembly and repair are differentially localized in *Chlamydomonas*. Plant Cell 19:3640–3654

Vowinckel J, Capuano F, Campbell K, Deery MJ, Lilley KS, Ralsera M (2013) The beauty of being (label)-free: sample preparation methods for SWATH-MS and next-generation targeted proteomics. F1000 Res 2:272

Wang L, Yamano T, Takane S, Niikawa Y, Toyokawa C, Ozawa S-i, Tokutsu R, Takahashi Y, Minagawa J, Kanesaki Y, Yoshikawa H, Fukuzawa H (2016) Chloroplast-mediated regulation of CO_2-concentrating mechanism by Ca^{2+}-binding protein CAS in the green alga *Chlamydomonas reinhardtii*. Proc Nat Acad Sci USA 113:12586–12591

Zhan Y, Marchand CH, Maes A, Mauries A, Sun Y, Dhaliwal JS, Uniacke J, Arragain S, Jiang H, Gold ND, Martin VJJ, Lemaire SD, Zerges W. (2018) Pyrenoid functions revealed by proteomics in *Chlamydomonas reinhardtii*. PLoS One. Feb 26:1–20

Supporting information

Aditional supporting information is available in the online version of this article:

Appendix Fig. S1. Sequences of oligonucleotides used in this work and alignment of sequences of *I. galbana* and *E. huxleyi* eukaryotic cytochromes c_{550}.

Appendix Fig. S2. Summary for peptide fingerprint analysis of *P. tricornutum* Cc_{550} directly resolved on polyacrylamide gel electrophoresis.

Appendix Fig. S3. Co-immunoseparated proteins identified by LC-MS/MS in soluble, membrane and crosslinked fractions of *P. tricornutum*.

Appendix Fig. S4. Western blot analysis of crude cellular fractions of *P. tricornutum* with antibodies against either Cc_{550} or the RubisCO large and small subunits.

Data Deposition

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD008763. The *psbV* gene sequences from *Chaetoceros muelleri* and *Isochrysis galbana* are deposited in the NCBI databank, GenBank accession numbers MG779498 and MG779497, respectively.

FIGURE LEGENDS

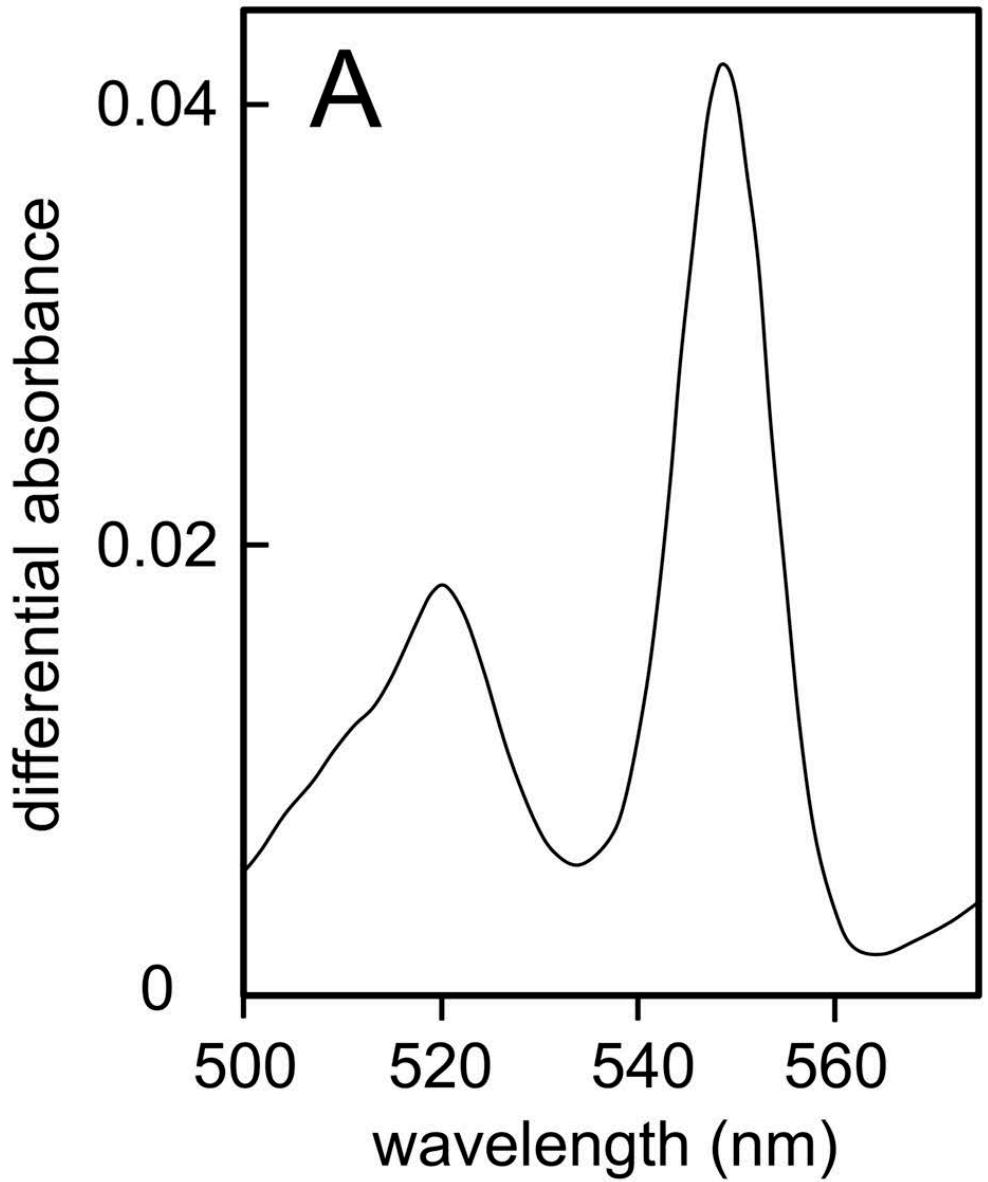
Figure 1. (A) Reduced *minus* oxidized (dithionite *minus* ascorbate) differential absorption spectra of Cc_{550} from *Phaeodactylum tricornutum* extracted by solubilization from membrane fraction. The concentration of cytochrome was 3 μM. (B) Molecular weight MS-analysis of the complete holo-cytochrome. See the Materials and Methods section for more details.

Figure 2. (A) Backbone model of Cc_{550} from *Chaetoceros muelleri* obtained using the crystal structures of Cc_{550} from the cyanobacterium *Synechocystis* sp. PCC 6803 (PDB entry 1E29) and the red alga *Cyanidium caldarium* (PDB 4YUU) as main templates. (B-D) Surface electrostatic potential distribution of the structural models of Cc_{550} from (B) *Chaetoceros muelleri*, (C) *Nannochloropsis gaditana*, and (D) *Isochrysis galbana*. The view displays the heme groups in the same orientation as in (A), showing in front the cofactor exposed area, and in the top the protein C-terminal hydrophobic protuberance. Simulations of surface electrostatic potential distribution were performed using the Swiss-PDB Viewer Program assuming an ionic strength of 500 mM at pH 7.0. Positively and negatively charged regions are depicted in blue and red, respectively.

Figure 3. Potential Cc_{550} interacting proteins as identified by co-immunoseparation and LC-MS/MS in soluble, membrane and crosslinked fractions of the diatom *Phaeodactylum tricornutum*. (*Left*) Venn diagram of the Cc_{550}-protein interaction dataset showing intersections of the three co-immunoseparated fractions. (*Right*) Chloroplastic protein targets selected as appearing both in the crosslinked fraction and at least in one of the other two (soluble or membrane) fractions. [a)] SF, co-immunoseparated in the soluble fraction; MF, co-immunoseparated in the membrane fraction; CF, co-immunoseparated in the crosslinked fraction. [b)] Cc_{550} was identified in all the samples. See text for more details.

Figure 4. (*Upper*) Immunoelectron microscopy images of *Phaeodactylum tricornutum* cells showing localization of (A) photosynthetic Cc_{550}, and (B) the soluble luminal Cc_6. (*Lower*) Expansion of the pyrenoid area. (Cp) chloroplast, (Py) pyrenoid.

Fig 1


A
0.04
0.02
0
differential absorbance
500
520
540
560
wavelength (nm)

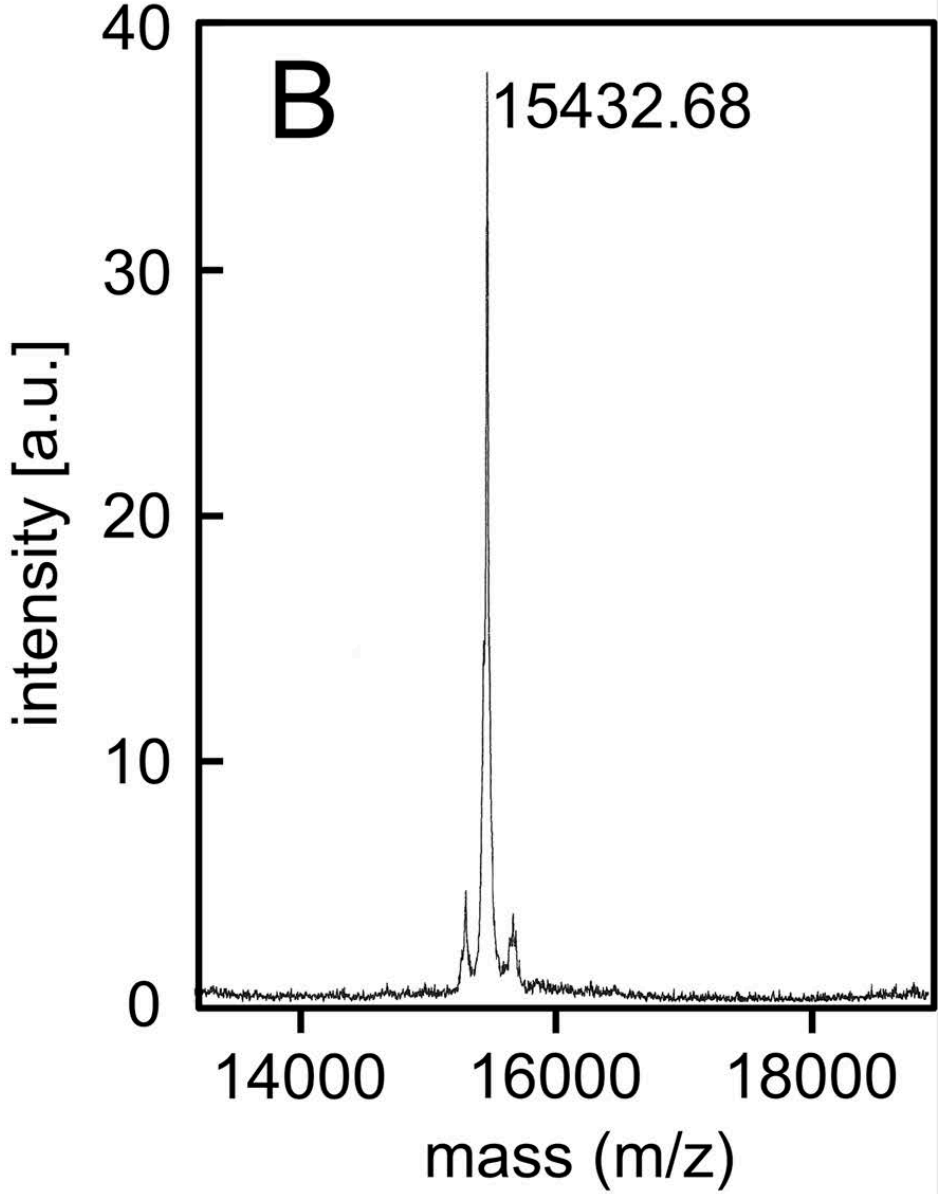

B
15432.68
40
30
20
10
0
intensity [a.u.]
14000
16000
18000
mass (m/z)

Figure 2

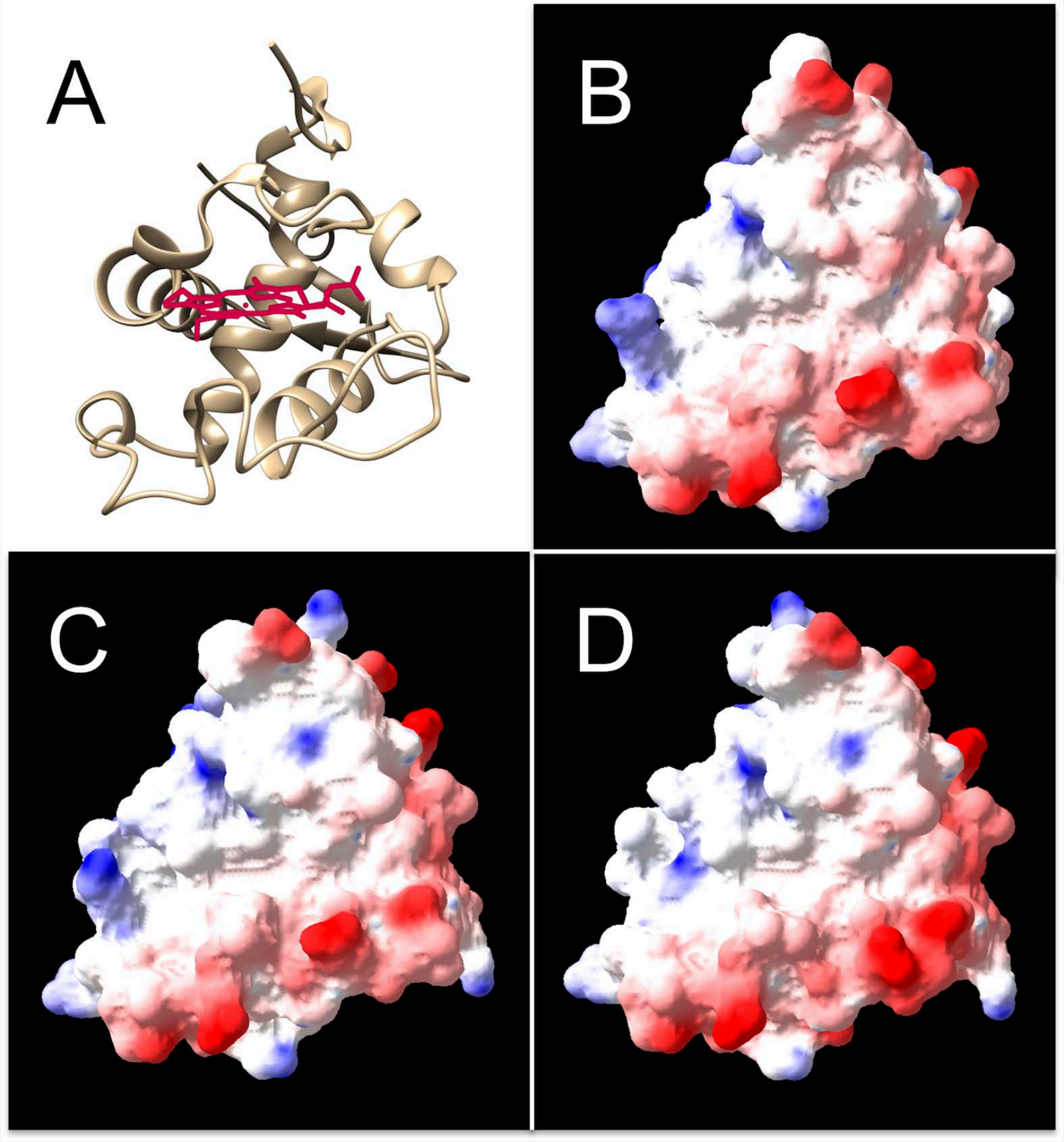
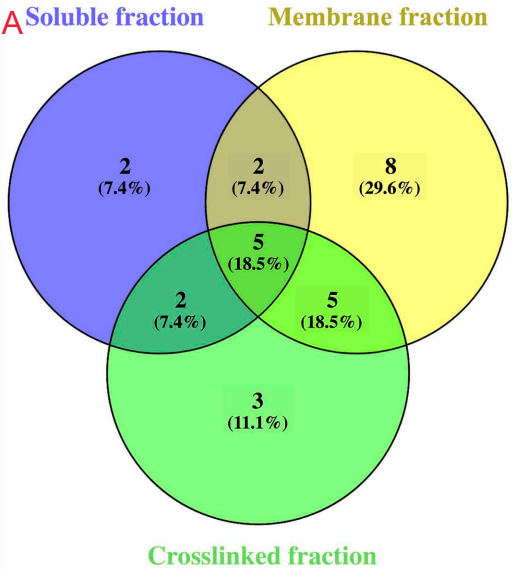


Figure 3

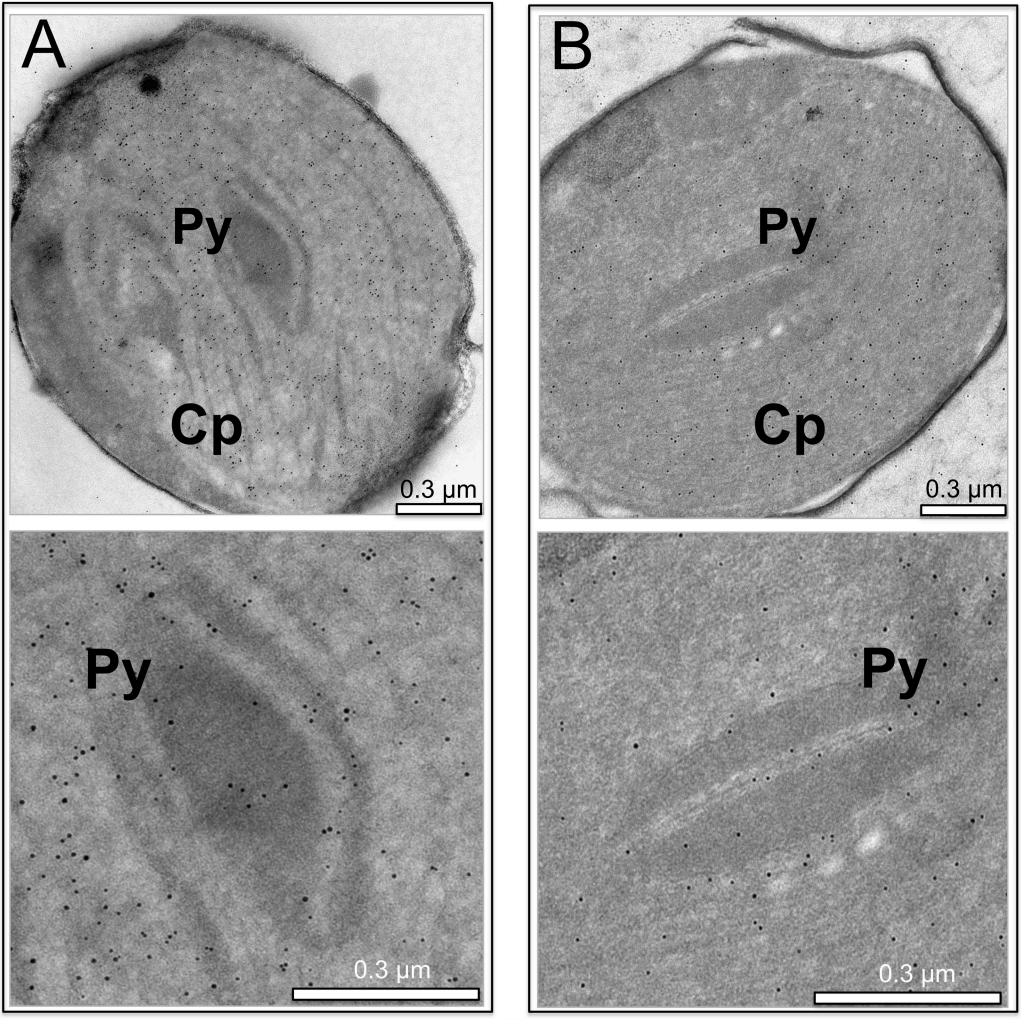
A



B

Protein	Fraction[a)]
Cytochrome c_{550}[b)]	SF, MF, CF
RubisCO large chain, pyrenoid	SF, MF, CF
RubisCO small chain, pyrenoid	SF, MF, CF
CreA-like protein, chloroplast	SF, MF, CF
PtCA1-carbonic anhydrase, pyrenoid	SF, MF, CF
Psb31-photosystem II	SF, CF
FbaC5-Fructose-bisphosphate aldolase, pyrenoid	SF, CF
ATP-synthase subunit α, chloroplast	MF, CF
CP43-photosystem II	MF, CF
D1-photosystem II	MF, CF
D2-photosystem II	MF, CF
PsaD-photosystem I	MF, CF

Figure 4



Amplified *psbV* genes	Oligo names[a)]	Sequence
Chaetoceros muelleri	C1-F	ATGTTAAAAAAATATTCAAAATTTTGTGCG
	C2-R	TTAGTAATAAATTTTTCCACCACCCCAT
Isochrysis galbana (3'-end)	I1-F	CCWGAAGCWTTAAGTTTAGC
	I2-R	GAAAAATGGGGTGGTGGAAAAATTTA
Isochrysis galbana (5'-end)	I3-F	ATGGCACTAAAAAGTAAATTTTTAGTCG
	I4-R	TTATCACGGGGCGGGGTTGC

```
Isochrysis      LELDEDTRTVTLDPKGNTVVLSVEQIKRGKRLFNNACAICHVGGLTKTNPNVGLDVEALS
Emiliania       LELDEDTRTVTLDGKGNTVVLSVEQIKRGKRLFNNACAICHVGGLTKTNPNVGLDVEALS
                ********** **************************************************

Isochrysis      LATPPRDNIANLVAYLKDPMTYDGADSIAELHPSIKSADIFPKMRSLTDEDLFAISGHIL
Emiliania       LATPPRDNVSSLVSYLKDPMTYDGADSIAELHPSIKSADIFPKMRSLTDEDLFAISGHIL
                ********:::.**:**************************************************

Isochrysis      VQPKVVNEKWGGGKIYY
Emiliania       VQPKVVNEKWGGGKIYY
                *****************
```

Figure S1.

(*Upper*) Sequences of oligonucleotides used in this work (5´to 3´). [a)] F, forward primer; R, reverse primer.

(*Lower*) Alignment of sequences of cytochromes c_{550} from the Isochrysidal algae *Emiliania huxleyi* (NCBI accession YP_277399.1) and *Isochrysis galbana*, as deduced from the *psbV* gene sequences

Names	Confidence	Sequence
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	IDLDEATRTVVVDSSGK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	IDLDEATRTVVVDSSGKTIVLTPEQVK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	IDLDEATRTVVVDSSGKTIVLTPEQVKR
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	IVTEKWGGGKIYY ←
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	LFNATCGACHVGGVTK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	MRSVTDEDLTAMAGHILLQPK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	NPTTYDGLESIAEVHPSIK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	RDNIAGLVDFLK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	RLFNATCGACHVGGVTK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	SVTDEDLTAMAGHILLQPK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TDEDLTAMAGHILLQPK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TIVLTPEQVK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TIVLTPEQVKR
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TIVLTPEQVKRGK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TNPNVGLDPEALSLATPR
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TNPNVGLDPEALSLATPRR
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TVVVDSSGKTIVLTPEQVK
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TVVVDSSGKTIVLTPEQVKR
Cytochrome c-550 OS=Phaeodactylum tricornutum	98,52	SADIYPR
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	VTEKWGGGKIYY ←
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	PEALSLATPR
Cytochrome c-550 OS=Phaeodactylum tricornutum	76,45	SADIYPRMR
Cytochrome c-550 OS=Phaeodactylum tricornutum	99,00	TEKWGGGKIYY ←

Figure S2. Summary for the in-gel peptide fingerprint analysis of *P. tricornutum* Cc_{550} obtained from crude cell extracts resolved on polyacrylamide gel electrophoresis. Only the peptides that have contributed to the protein identification are shown, which cover 100 % of the cytochrome mature protein. Arrows indicate peptides fitting for a complete (no truncated) C-terminal end. Peptides corresponding to the truncated protein could not be correctly fitted to the obtained data.

Soluble fraction	Membrane fraction	Crosslinked fraction
Cytochrome c_{550}	Cytochrome c_{550}	Cytochrome c_{550}
RubisCO large chain	**RubisCO large chain**	**RubisCO large chain**
RubisCO small subunit	**RubisCO small subunit**	**RubisCO small subunit**
	RubisCO expression protein	
CreA-like protein	**CreA-like protein**	**CreA-like protein**
	ATP synthase subunit α	**ATP synthase subunit α**
	ATP synthase subunit β	
	ATP synthase subunit λ	
PtCA1-carbonic anhydrase	**PtCA1-carbonic anhydrase**	**PtCA1-carbonic anhydrase**
	ptCA2-carbonic anhydrase	
PsbQ'-photosystem II	PsbQ'-photosystem II	
Psb31-photosystem II		**Psb31-photosystem II**
PsbO-photosystem II		
	CP43-photosystem II	**CP43-photosystem II**
		CP47-photosystem II
	D1-photosystem II	**D1-photosystem II**
	D2-photosystem II	**D2-photosystem II**
	PsaD-photosystem I	**PsaD-photosystem I**
	PsaA-photosystem I	
	PsaB-photosystem I	
	PsaC-photosystem I	
		PsaL-Photosystem I
		PsaF-photosystem I subunit
FBAC5-Fructose-bisphosphate aldolase		**FBAC5-Fructose-bisphosphate aldolase**
FKBP Peptidylprolyl isomerase	FKBP Peptidylprolyl isomerase	
	Phosphoenolpyruvate/ phosphate translocator	
Phosphoglycerate kinase		

Figure S3. Cc_{550} co-immunoseparated proteins identified by LC-MS/MS in soluble, membrane and crosslinked fractions of the diatom alga *Phaeodactylum tricornutum*. Raw LC-MS/MS data were filtered according the following two criteria: i) possible protein targets have to be localized in the chloroplast; and ii) undefined predicted proteins, ribosomal proteins and fucoxanthin-chlorophyll light-harvesting antenna proteins were discarded. Validated protein targets shown in Figure 3 in the text, as appearing both in the crosslinked fraction and at least in one of the other two (soluble or membrane) fractions, are marked in bold. Cc_{550} appeared in all the fractions.

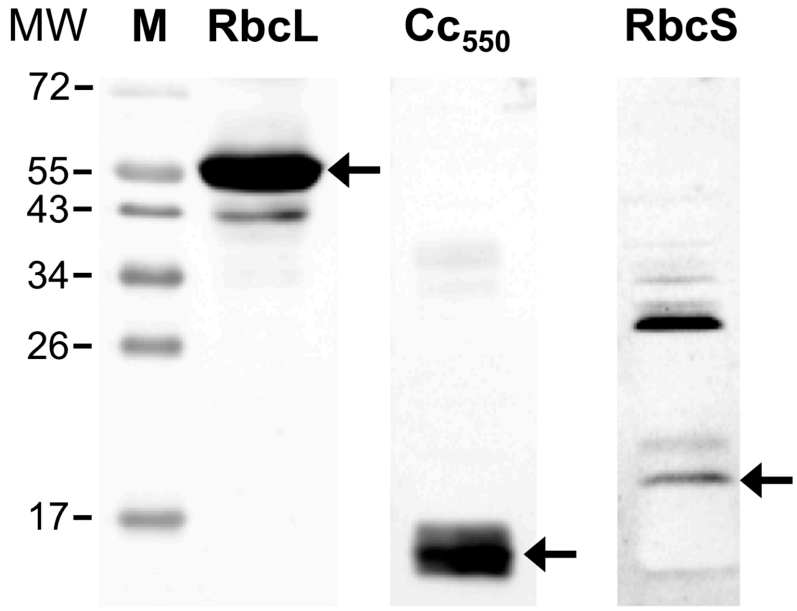


Figure S4.

Western blot analysis of crude cellular fractions of *P. tricornutum* with antibodies against either Cc_{550} or the RubisCO large (RbcL) and small (RbcS) subunits, as indicated. Arrows point to bands corresponding to the proteins against which the specific antibodies were used. The total amount of protein loaded and the antibody dilution was: 10 μg and 1:50000 (Cc_{550}), 50 μg and 1:1000 (RbcL), and 100 μg and 1:1300 (RbcS). M, molecular weight standard.