



# Multiple Profile Models Extract Features from Protein Sequence Data and Resolve Functional Diversity of Very Different Protein Families

Vasco Giovagnetti, Marianne Jaubert, Mahendra Shukla, Petra Ungerer, Jean-Pierre Bouly, Angela Falciatore, Alexander Ruban

## ► To cite this version:

Vasco Giovagnetti, Marianne Jaubert, Mahendra Shukla, Petra Ungerer, Jean-Pierre Bouly, et al.. Multiple Profile Models Extract Features from Protein Sequence Data and Resolve Functional Diversity of Very Different Protein Families. *Molecular Biology and Evolution*, 2022, 39 (4), pp.509-525. 10.1093/molbev/msac070 . hal-03850483

**HAL Id: hal-03850483**

**<https://hal.science/hal-03850483>**







Submitted on 13 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Biochemical and molecular properties of LHCX1, the essential regulator of dynamic photoprotection in diatoms

Vasco Giovagnetti <sup>1</sup>, Marianne Jaubert <sup>2</sup>, Mahendra K. Shukla <sup>1</sup>, Petra Ungerer,<sup>1</sup>  
Jean-Pierre Bouly <sup>2</sup>, Angela Falciatore <sup>2</sup> and Alexander V. Ruban <sup>1,\*†</sup>

<sup>1</sup> School of Biological and Chemical Sciences, Queen Mary University of London, London E1 4NS, UK

<sup>2</sup> Laboratoire de Biologie du Chloroplaste et Perception de la Lumière Chez les Micro-algues, UMR7141, CNRS, Sorbonne Université, Institut de Biologie Physico-Chimique, Paris 75005, France

\*Author for communication: a.ruban@qmul.ac.uk

†Senior author

V.G. and A.V.R. conceived and designed the study; V.G., M.J., M.K.S., P.U., and J.P.B. performed experiments; V.G., M.J., M.K.S., P.U., J.P.B., A.F., and A.V.R. analyzed the data. V.G. and A.V.R. wrote the manuscript; all authors discussed and commented on the results, as well as revised the manuscript. A.V.R. agrees to serve as the author responsible for contact and ensures communication.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (<https://academic.oup.com/plphys/pages/General-Instructions>) is Alexander V. Ruban (a.ruban@qmul.ac.uk).

## Abstract

Light harvesting is regulated by a process triggered by the acidification of the thylakoid lumen, known as nonphotochemical “energy-dependent quenching” (qE). In diatoms, qE is controlled by the light-harvesting complex (LHC) protein LHCX1, while the LHC stress-related (LHCSR) and photosystem II subunit S proteins are essential for green algae and plants, respectively. Here, we report a biochemical and molecular characterization of LHCX1 to investigate its role in qE. We found that, when grown under intermittent light, *Phaeodactylum tricornutum* forms very large qE, due to LHCX1 constitutive upregulation. This “super qE” is abolished in LHCX1 knockout mutants. Biochemical and spectroscopic analyses of LHCX1 reveal that this protein might differ in the character of binding pigments relative to the major pool of light-harvesting antenna proteins. The possibility of transient pigment binding or not binding pigments at all is discussed. Targeted mutagenesis of putative protonatable residues (D95 and E205) in transgenic *P. tricornutum* lines does not alter qE capacity, showing that they are not involved in sensing lumen pH, differently from residues conserved in LHCSR3. Our results suggest functional divergence between LHCX1 and LHCSR3 in qE modulation. We propose that LHCX1 evolved independently to facilitate dynamic tracking of light fluctuations in turbulent waters. The evolution of LHCX(-like) proteins in organisms with secondary red plastids, such as diatoms, might have conferred a selective advantage in the control of dynamic photoprotection, ultimately resulting in their ecological success.

## Introduction

Oxygenic photosynthesis has sustained life on Earth. Around 1.5 billion years ago, an ancestor of extant cyanobacteria was engulfed by a heterotrophic unicellular eukaryote, resulting in a successful endosymbiosis that evolved into a novel and

vertically inherited eukaryotic organelle, the plastid (Falkowski et al., 2004; Keeling, 2013). Contemporary oceans house a multitude of photosynthetic microbes that are responsible for approximately half of planetary carbon fixation

(Field et al., 1998; Falkowski et al., 2004). Among the eukaryotic clades, diatoms often dominate algal communities in well-mixed coastal and upwelling waters (Falkowski et al., 2004; Armbrust, 2009; Falcatore et al., 2020). While green algae and land plants inherited their “green plastids” through primary endosymbiosis, diatoms harbor a complex, “secondary red plastid” that originated via a secondary endosymbiotic event (Armbrust, 2009; Keeling, 2013). Their genome retains an unusual mix of genes (Armbrust, 2009; Falcatore et al., 2020) often argued to be associated with their remarkable capacity to cope with changing abiotic factors, such as light (Armbrust, 2009; Bailleul et al., 2010; Giovagnetti and Ruban, 2017; Falcatore et al., 2020). Indeed, the intrinsic large variability of terrestrial and aquatic light environments requires fine adjustment of light harvesting to avoid photodamage of the delicate photosynthetic machinery (Ruban et al., 2012; Giovagnetti and Ruban, 2018). Similar to other photosynthetic organisms, diatoms achieve such flexibility in light-harvesting regulation mainly through the energy-dependent quenching (qE) of nonphotochemical quenching (Bailleul et al., 2010; Giovagnetti and Ruban, 2017, 2018). qE is feedback-modulated by specialized members of the light-harvesting complex (LHC) protein superfamily and the activation of the xanthophyll cycle (Ruban et al., 2012).

Different functions in qE have been put forward for the stress-related members of the LHC protein superfamily identified so far (Giovagnetti and Ruban, 2018). Plants require the constitutive accumulation of the photosystem (PS) II subunit S (PsbS) to efficiently induce qE (Li et al., 2000). PsbS is a unique four-helix, integral membrane protein (Li et al., 2000; Fan et al., 2015) that has been shown not to bind pigments or not as tightly as other Lhcb members (Dominici et al., 2002; Bonente et al., 2008; Wilk et al., 2013). Able of sensing lumen pH (Li et al., 2004), PsbS appears to promote conformational changes by interacting with PSII chlorophyll *a/b*- and carotenoid-binding LHCs (Correa-Galvis et al., 2016a; Sacharz et al., 2017). Nonetheless, the molecular mechanism of PsbS action on PSII LHCs remains unknown (Sacharz et al., 2017; Giovagnetti and Ruban, 2018).

LHC stress-related (LHCSR) and LHCX proteins are predicted to comprise three membrane-spanning  $\alpha$ -helices and short amphipathic helices exposed to the thylakoid lumen (Peers et al., 2009; Bailleul et al., 2010; Bonente et al., 2011; Ballottari et al., 2016; Taddei et al., 2016). In the green algal model, *Chlamydomonas reinhardtii*, light-induced LHCSR proteins mediate qE (Peers et al., 2009; Bonente et al., 2011; Ballottari et al., 2016; Dinc et al., 2016; Kosuge et al., 2018) together with rapid and transient accumulation of PsbS under high light, and expression in response to UV-B exposure or CO<sub>2</sub> availability (Allorent et al., 2016; Tibiletti et al., 2016; Correa-Galvis et al., 2016b). Moreover, both types of proteins are induced by different wavelengths of light (Allorent et al., 2016; Petroutsos et al., 2016). LHCSR3, the main regulator of qE in *C. reinhardtii* (Peers et al., 2009), was found to bind chlorophyll *a/b* and carotenoid pigments, and to be a quencher of chlorophyll excited states upon protonation of lumen-exposed acidic

residues and interaction with other Lhcb antenna proteins, such as Lhcbm1 (Elrad et al., 2002; Bonente et al., 2011; Ballottari et al., 2016). LHCSR proteins thus seem to be peculiar in merging both pH-sensing and quenching capacities (Bonente et al., 2011; Dinc et al., 2016; de la Cruz Valbuena et al., 2019), although PsbS also appears to contribute to qE activation of *C. reinhardtii* (Allorent et al., 2016; Tibiletti et al., 2016; Correa-Galvis et al., 2016b; Nawrocki et al., 2020).

Diatoms lack *PsbS* genes. However, their quenching formation depends on the accumulation of LHCX proteins (Bailleul et al., 2010; Lepetit et al., 2013, 2017; Taddei et al., 2016, 2018; Buck et al., 2019). LHCX-like proteins are present in algae containing red secondary plastids, such as diatoms, but absent in red algae and cryptophytes (Bailleul et al., 2010; Giovagnetti and Ruban, 2018). Four LHCX genes (LHCX1–4) have been identified in the genome of the pennate diatom *Phaeodactylum tricornutum* (Bailleul et al., 2010; Lepetit et al., 2013, 2017; Taddei et al., 2016). The LHCX1 isoform was found to be constitutively expressed in *P. tricornutum* and to play a key role in qE (Bailleul et al., 2010), while other LHCX isoforms accumulate in response to prolonged light stress, dynamic light changes, nutrient starvation, and even darkness (Lepetit et al., 2013, 2017; Taddei et al., 2016). Diatoms possess a unique type of light-harvesting antenna formed by fucoxanthin chlorophyll *a/c* binding protein (FCP) complexes, which bind fucoxanthin as main carotenoid and chlorophylls *a* and *c* to a high carotenoid/chlorophyll ratio (Wang et al., 2019; Büchel, 2020). The FCP basic unit of *P. tricornutum* has been recently purified as a homodimer (Wang et al., 2019), while the presence of PSII-FCPII and PSI-FCPI supercomplexes has been only confirmed recently in the centric diatoms *Chaetoceros gracilis* (Pi et al., 2019; Xu et al., 2020) and *Thalassiosira pseudonana* (Calvaruso et al., 2020). FCP complexes comprise the major light-harvesting antenna LHCF proteins (PSI-related) LHCR and LHCZ proteins, together with stress-induced LHCX isoforms (Büchel, 2020). Currently, LHCX1 role in the qE mechanism is unknown. No direct evidence that LHCX1 can bind pigments, similarly to LHCSR proteins, is currently available (Giovagnetti and Ruban, 2018). Here, we tested if LHCX1 shows such property by using *P. tricornutum* cultures acclimated to continuous and intermittent light (IL) conditions, the latter known to induce large levels of qE in this species (Lavaud et al., 2002; Ruban et al., 2004; Giovagnetti and Ruban, 2017). Then we explored the pH-sensing ability of LHCX1 by addressing the function of specific, putative protonatable residues conserved with LHCSR3 (Ballottari et al., 2016; Taddei et al., 2016). Functional and evolutionary implications for qE modulation in diatoms are discussed.

## Results

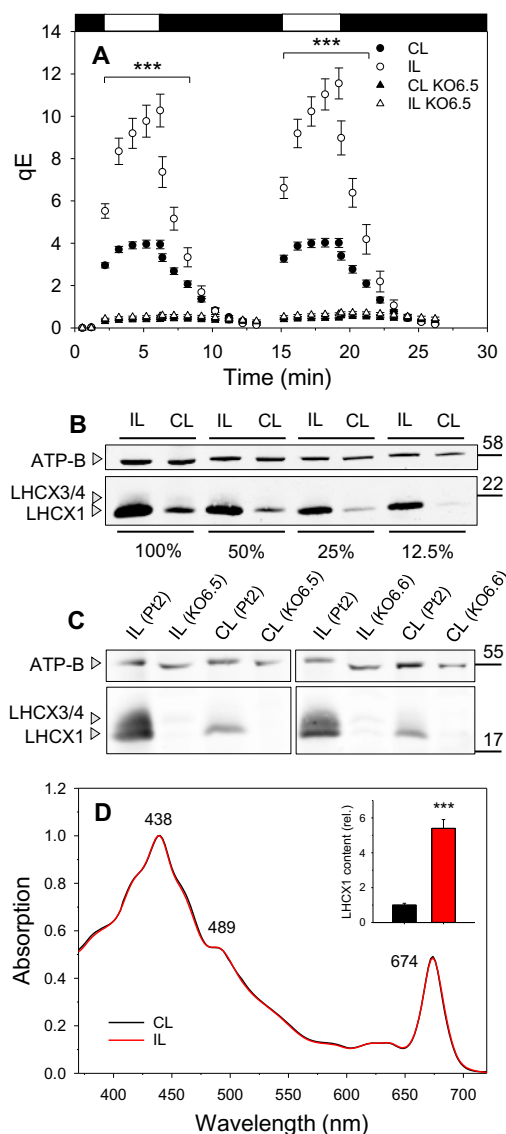
### Constitutive upregulation of LHCX1 protein in wild-type *P. tricornutum* confers large, dynamic photo-protection (qE)

Constitutive accumulation of LHCX1 (Bailleul et al., 2010), together with a very low content of a second isoform (most

likely LHCX3; Taddei et al., 2016), has been reported in *P. tricornutum*, where LHCX1 levels and qE were also observed to be associated when the de-epoxidation state of xanthophyll cycle remained unchanged (Bailleul et al., 2010). Variable upregulation of the four LHCX isoforms is instead induced by different abiotic stress conditions, for example, prolonged high light and dynamic light exposure (Lepetit et al., 2013, 2017; Taddei et al., 2016).

In agreement with previous reports (e.g. Lavaud et al., 2002; Giovagnetti and Ruban, 2017), we found that *P. tricornutum* cells acclimated to IL were able to form very high levels of qE (“super qE”) when exposed to high actinic light, which are significantly larger than those measured in cells grown under continuous light (CL;  $P < 0.001$ ,  $t$  test; Figure 1A). Moreover, qE enhancement in IL cells was accompanied by an unusual rapid recovery of quenching in the dark, which was comparable to that found in CL cells (after 3 min of darkness, no significant difference was observed;  $P > 0.05$ ,  $t$  test) (Figure 1A), owing to analogous activation of the xanthophyll cycle in both light conditions (~35%; Supplemental Table S1; Giovagnetti and Ruban, 2017). Moreover, no significant difference was found in qE amplitude and extent of its recovery ( $P > 0.05$ ,  $t$  test), as well as in qE kinetics, between the first and second cycle of illumination (Figure 1A). Overall, these results support that very large qE levels can be formed in *P. tricornutum* without enhancement of xanthophyll cycle (Giovagnetti and Ruban, 2017).

Here, we found that this unusually large qE is related to a strong, constitutive upregulation of LHCX1 expression and consequent LHCX1 protein accumulation in response to IL growth (Figure 1B). LHCX1 protein identity and upregulation under IL were confirmed by tandem liquid-chromatography mass-spectrometry (LC-MS/MS) analysis (Supplemental Figure S1 and Supplemental Table S2). The causation between the observed over-accumulation of LHCX1 and large extent of quenching was proven by using LHCX1 knockout (KO) cell lines of *P. tricornutum* (Pt2) (KO6.5 and KO6.6; Supplemental Figure S2, A and B), which showed no LHCX1 accumulation (Figure 1C; Supplemental Figure S2C). Quenching was decreased very strongly, being almost abolished, and the observed qE reduction was comparable in both KO lines tested (Figure 1A; Supplemental Figures S3 and S4), with only minor—despite significant ( $P < 0.05$ ,  $t$  test)—differences in qE levels between the two light acclimations investigated (see Supplemental Figure S4). Relevantly, cell physiological state and growth capacity of KO mutants were optimal and similar to those of wild-type (WT) cells (Supplemental Figure S4). Complementation of LHCX1 KO 6.6 line with a genomic LHCX1 sequence under the control of the endogenous LHCX1 promoter and terminator was carried out (Supplemental Figure S5). LHCX1 complementation successfully restored qE function to comparable or even larger extent of WT cells, as previously shown in another *P. tricornutum* ecotype (Buck et al., 2019). The variable rescue of qE corresponded to the different



**Figure 1** LHCX1 modulates the extent of qE in *P. tricornutum* cells acclimated to CL and IL. A, Chlorophyll fluorescence quenching induction (actinic light intensity,  $1,650 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ) and relaxation in the dark of WT cells grown either under CL (closed circles; seven biological replicates,  $n = 24 \pm \text{SE}$ ) or IL (open circles; eight biological replicates,  $n = 32 \pm \text{SE}$ ; \*\*\* $P < 0.001$ ,  $t$  test). LHCX1 KO cell line (KO 6.5) acclimated to CL (closed triangles; five biological replicates,  $n = 8 \pm \text{SE}$ ) and IL (open triangles; six biological replicates,  $n = 9 \pm \text{SE}$ ) are shown. B, Western blot analysis of LHCX proteins accumulated in TMs isolated from dark-adapted CL- and IL-grown cells. Same total protein amounts were loaded for CL and IL conditions (100–12.5% range corresponds to 10–1.25  $\mu\text{g}$ ). ATP-B levels are shown as loading control. C, Western blot analysis of TMs isolated from dark-adapted CL- and IL-grown LHCX1 KO lines (KO6.5 and KO6.6). WT samples (Pt2) are shown as controls. Same total protein amounts (5  $\mu\text{g}$ ) were loaded for each condition. D, Absorption spectra of TMs isolated from dark-adapted CL- and IL-grown cells (averages of four-six independent biological samples). The inset shows the upregulation of ATP-B-normalized LHCX1 content due to IL acclimation (averages of six independent biological samples,  $n = 22 \pm \text{SE}$ ; \*\*\* $P < 0.001$ ,  $t$  test).



*LHCX1* expression levels found in independent complemented lines, due to random and possibly multiple copy insertion of the complementation constructs (Supplemental Figure S5).

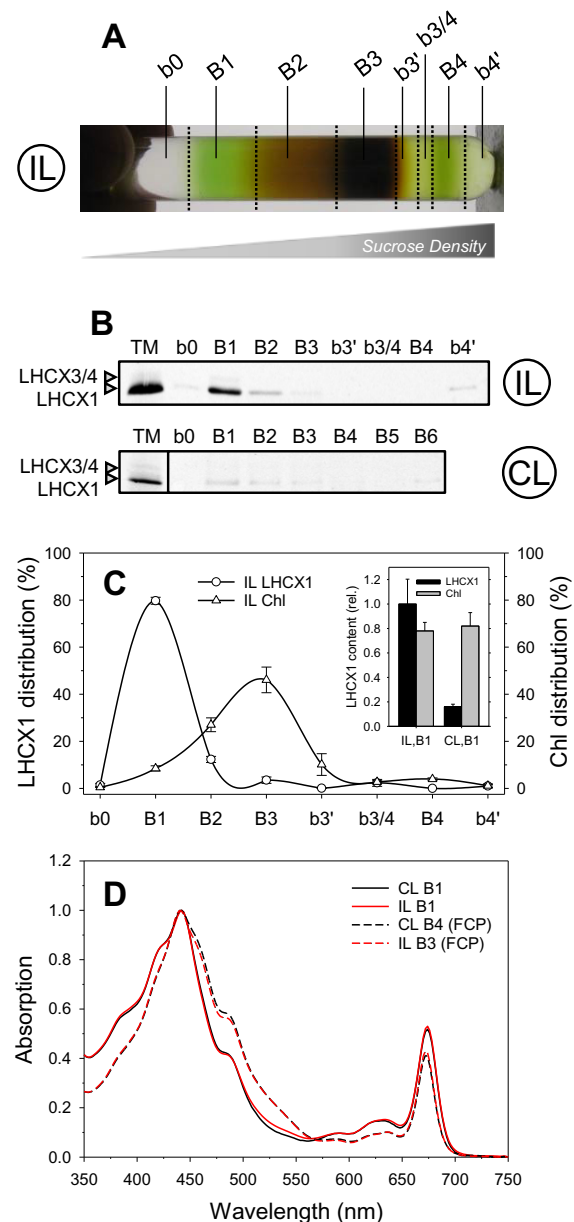
Notably, we observed no alteration in the absorption spectra of thylakoid membranes (TMs) isolated from dark-adapted cells grown under both conditions (Figure 1D), although *LHCX1* accumulation in IL TMs was more than five-fold greater than in CL membranes ( $P < 0.001$ ,  $t$  test; inset in Figure 1D). Densitometric analysis of InstantBlue-stained SDS–PAGE protein bands suggested that *LHCX1* contributed to  $7.2 \pm 1.2\%$  of pigment-binding FCP proteins (10–30 kDa range) and  $\sim 3\%$  of total thylakoid proteins in IL thylakoids (see Supplemental Figure S1). Although measuring the densitometry of sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS–PAGE) protein bands is not a quantitative analysis, the increase in *LHCX1* accumulation found under IL did not induce absorption changes in thylakoids, consistently with previous *in vivo* absorption measurements on *P. tricornutum* cells (Giovagnetti and Ruban, 2017).

Overall, these data suggest that *LHCX1* can undergo a strong upregulation *in vivo* and its higher basal levels under IL relative to CL acclimation are almost exclusively responsible for increasing qE capacity and flexibility, without requiring an enhanced xanthophyll cycle operation (Supplemental Table S1).

### Isolation of *LHCX1* protein: biochemical and spectroscopic analyses reveal its different capacity of binding pigments relative to the major FCP antenna

To isolate *LHCX1* and elucidate if such protein is able to bind pigments, we used IL-acclimated *P. tricornutum* cells that accumulate a very large amount of such protein (Figure 1), much greater than that observed in transgenic lines overexpressing the *LHCX1* gene (Bailleul et al., 2010; Taddei et al., 2016). *LHCX1* was isolated by linear sucrose density-gradient ultracentrifugation and its localization among photosynthetic membrane protein complexes was established (Figure 2). TMs were isolated from dark-adapted *P. tricornutum* IL cells and mildly solubilized to a final concentration of 0.6% n-dodecyl- $\beta$ -D-maltoside ( $\beta$ -DDM, w/v; at a final detergent/chlorophyll  $a$  ratio of 13.5) to preserve protein–pigment interactions and avoid formation of free pigment (FP), despite reducing the capacity of separating discrete bands.

Four pigmented bands (B1–4) and pigmented fractions (b3', b3/4, and b4') were separated and collected, together with the unpigmented fraction b0 (Figure 2A). The first green band likely contained monomeric proteins (B1), while brown bands B2 and B3 mostly comprised FCP antenna complexes, previously reported to be trimeric (Lepetit et al., 2007; Gundermann et al., 2013) and only recently found to form stable homodimers of LHCF4 (LHCF3) protein in *P. tricornutum* (Wang et al., 2019; Figure 2, A; Supplemental Figures S6, A and B and S7). Notably, the mild-detergent



**Figure 2** *LHCX1* and chlorophyll distribution in photosynthetic membrane complexes isolated from dark-adapted *P. tricornutum* cells grown under IL. A, Sucrose density-gradient separation of TMs isolated from IL-acclimated cells (bands B1, B2, B3, and B4; fractions b0, b3', b3/4, and b4'). B, Western blot analysis of *LHCX* isoforms found in bands and fractions harvested from sucrose density gradients. IL and CL conditions are compared. Equal sample volumes were loaded onto each lane. About 10  $\mu$ g of TM total protein content was loaded as control. C, *LHCX1* (open circles) and chlorophyll (open triangles) relative distribution (%) among bands and fractions harvested from dark-adapted IL TMs. Data are averages of three independent biological samples ( $n = 4 \pm \text{SE}$ ). The inset compares the normalized, absolute content of *LHCX1* and chlorophylls found in band B1 between CL and IL condition (three independent biological samples,  $n = 3-4 \pm \text{SE}$ ). Absolute content of chlorophylls was retrieved from chlorophyll relative distribution calculated using absorption spectra. D, Absorption spectra of bands B1 (enriched in *LHCX1*), and bands B3 and B4 (corresponding to the FCP antenna complexes), obtained by sucrose density gradient separation of CL and IL TMs (averages of at least three independent biological samples).

solubilization applied to TMs did not induce any release of FP, visible instead as a yellow band upon solubilization to a final concentration of 2%  $\beta$ -DDM (w/v, at a final detergent/chlorophyll *a* ratio of 100; [Supplemental Figure S8](#)). Green bands and fractions denser than B3 were instead enriched in PS II and I core complexes, associated with a minor content of FCP polypeptides ([Figure 2A](#); and [Supplemental Figures S6, A and B and S7](#)). The poor segregation of PSII and PSI is in agreement with previous studies on *P. tricornutum* ([Lepetit et al., 2007, 2010](#)), especially given the low amount of detergent used here. Consistently, when TMs were solubilized using more detergent, a greater separation of membrane complexes was achieved, yielding resolved bands for FCP, PSII, and PSI, as observed earlier ([Taddei et al., 2018](#); [Supplemental Figure S8A](#)).

LHCX1 protein (with very minor amount of LHCX3/4 isoform; [Taddei et al., 2016](#)) was identified by SDS–PAGE and immunoblotting analysis ([Figure 2B](#); [Supplemental Figure S6, A and B](#)), and its relative distribution over the sucrose density gradient were assessed ([Figure 2C](#)). LHCX1 was mainly found in band B1 ( $80 \pm 1.5\%$ ) with minor contribution of band B2 ( $12 \pm 1.3\%$ ; [Figure 2C](#)), and clearly upregulated when compared to CL condition ([Figure 2B](#)), where the highest content of LHCX1 was also measured in correspondence of band B1 ([Figure 2B](#); [Supplemental Figure S6](#)). Under IL condition, LHCX1 content was  $40 \pm 1.3\%$  of antenna pigment-binding polypeptides (17–28 kDa range) and  $17 \pm 1.5\%$  of total polypeptide content found in band B1, as assessed by densitometric analysis of SDS–PAGE protein bands ([Supplemental Figure S6, A and C](#)). This indicates that, although band B1 is not pure, LHCX1 contributes to almost half of the proteins that binds the majority of pigments.

Sucrose density-gradient ultracentrifugation of 2%  $\beta$ -DDM-solubilized TMs resulted in altered distribution of LHCX1 ([Supplemental Figure S8B](#)) relative to that induced by mild solubilization ([Figure 2B](#)). In this condition, LHCX1 was mainly detected in the FCP antenna complexes as well as the adjacent and denser unpigmented fraction (Unpigm. F, possibly due to protein aggregation), but not in FP ([Supplemental Figure S8B](#)). Overall, these data indicate that the enrichment of LHCX1 observed in B1 does not relate to FP formation, hence providing an experimental condition suitable to address the pigment-binding capacity of this protein.

Band B1 possessed low chlorophyll amount ( $8.5 \pm 1.1\%$ ; [Figure 2C](#); [Supplemental Figure S9A](#); [Supplemental Table S3](#)), while the highest chlorophyll content ( $46 \pm 5.5\%$ ) was found by the FCP-containing band B3 ([Figure 2C](#); [Supplemental Figure S9A](#); [Supplemental Table S3](#)), which was the most intact (and likely dimeric) antenna complex obtained in our solubilization condition ([Figure 2C](#)). This is in agreement with very minor amounts of chlorophyll *a* and *c*, fucoxanthin, and diadinoxanthin found in band B1, relative to FCP antenna (B3; [Supplemental Table S3](#)). Notably, we found that the absolute content of chlorophylls in B1

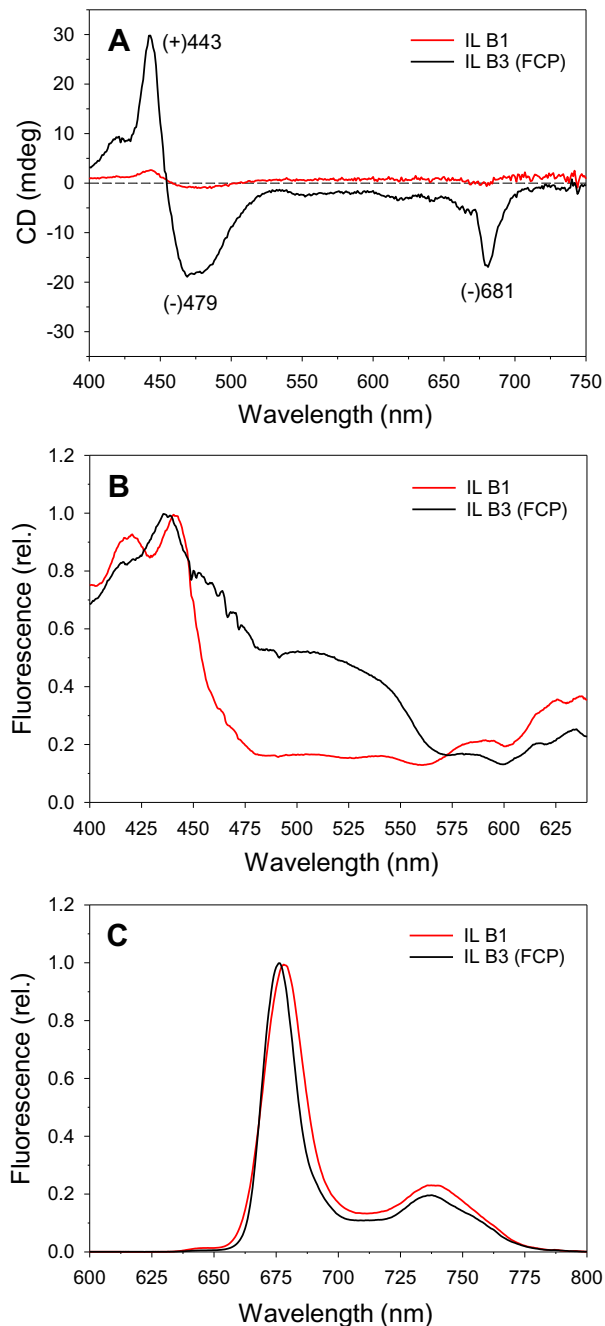
retrieved by absorption spectroscopy was almost identical between CL and IL conditions, despite the latter accumulates around six-fold more LHCX1 protein (see inset in [Figure 2C](#)). Moreover, the amount of proteins, and especially FCP pigment-binding components, found in band B1 is similar between IL and CL samples ([Supplemental Figure S6, A and B](#)). This indicates that LHCX1 over-accumulation under IL (contributing to 40% of antenna proteins in band B1) did not alter the chlorophyll content found in band B1. Indeed, the absorption spectra from bands B1, as well as the FCP-related bands B3 (IL)/B4 (CL), were almost identical between light acclimations ([Figure 2D](#)). This suggests that LHCX1 might be unable of binding pigments, and the absorption capacity of the isolated FCP complexes is comparable between the two light conditions tested.

Because minor amounts of pigments were co-purified with monomers, spectroscopic properties of band B1 were investigated and compared to those of FCP complexes (band B3) to address the presence of pigment–pigment interactions and functional energetic connectivity ([Figure 3](#)). The circular dichroism (CD) spectrum of FCP showed features identical to those reported for FCP isolated from *P. tricornutum*, with excitonic interactions involving chlorophylls and carotenoids in the Soret region (approximately +443 nm/–479 nm) and a single negative signal (approximately –680 nm, originating from chlorophyll *a* molecules) in the  $Q_y$  region ([Lepetit et al., 2007](#); [Szabó et al., 2008](#); [Figure 3A](#)). No optical activity was instead revealed for band B1 ([Figure 3A](#)). Upon detection of excitation fluorescence at 690 nm, functional energetic coupling between chlorophylls and carotenoids was preserved in the FCP antenna ([Figure 3B](#)) and fluorescence emission typical of diatom FCP antenna was measured ( $\sim 676$  nm; [Lavaud et al., 2003](#); [Lepetit et al., 2007](#); [Figure 3C](#)). This suggests that most of the FCP complexes isolated were intact. Conversely, we found that chlorophylls and carotenoids were not energetically coupled in band B1 ([Figure 3B](#)), which instead displayed a slightly red-shifted and nonhomogeneously broadened fluorescence spectrum relative to that of FCP antenna ([Figure 3C](#)). While the former feature is in line with the absence of free chlorophylls in band B1, the latter is indicative of loss of pigment functional organization and monomerization of antenna complexes ([Peterman et al., 1996](#)).

Because the optical activity and energetic connectivity that are typical of an intact FCP antenna were retained in band B3, but absent in band B1, we conclude that LHCX1 might be unlikely to bind pigments or as tightly as the main pool of light-harvesting proteins that constitute the antenna of *P. tricornutum*.

### LHCX1 and pigments do not co-elute during further protein purification

Size exclusion chromatography (SEC) was employed to further purify the LHCX1-enriched band B1, showing the appearance of protein signal (absorption at 280 nm) after



**Figure 3** Spectroscopic analysis of band B1 (enriched in LHCX1) shows lack of optical activity and energetic coupling between LHCX1 and pigments. A, CD representative spectra in the visible region. Bands B1 and B3 (FCP) are compared for the IL condition. Each sample corresponded to 12.5  $\mu$ g of total protein. B, C, 77K fluorescence excitation (B) and emission spectra (C) measured for IL bands B1 and B3. Spectra are averages of three independent biological samples ( $n = 5$ ).

23 min of elution (Figure 4A). While FP was not present during SEC, pigment peaks (mainly due to carotenoids, 480 nm) were detected after 28 min of elution and in correspondence of a protein peak (Figure 4A). However, western blot detection of LHCX1 in fractions eluted (F5–F11, Figure 4B) and

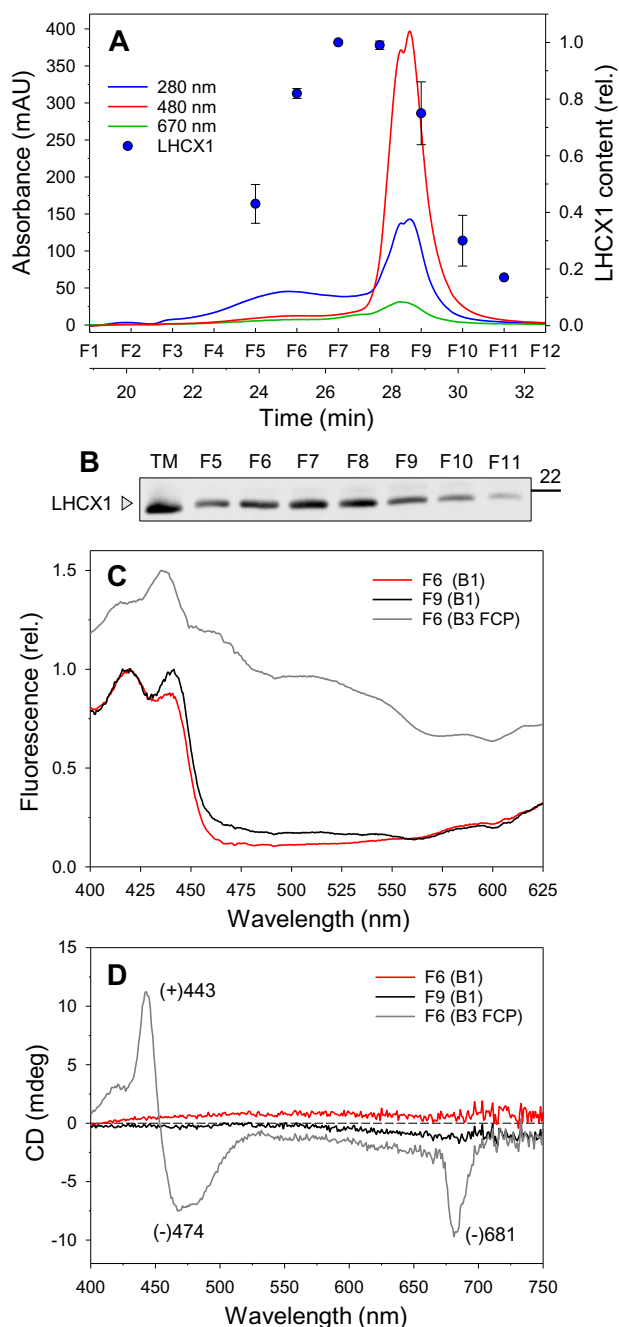
redistribution of its relative content over time (blue dots in Figure 4A) show that the elution of LHCX1 protein was maximal at around 26 min and before the detection of pigments (mainly carotenoids, after 28 min; Figure 4A; Supplemental Figure S9B; Supplemental Table S3). Therefore, the protein signal found in correspondence to the pigment absorption peaks (Figure 4A) reflects only partial contribution of LHCX1, while the remaining protein signal corresponds to other monomeric polypeptides (such as antenna LHCF proteins; see 18–23 kDa range in Supplemental Figure S10). LHCX1 contribution to antenna pigment-binding polypeptides (17–30 kDa) found in SEC fractions (F5–F11) ranged indeed from 32% to 54% (on average  $41 \pm 0.02\%$ ), with  $21 \pm 0.01\%$  average contribution to total polypeptide content (Supplemental Figure S10). These data suggest that the pigments found in band B1 are not carried by LHCX1, but rather by other pigment-binding (antenna) proteins. This is also consistent with finding a similar content of FCP antenna proteins in CL and IL band B1 in response to mild solubilization of thylakoids (Supplemental Figure S6, A and B), despite the different LHCX1 accumulation observed (Figure 2).

SEC of FCP (band B3) was also performed to assess the elution profile of the FCP antenna (Supplemental Figure S11), likely to be dimeric (Wang et al., 2019), and compared to that of monomeric proteins comprised in band B1 (Figure 4A). This revealed that although LHCX1 started to appear as aggregates or dimers (24–26 min), it was eluted later as a monomer (compare Figure 4A to Supplemental Figure S11A). Notably, similar carotenoid/chlorophyll ratios were measured for the FCP complexes either isolated in band B3 or in SEC fractions (F6–F7 from band B3; Supplemental Table S3), suggesting that the SEC procedure did not strip away pigments.

Spectroscopic characterization of fractions F6 (shifted from pigment peak) and F9 (co-eluted with pigments) collected from SEC of band B1 exhibited absence of functional energetic coupling between chlorophylls and carotenoids (Figure 4C), as well as optical activity (Figure 4D), whether pigments were present or not. Differently, the fraction F6 eluted during SEC of FCP (Supplemental Figure S11A) maintained energetic connectivity among pigments (Figure 4C) and pigment–pigment excitonic interactions, displaying similar features to those of FCP antenna (although with reduced CD signal; compare Figure 4, C and D to Figure 3, A and B). Overall, no pigment–LHCX1 (and pigment–antenna protein) functional energetic coupling was found in band B1, in contrast to the proteins that form the FCP antenna (band B3). This indicates that band B1 contained some pigment-binding (antenna) proteins that have dissociated from functional FCP antenna complexes.

### Residues D95 and E205 in LHCX1 protein have no role in pH sensing

The pH-sensing ability of LHCX1 has not been investigated yet. PsbS and LHCSR proteins possess specific residues that



**Figure 4** SEC of band B1 (IL) reveals no comigration between LHCX1 and pigments. **A**, Representative SEC elution profile of band B1 harvested after sucrose density-gradient separation of IL TMs. Protein, carotenoid, and chlorophyll elution was probed by absorbance at 280 (blue), 480 (red), and 670 nm (green), respectively. LHCX1 content was quantified after western blot analysis and its relative distribution overlaid to the elution pattern (blue dots; three independent biological samples  $\pm$ SE). **B**, Western blot analysis of LHCX1 found in fractions eluted by SEC (F5–F11). Equal sample volumes were loaded onto each lane. About 10  $\mu$ g of TM total protein content was loaded as control. **C** and **D**, Representative spectra of fluorescence excitation (**C**) and CD in the visible region (**D**) for fractions F6 and F9 (SEC of band B1), and F6 (SEC of band B3, that is, FCP; see Supplemental Figure S11 for SEC of band B3). Samples corresponding to 12.5  $\mu$ g of total protein were used for CD measurements.

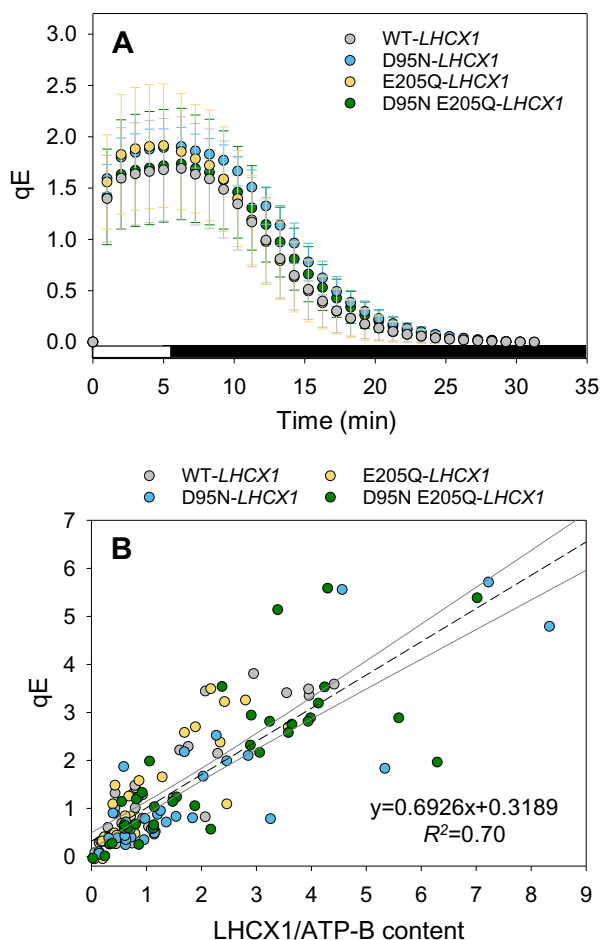
are protonated to sense lumen acidification. While PsbS senses lumen pH via two glutamates (referred as E69 and E173 (Fan et al., 2015), or Glu122 and Glu226 (Li et al., 2004)), residues D117, E221, and E224 have been shown to be crucial in LHCSR3, which also bears a number of possibly protonatable acidic residues located at the C-terminal domain (Liguori et al., 2013; Ballottari et al., 2016). Although five acidic residues are predicted to face the thylakoid lumen in LHCX1, four are conserved with LHCSR3, among which D95 and E205 correspond to the key pH-sensing residues identified in LHCSR3 (i.e. D117 and E224; Ballottari et al., 2016; Taddei et al., 2016; Supplemental Figure S12). For these reasons, the involvement of both amino acids in sensing pH was tested here by complementing *LHCX1* KO lines with *LHCX1* gene sequences where D95 and E205 were mutated into asparagine (D95N) and glutamine (E205Q), respectively (Supplemental Figure S13). Single point mutations (D95N and E205Q) and double point mutation (D95N E205Q) were characterized and compared to the WT-*LHCX1* complementation construct (Figure 5; Supplemental Figure S13).

Despite the variability in *LHCX1* expression found in cell lines of each construct, all *LHCX1* complementation mutants were able to form qE levels comparable to those of WT-*LHCX1* complementation lines (Figure 5; Supplemental Figure S14), with similar kinetics of formation and relaxation (Figure 5A; Supplemental Figure S15). This indicates that qE function is not impaired in the absence of one or both conserved residues. LHCX1 accumulation and qE formation were found to be linearly correlated in the four *LHCX1* complementation constructs (Figure 5B; Supplemental Figure S16), in agreement with LHCX1 essential, constitutive control on qE (Bailleul et al., 2010; Taddei et al., 2016; Buck et al., 2019). The “LHCX1 content: qE levels” relationships observed among constructs were compared and tested for statistical significance (Supplemental Figures S16 and S17). A mixed effects model with random slopes and intercepts found no significant differences in either the slopes (“cell line: LHCX1” interaction term:  $F_{3,109.9} = 0.337$ ,  $P = 0.798$ ) or intercepts of the linear regressions (i.e. once the interaction term was removed from the model:  $F_{3,116.6} = 0.7728$ ,  $P = 0.512$ ). Regardless of *LHCX1* complementation constructs, the model supports the presence of a positive, significant relationship between LHCX1 content and qE formation ( $F_{1,10.8} = 35.42$ ,  $P = 0.0001$ , slope = 0.739, SE = 0.124). These data suggest that D95 and E205 are not proton-binding residues that affect LHCX1 function, despite being conserved relative to LHCSR3 amino acids that are essential to sense pH and trigger qE.

## Discussion

Here we show that unusually large, dynamic qE can be formed in *P. tricornutum* cells acclimated to IL due to strong upregulation of the LHCX1 protein and independently from enhanced xanthophyll cycle activity (Figure 1; Supplemental Table S1; Giovagnetti and Ruban, 2017). Knocking out of *LHCX1* gene results in cell lines essentially lacking qE





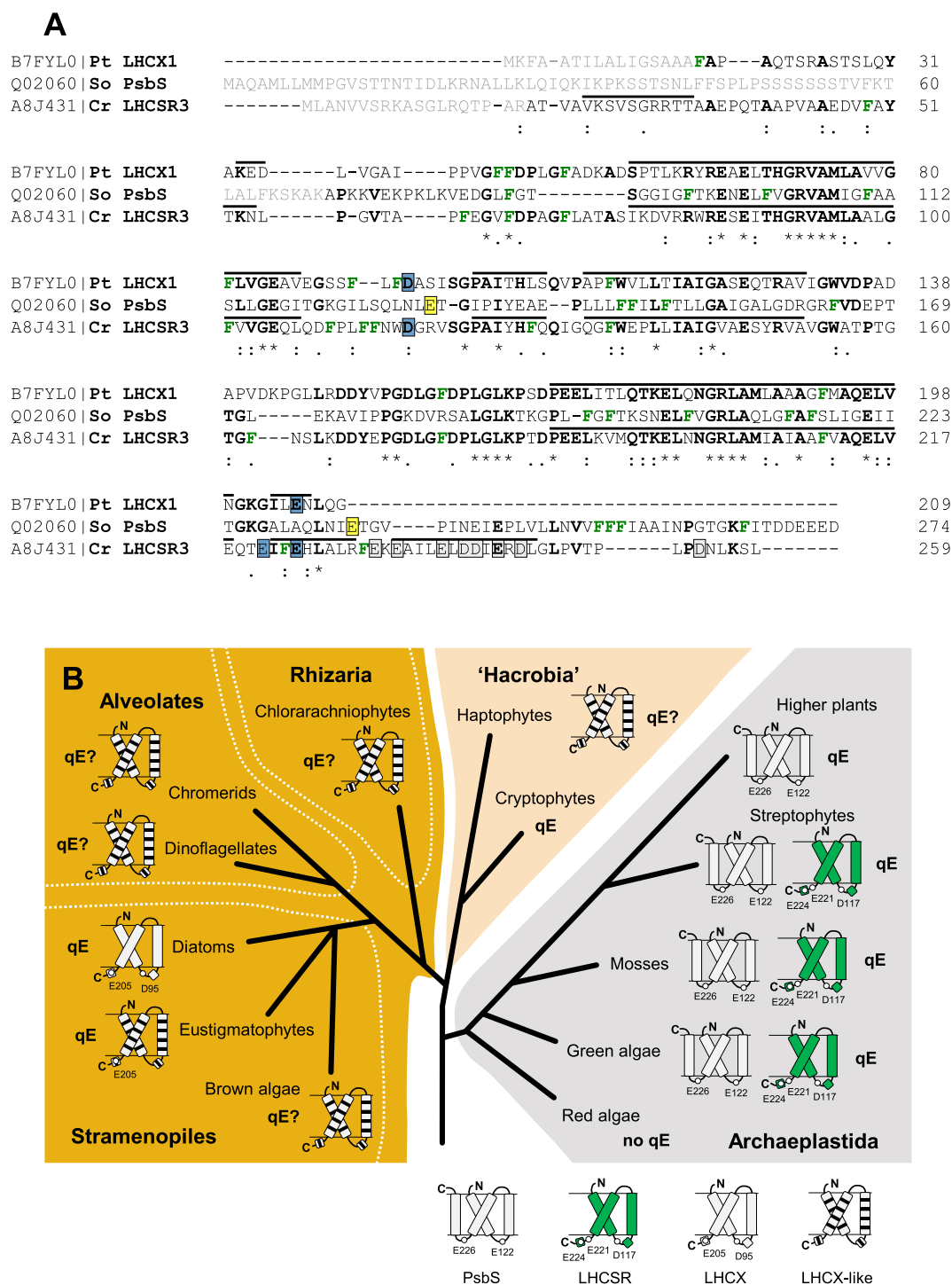
**Figure 5** Residues D95 and E205 in LHCX1 are not involved in pH sensing. A, qE kinetics of formation and relaxation measured in WT- (gray), D95N- (cyan), E205Q- (yellow), and D95N E205Q-LHCX1 (green) complementation constructs. Complementation cell lines were grown under CL ( $10 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ ). qE was induced by a  $600 \mu\text{mol photons m}^{-2} \text{s}^{-1}$  actinic light during 6 min. Averages ( $\pm$ SD) of mutant lines with qE values between 1 and 2.5 are shown (three independent biological replicates; WT ( $n = 20$ ), D95N ( $n = 12$ ), E205Q ( $n = 25$ ), and D95N E205Q ( $n = 25$ )). B, Relationship between LHCX1 accumulation and qE formation (three independent biological replicates,  $n = 30$ – $36$ ). LHCX1 protein content was retrieved by western blot analysis (see Supplemental Figures S13 and S14), and normalized by that of the loading control (ATP-B).

(Figure 1; Supplemental Figures S3 and S4A), supporting the key role of LHCX1 in controlling fast and reversible regulation of light harvesting in diatoms (Bailleul et al., 2010).

Because of such phenotype, we used IL-acclimated *P. tricornutum* cells to elucidate the LHCX1 role in diatom qE by studying its ability to bind pigments, a feature previously reported for the LHCX homologs of green algae, LHCSR3 (Bonente et al., 2011; Dinc et al., 2016; de la Cruz Valbuena et al., 2019). LHCX1 was isolated from IL-acclimated *P. tricornutum* cells that over-accumulate this protein (Figure 1). Mild  $\beta$ -DDM-detergent solubilization of thylakoids, which induced no FP formation, was employed to maintain

pigment–pigment and pigment–protein interactions found in vivo, with a decrease in purity of LHCX1 as a methodological trade-off. This approach was preferred to in vitro reconstitution of recombinant LHCX1, a method that can generate large differences in pigment/protein stoichiometry and pigment site occupancy during in vitro refolding relative to the pigments bound in vivo, as well as artifacts (e.g. Croce et al., 1999, 2002; Aspinall-O’Dea et al., 2002; Dall’Osto et al., 2006). Here, we attempted to isolate LHCX1 via sucrose density gradient separation of solubilized TMs, which yielded no FP and 80% of total LHCX1 content present in a monomeric band (B1; Figure 2), despite showing some contamination with other polypeptides (Supplemental Figure S6). Absorption and pigment content of band B1 are strongly reduced relative to those of isolated FCP complexes (Supplemental Figure S9; Supplemental Table S3), and the larger accumulation of LHCX1 in IL TMs (Figure 1D) and band B1 (Figure 2C, inset) relative to CL acclimation does not induce changes in absorption (Figures 1, D and 2, D). Moreover, we show that LHCX1-containing band B1 or SEC fractions exhibit no energetic connectivity between pigments and optical activity, in contrast to FCP complexes isolated and fractions collected during SEC (Figures 3 and 4; Supplemental Figure S11). Purification by SEC reveals that LHCX1 and pigments do not co-elute (Figure 4A). Instead, other monomers (such as antenna LHCF proteins) found in B1 SEC fractions (compare Figure 4A; Supplemental Figure S10 to S11) are responsible for carrying pigments, which however are not functionally organized to provide energy transfer and optical activity (Figure 4, C and D). Overall, these data indicate that LHCX1 might possess a different capacity or be unable to bind pigments relative to other LHCF (and LHCR) members that harvest light in diatoms (Pi et al., 2019; Wang et al., 2019; Xu et al., 2020). This implies that LHCX1 would not be a quencher in qE, differently to what has been shown for LHCSR3 (Bonente et al., 2011; de la Cruz Valbuena et al., 2019), as well as that LHCX1 and LHCSR proteins would play a different role in qE of diatoms and green algae, respectively.

Then we started to explore if LHCX1 can sense pH changes in the lumen that would trigger its activity. We show that exchanging the acidic candidates D95 and E205 (Taddei et al., 2016) into neutral amino acids does not alter LHCX1-dependent qE (Figure 5; Supplemental Figures S14 and S15), in contrast to the protonatable residues, D117 and E224, that are conserved in LHCSR3 (Ballottari et al., 2016). This indicates that D95 and E205 have no pH-sensing role in LHCX1, and further underlines possible functional divergence between LHCX1 and LHCSR3, despite their primary structure sequence identity (46%; Figure 6A). In line with the redundancy of the proton-sensing residues reported in LHCSR3 (Liguori et al., 2013; Ballottari et al., 2016), LHCX1 might possess other lumen-facing amino acids, not conserved with key LHCSR protonatable residues (e.g. E85, E88, and E196; Figure 6A), capable of binding protons and modulate its activation. Alternatively, LHCX1 response to lumen



**Figure 6** Diversity and evolutionary distribution of the proteins controlling qE in photosynthetic eukaryotes. A, Multiple protein sequence alignment of LHCX1 (*P. tricornutum*), PsbS (*S. oleracea*), and LHCSR3 (*C. reinhardtii*) proteins. Aspartic (D) and glutamic acid (E) residues responsible for pH-sensing in PsbS (Li et al., 2004) and LHCSR3 (Ballottari et al., 2016) are highlighted (yellow and blue boxes, respectively), together with relevant acidic residues in LHCSR3 C-terminus region (gray boxes) (Liguori et al., 2013; Ballottari et al., 2016). Residues D95 and E205 in LHCX1 are highlighted in blue boxes. Transit peptides (gray amino acids), sequence identity (bold letters and asterisks) and degree of similarity (colon, strong similarity; period, weak similarity), hydrophobic phenylalanine residues (F, highlighted in green), and helices predicted for LHCX1 and LHCSR3 (dark bars) are also shown. B, Diagram of qE evolution in photosynthetic eukaryotes in relation to the current knowledge on proteins that regulate dynamic quenching (Giovagnetti and Ruban, 2018). Direct involvement in qE, pH-sensing, and pigment-binding capacities of LHCX-like proteins are currently unknown (e.g. Dittami et al., 2010; Xiang et al., 2015; Chukhutsina et al., 2017; Giovagnetti and Ruban, 2018; Park et al., 2019). LHCX/LHCX-like proteins appear to be absent in red algae and cryptophytes possibly due to ecosystem adaptation (the former are often restricted to rocky shorelines) and different mechanisms of qE (Giovagnetti and Ruban, 2018; Haniewicz et al., 2018).

acidification could be regulated by a mechanism different than those found in LHCSR3 (Liguori et al., 2013; Ballottari et al., 2016) and PsbS (Li et al., 2004), or even absent. Further studies are required to understand LHCX1 activity modulation by pH.

Despite having been initially postulated as a pigment-binding protein (Li et al., 2000), studies of PsbS crystal structure (Fan et al., 2015) and molecular function in qE (Dominici et al., 2002; Bonente et al., 2008; Wilk et al., 2013) revealed its inability to bind pigments. In the current absence of experimental evidence for transient binding of pigments by PsbS, it is accepted that PsbS is not directly involved in quenching, but it is rather a sensor of lumen pH that prompts the formation of quencher(s) (Dominici et al., 2002; Bonente et al., 2008; Sacharz et al., 2017; Giovagnetti and Ruban, 2018). Similar functions in qE have been previously hypothesized for PsbS and LHCX1 (Lepetit et al., 2012). Here we provide biochemical and spectroscopic evidence that LHCX1 does not bind pigments in a way the other LHCF/LHCR antenna components of diatoms bind, although transient binding of pigments upon qE induction cannot be ruled out. One possible explanation to our results is that LHCX1 role in qE might be analogous to that of PsbS. Despite the mechanism of action of LHCX1 and PsbS is unknown and likely to be different, their inability of tightly binding pigments might have evolved to constitutively promote and modulate changes in light-harvesting antenna complexes that lead to quenching.

A number of observations further supports our data interpretation. Very few LHCF4 residues coordinating chlorophylls and carotenoids are conserved in the sequence of the mature LHCX1 protein (four chlorophyll- and one fucoxanthin-binding residues out nine chlorophyll- and seven fucoxanthin-binding sites reported per FCP monomer, Wang et al., 2019; Supplemental Figure S18A). This is in line with the conservation of some residues associated to chlorophyll binding between PsbS and Lhcb1 (LHCII; Liu et al., 2004) and Lhcb4 (CP29; Pan et al., 2011; Supplemental Figure S18B). Despite this, PsbS is not a canonical pigment-binding protein (Dominici et al., 2002; Bonente et al., 2008; Fan et al., 2015).

The oligomerization states of the pennate *P. tricornutum* FCP complexes are under debate, with FCP basic units reported either as trimers or (homo)dimers (Lepetit et al., 2007; Gundermann et al., 2013; Wang et al., 2019; Büchel, 2020; Nagao et al., 2021). *Phaeodactylum tricornutum* FCP antenna was crystallized in a homodimeric state, corresponding to the product of *LHCF3–4* genes that encode for identical sequences (Wang et al., 2019). Differently, Fcp6 protein (an ortholog of LHCX1) was shown to be a stable subunit of antenna trimeric complexes (FCPa) of centric diatoms, such as *Cyclotella meneghiniana* (Lepetit et al., 2010; Gundermann and Büchel, 2012; Gundermann et al., 2013; Ghazaryan et al., 2016). However, using knockdown mutants, Fcp6 was found to affect large-scale pigment–pigment interactions during FCP aggregation, rather than pigment

interactions within FCPa (Ghazaryan et al., 2016). Moreover, LHCX proteins were not found in the structure of *C. gracilis* PSII-FCPII supercomplex recently elucidated (Pi et al., 2019), in line with the absence of PsbS from the structure of plant  $C_2S_2M_2$ -type PSII-LHCII supercomplex (Su et al., 2017). This could indicate that LHCX1 is loosely associated/interacting with supercomplexes, rather than being a “structural component” of the supercomplex, hence explaining its ability to undergo strong overexpression. Additionally, LHCX1 hydrophobicity (10 phenylalanine and 2 tryptophan residues; Figure 6A) could suggest its promiscuity in interacting with FCP antenna complexes, which would agree with the highly hydrophobic character of PsbS (17 phenylalanine residues; Figure 6A) and reported interactions with major and minor antenna complexes of plants (Correa-Galvis et al., 2016a; Sacharz et al., 2017).

We hypothesize that two evolutionarily distant proteins might have emerged de novo to fulfill dynamic photoprotection in turbulent waters (LHCX1) and terrestrial habitats (PsbS), allowing rapid control of the photosynthetic electron and proton transport. Although the evolution of LHCSR/LHCX proteins is currently unresolved (Dittami et al., 2010; Giovagnetti and Ruban, 2018), LHCX or “LHCX-like” proteins—which appear not to be found in red algae and cryptophytes, besides plants (Peers et al., 2009; Giovagnetti and Ruban, 2018; Haniewicz et al., 2018)—are largely distributed in other organisms that possess red algal-derived secondary plastids (Figure 6B; Supplemental Figure S19). Together with diatoms, they are the major and most successful eukaryotic taxa populating the oceans (Falkowski et al., 2004; Armbrust, 2009). “LHCX-like” proteins were observed to respond to light stress, although their direct involvement in qE, and capacities to bind pigments or sense pH are currently unknown (Dittami et al., 2010; Xiang et al., 2015; Chukhutsina et al., 2017; Giovagnetti and Ruban, 2018; Figure 6B; Supplemental Figure S19). Recently, binding of zeaxanthin to *Nannochloropsis oceanica* LHCX1, a homolog of diatom LHCX1, has been proposed to occur during qE (Park et al., 2019), although biochemical evidence for such a function has yet to be shown. We propose that the adaptation to the ever-changing light environment of turbulent waters resulted in the evolution of LHCX(-like) proteins in photosynthetic organisms equipped with secondary red plastids, such as diatoms. LHCX(-like) proteins, essential for the dynamic tracking of light fluctuations, might have contributed to their ecological success in contemporary oceans.

## Materials and methods

### Cell cultures and growth conditions

*Phaeodactylum tricornutum* Bohlin CCAP 1052/1A (accession Pt2; Culture Collection of Algae and Protozoa, Scotland, UK) (De Martino et al., 2007) was grown in sterile artificial seawater F/2 medium, supplemented with silicate (Guillard and Ryther, 1962). Nonaxenic, semi-continuous batch cultures were kept at 18°C in glass bottles and continuously flushed with sterile air. Cells were grown either under CL

(40  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ , in a photoperiod consisting of 14 h of light/10 h of darkness) or IL (40  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ , with cycles of 5 min light/55 min dark) (Lavaud et al., 2002; Giovagnetti and Ruban, 2017). Light conditions and temperature were controlled using Sanyo MLR-351-PE growth chambers. Cell concentration and growth rates were measured on three-six biological samples to ensure experiments were performed with cells in exponential phase. An improved Neubauer haemocytometer (VWR International Ltd., UK) and a CH20 microscope (Olympus, UK) were used. *LHCX1* KO lines were grown either under CL or IL conditions, similarly to Pt2 WT cells. WT-, D95N-, E205Q- and D95N E205Q-*LHCX1* complementation cell lines were grown under CL (10  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ , in a photoperiod of 12 h of light/12 h of darkness) and kept at 19°C. Cell concentration and growth rates were measured in mutant lines, and experiments were performed using cells in exponential phase. All experiments were repeated at least 3 times on independent biological samples, unless otherwise stated.

### Generation of *LHCX1* KO and complemented cell lines

TALE nuclease (TALEN) pairs (pCLS25637 and pCLS25527) were designed and produced by Collectis SA (8 Rue de la Croix Jarry, 75013, Paris, France) to bind upstream and downstream the targeted cleavage site of *LHCX1* gene (TGCTTCGCTCCGGCCC sequence and the sequence corresponding to TCTACTAGCCTTCAGTA on the opposite strand, respectively, from +45 to +92 relative to the start codon), and co-transformed in Pt2 with NAT resistance vector by microparticle bombardment as described in (Falcitatore et al., 1999). Three nourseothricin N-acetyl transferase (NAT)-resistant transgenic lines containing both TALENs were screened for mutagenic events by T7 endonuclease assay (Daboussi et al., 2014). These lines were subcloned and the *LHCX1* gene for individual subclones was sequenced. From the KO6 line, the five subclones analyzed exhibited all of the same biallelic mutation in *LHCX1* gene, that is, +65- and +4-nt insertion between the TALEN target sites, resulting in a frame shift and premature stop codon at the 57th and 28th residue, respectively. For the complementation of *LHCX1* KO6.6 mutant, the *LHCX1* nucleotide sequence has been modified at the TALEN target sites, without changing the translated amino acid sequence, by overlap extension PCR of the two fragments amplified using the *LHCX1*.C.Fw and *LHCX1*tal.Rv, and *LHCX1*tal.Fw and *LHCX1*.E.Rv primer pairs (Supplemental Table S4) and the Pt2 genomic DNA as template. As summarized in Supplemental Table S5, this modified *LHCX1* sequence (WT-*LHCX1*), as well as the 671-nt long region upstream *LHCX1* start codon and the 284-nt long region downstream *LHCX1* stop codon, amplified with the primer pairs described in Supplemental Table S4, were individually inserted into level 0 vector (pL0, MN436816), and then assembled together using BsaI restriction sites into level L1 vector (pCAo2, MN436805) with the uLoop assembly system (Pollak et al., 2020). D95N or/and E205Q site-directed mutant constructs were generated by overlap extension PCR of the two

fragments obtained using primer pairs shown in Supplemental Table S4: *LHCX1*.C.Fw and *LHCX1*D95N.Rv or *LHCX1*E205Q.Rv, and *LHCX1*.E.Rv and *LHCX1*D95N.Fw or *LHCX1*E205Q.Fw with WT-*LHCX1* or D95N-*LHCX1* level 1 vector, as template. Then, the *LHCX1* constructs were either assembled into level 2 vector (pCAe1, MN436800) using SapI (Pollak et al., 2020), with Bleomycin resistance cassette (amplified from pAF6 vector (Falcitatore et al., 1999) with FcpFp.A.Fw and FcpAt.F.Rv primer pairs (Supplemental Table S4) and cloned into level 1 vector (pCAo1, MN436804)), and introduced in *LHCX1* KO6.6 mutant as single plasmid, or recombined into *Sma*I-linearized pUC19 to improve plasmid yield and co-transformed with the Bleomycin-resistant cassette-carrying plasmid pAF6 (Falcitatore et al., 1999). Bleomycin-resistant colonies were obtained, screened by PCR for the presence of the transgene, and then analyzed for *LHCX1* expression by western blot as described in Taddei et al. (2016).

### Chlorophyll fluorescence measurements

Chlorophyll fluorescence quenching induction and relaxation kinetics were assessed using a DUAL-PAM-100 measuring system (Walz Effeltrich, Germany) on 20–30 min dark-adapted cells of *P. tricornutum* (Pt2 WT and KO lines) acclimated to CL or IL. Measurements were performed in a quartz cuvette with cells gently stirred. Cells (in exponential phase) were harvested by centrifugation (1,500  $\times g$  for 3 min) and resuspended in their culture medium to reach a final chlorophyll *a* concentration of 10  $\mu\text{g chlorophyll a mL}^{-1}$ . Chlorophyll content was quantified as previously described (Giovagnetti and Ruban, 2017). Chlorophyll *a* and *c* concentration was determined spectrophotometrically (Ultrospec 2100 pro, GE Healthcare Ltd., UK) (Jeffrey and Humphrey, 1975). Two consecutive cycles, each consisting of 5 min of actinic light followed by 7 min of recovery in the dark, were used (seven-eight biological replicates per light acclimation, with a total 24 and 32 repeats for CL and IL cells, respectively). Actinic light and saturating pulses were provided by an array of 635 nm light-emitting diodes. Actinic light intensity corresponding to 1,650  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$  was used. A saturating pulse (4,000  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ , 600 ms) was applied every minute, and 10 s after the actinic light was switched off.  $F_0$  and  $F_m$  are the minimum and maximum fluorescence levels in the dark, and  $F_v/F_m$  is the maximum quantum yield of PSII in the dark, where  $F_v = (F_m - F_0)$ . Nonphotochemical quenching is calculated as  $([F_m - F_m'] / F_m')$ , where  $F_m'$  is the maximum fluorescence during the course of actinic illumination. For *LHCX1* complementation lines, chlorophyll fluorescence quenching induction and relaxation were measured on a fluorescence CCD camera recorder (JTS-10; BeamBio, France). About 0.5–1 million cells  $\text{mL}^{-1}$  in exponential phase were harvested and concentrated  $\sim 10$  times before measurements.  $qE$  formation was induced by 600  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$  actinic light (6 min), while recovery was monitored in the dark (25 min). Twelve lines per *LHCX1* complementation construct were tested (three independent biological replicates,  $n = 6 \pm \text{SD}$  for each



line). Fitting of qE relaxation data was performed on Origin software (sigmoidal, Hill equation;  $y = y_0 + qE \cdot x^n / (t_{1/2}^n + x^n)$ ). Lines with  $qE < 0.25$  were discarded from the average data shown to ensure accurate fitting. Twelve lines per construct were fitted (three independent biological replicates,  $n = 60 \pm \text{SD}$ ; for Pt2  $n = 24 \pm \text{SD}$ ).

### Thylakoid isolation and sucrose-gradient ultracentrifugation

TMs were isolated from cells dark-adapted for 30 min. TMs were isolated as previously described (Veith and Büchel, 2007) with some modifications. All following preparation steps were carried out in dim light and at 4°C. Cells were harvested by centrifugation (1,500g for 5 min) and resuspended in a homogenization buffer (10 mM HEPES pH 7.6, 6 mM EDTA, 0.6 M sorbitol, 1 mM aminocaproic acid, 0.2 mM benzamidine). They were broken by one passage through a LM20 Microfluidizer (Microfluidics, USA) at 18,000 psi. Cell debris and unbroken cells were removed by centrifugation at 1,500g for 3 min. The supernatant was collected and centrifuged for 20 min at 20,000g, thus obtaining a final pellet that was resuspended in osmotic resuspension buffer (10 mM HEPES pH 7.6, 6 mM EDTA, 1 mM aminocaproic acid, 0.2 mM benzamidine). Chlorophyll *a* concentration of TMs was determined as described above. Separation of photosynthetic protein complexes was obtained by linear sucrose density-gradient ultracentrifugation. Equal amounts of TMs (0.75 mg chlorophyll *a* mL<sup>-1</sup>) isolated from CL and IL cells were mildly solubilized to a final concentration of 0.6% n-dodecyl-β-D-maltoside (β-DDM, Generon, UK; w/v) at a final detergent/chlorophyll *a* ratio of 13.5 for 30 min on ice. Solubilization of TMs (0.35 mg chlorophyll *a* mL<sup>-1</sup>) to a final concentration of 2% β-DDM (w/v, at a final detergent/chlorophyll *a* ratio of 100) was also tested. During solubilization, samples were gently mixed every 5 min. The solubilized membranes were centrifuged at 14,000 rpm for 1 min to remove unsolubilized material. The supernatant, containing the solubilized pigment–protein complexes, was collected and immediately loaded onto linear sucrose gradients (0.1–0.7 M sucrose [w/v] in osmotic resuspension buffer supplemented with 0.03% β-DDM). Sucrose density gradient separation was performed on biological triplicates, each replicate consisting of five-six tubes (technical replicates). A final chlorophyll *a* concentration of 0.35 mg was loaded onto each tube. Photosynthetic protein complexes were fractionated by ultracentrifugation using a swinging bucket rotor (SW41 Ti, Beckman Coulter, Wycombe, UK) for 18 h at 273,620g and at 4°C (Optima L-80 XP Ultracentrifuge, Beckman Coulter, UK). After the separation, sucrose-gradient bands and fractions were rapidly collected and characterized by UV–visible absorption and low-temperature fluorescence spectroscopy (see below). Bands of interest were pooled together and used for further purification. Immediately after collection, band/fraction aliquots were flash-frozen with liquid nitrogen and kept at –20°C for further gel electrophoresis, immunoblotting, and pigment analysis.

### SEC

SEC was performed on bands collected after sucrose density gradient ultracentrifugation. Same bands were harvested from technical replicates (five-six tubes) of the same biological replicate, pooled together, and concentrated for SEC. Purifications were repeated on biological triplicates. Samples were injected through a 100 μL loop and loaded onto a Superdex 200 10/300 GL size exclusion column (GE Healthcare, Hatfield, UK). Fractions were eluted with filtered and degassed buffer containing 10 mM HEPES (pH 7.6) and 0.03% β-DM. Collected fractions were subjected to further analysis.

### Absorption spectroscopy

Absorption spectra were measured at room temperature. Measurements were performed on samples diluted to OD in the range 0.2–0.3 at the Q<sub>y</sub> band. Absorption spectra were measured between 350 and 750 nm with 1-nm increments, on a modernized Aminco DW-2000 UV/Vis spectrophotometer (Olis Inc., Athens, GA, USA). Spectra were normalized at their absolute maximum in the Soret region. Each spectrum shown is the average of three independent biological replicates, unless otherwise stated. Absorption spectra were used to calculate the relative distribution of chlorophyll molecules in the bands and fractions harvested from sucrose density gradients as  $[\text{chlorophyll}] = \left( \int_{600}^{750} A(\lambda) \times d(\lambda) \times V \times D \right) / \left( \int_{600}^{750} A(\lambda) \times d(\lambda) \right)$ , where  $\int_{600}^{750} A(\lambda) \times d(\lambda)$  is the integral of the absorption in function of wavelength (from 600 to 750 nm), *V* is the volume of the band harvested, and *D* is the dilution factor applied to samples prior to absorption spectrum acquisition.

### Low-temperature steady-state fluorescence spectroscopy

Low temperature (77 K) fluorescence emission and excitation spectra were measured using a FluoroMax-3 spectrophotometer (HORIBA Jobin Yvon, Longjumeau, France) equipped with a cryostat cooled by liquid nitrogen. Sample dilutions equal to those used for absorption spectroscopy were applied, which prevented re-absorption and re-emission artifacts during measurements. Samples were excited at 435 nm, and emission was acquired from 600 to 800 nm (fluorescence emission spectra). Fluorescence excitation spectra were detected at 690 nm and recorded between 380 and 640 nm. The emission spectral resolution was 1 nm, while excitation spectral resolution was 0.5 nm, with excitation and emission bandwidths of 1 and 5 nm, respectively. An integration time of 0.1 s was used to reduce the noise level. Five scans were taken for each spectrum and then averaged. Fluorescence excitation spectra were corrected for variations in the detector efficiency, and excitation lamp intensity and distribution, with files provided by the manufacturer.

### CD

Spectra were recorded in the visible region (400–750 nm) at 20°C, on a Chirascan instrument (Applied Photophysics, Leatherhead, UK). Sample points were taken every

nanometer (1 s acquisition time) and three repeat scans per spectrum were carried out using a 1 cm path-length cuvette. For direct comparison of CD spectra, 12.5 µg of total protein was used for each sample. Spectra are presented in CD ellipticity values ( $\theta$  millidegrees), after baseline spectrum subtraction.

### Pigment analysis

Pigments of bands harvested from sucrose gradients and SEC fractions were extracted in 90% methanol. Means of at least five biological samples  $\pm$ SE are shown for the depoxidation state of *P. tricornutum* cultures. Means of three biological samples ( $n = 4$ –5)  $\pm$ SE are shown for sucrose density-gradient bands, while means of technical replicates ( $n = 3$ –5)  $\pm$ SE from a representative SEC elution profile are presented. Samples were centrifuged at 10,000 g for 5 min and resulting supernatants were filtered (0.2 µm). An aliquot of 200 µL was loaded in vial inserts to perform high-performance liquid chromatography (HPLC) analysis. Pigment content and composition were determined by reversed-phase HPLC, using a LiChrospher 100 RP-18 column (Merck Chemicals Ltd., UK) and a Dionex Summit chromatography system (Thermo Fisher Scientific, UK) as previously described (Lavaud et al., 2002; Giovagnetti and Ruban, 2017).

### SDS–PAGE and immunoblotting

TMs isolated from dark-adapted CL and IL cells (a range of 1.25–10 µg of total protein amount), as well as samples of sucrose density gradient bands/fractions and SEC fractions (both loaded as identical sample volume) were loaded onto 16% Tricine–SDS–PAGE gels according to Schagger (2006). Protein content was quantified by Bradford assay (Bradford, 1976). Gels were either stained by InstantBlue protein stain (Expedeon Ltd., Over, UK) (gently shaking for 1 h and 30 min, at room temperature) or used for electroblotting onto nitrocellulose membranes (GE Healthcare, UK). Densitometric analysis of InstantBlue-stained SDS–PAGE protein bands was carried out to assess the contribution of LHCX1 to light-harvesting antenna and total protein content. An anti-LHCSR antibody was used (gift of K. K. Niyogi, University of California, Berkeley, USA), capable to detect all the four LHCX isoforms in *P. tricornutum* (Bailleul et al., 2010; Taddei et al., 2016) (1:5,000 dilution and overnight incubation at 4°C). Accurate blotting was verified by correct transfer of color prestained protein standard (P7712 or P7719, New England BioLabs, Hitchin, UK) on the nitrocellulose membranes. The  $\beta$ -subunit of ATP synthase (ATP-B; AS05085, Agrisera, Sweden; 1:4,000) was used as a loading control for western blots and to normalize LHCX1 protein signal. InstantBlue-stained gels were scanned using a ChemiDoc Touch Imaging System (Bio-Rad, USA). Antibody signals were detected after incubation with a secondary goat anti-rabbit antibody (IRDye 800CW, LI-COR Biosciences Ltd., Cambridge, UK; 1:20,000) and visualized by near-infrared fluorescence detection (Odyssey Imaging System, LI-COR Biosciences Ltd., UK). Quantitative densitometric analysis of protein signals was carried out on Image Studio Lite (LI-COR

Biosciences Ltd., UK) or ImageJ Fiji. Quantification of protein signals shown in this study is averages of western blots carried out on independent biological replicates  $\pm$ SE (six independent biological samples [ $n = 22$ ] for TMs, 3–4 independent biological samples for sucrose density-gradient bands/fractions and SEC fractions, and 3–4 independent biological replicates for KO and LHCX1 complementation mutants).

### Tandem LC–MS/MS analysis

LC–MS/MS analysis was performed at the Cambridge Centre for Proteomics, Cambridge Systems Biology Centre (Department of Biochemistry, University of Cambridge, UK). The 1D gel bands were transferred into a 96-well PCR plate. The bands were cut into 1 mm<sup>2</sup> pieces, destained, reduced (DTT) and alkylated (iodoacetamide), and subjected to enzymatic digestion with chymotrypsin overnight at 37°C. After digestion, the supernatant was pipetted into a sample vial and loaded onto an autosampler for automated LC–MS/MS analysis. LC–MS/MS experiments were performed using a Dionex Ultimate 3,000 RSLC nanoUPLC (Thermo Fisher Scientific Inc., Waltham, MA, USA) system and an Orbitrap Lumos mass spectrometer (Thermo Fisher Scientific Inc., USA). Peptides were loaded onto a precolumn (Thermo Scientific PepMap 100 C18, 5 µm particle size, 100A pore size, 300 µm i.d.  $\times$  5 mm length) from the Ultimate 3,000 auto-sampler with 0.1% formic acid (v/v) for 3 min at a flow rate of 10 µL/min. After this period, the column valve was switched to allow elution of peptides from the precolumn onto the analytical column. Separation of peptides were achieved by C18 reverse-phase chromatography at a flow rate of 300 nL/min and a Thermo Scientific reverse-phase nano Easy-spray column (Thermo Scientific PepMap C18, 2 µm particle size, 100A pore size, 75 µm i.d.  $\times$  50 cm length) at a flow rate of 300 nL/min. Solvent A was water + 0.1% formic acid (v/v) and solvent B was 80% acetonitrile, 20% water + 0.1% formic acid (v/v). The linear gradient employed was 2–40% B over 30 min. Total LC run time was 60 min including high organic wash step and column re-equilibration. The eluted peptides from C18 column LC eluant were sprayed into the mass spectrometer by means of an Easy-Spray source (Thermo Fisher Scientific Inc., USA). All  $m/z$  values of eluting peptide ions were measured in an Orbitrap mass analyzer, set at a resolution of 120,000, and were scanned between  $m/z$  380–1,500 Da. Data-dependent MS/MS scans (3 s cycle time) were employed to automatically isolate and fragment precursor ions and generate fragment ions by higher energy collisional-induced dissociation (normalized collision energy: 38%) in the ion routing multipole. The resolution of the Orbitrap was set to 15,000 for the measurement of fragment ions. Singly charged ions, ions with greater than seven charges, and ions with unassigned charge states were excluded from being selected for MS/MS and a dynamic exclusion window of 70 s was employed. Postrun, the data were processed using Protein Discoverer (version 2.1., Thermo Fisher Scientific Inc., USA). Briefly, all MS/MS data were converted to mgf files and the files were

then submitted to the Mascot search algorithm (Matrix Science, London, UK) and searched against the UniProt *Phaeodactylum tricornutum*\_20180123 database (10,795 sequences; 4,974,652 residues) and a database of common contaminant sequences (115 sequences, 38,274 residues). Variable modifications of oxidation (M), deamidation (NQ), and carbamidomethyl were applied. The peptide and fragment mass tolerances were set to 5 ppm and 0.1 Da, respectively. A significance threshold value of  $P < 0.05$  and a peptide cut-off score of 20 were also applied. Peak areas corresponding to the intact peptides were calculated from the chromatograms within PD. All data were then imported into the Scaffold program (version\_4.5.1, Proteome Software Inc., Portland, OR, USA).

### Protein sequence alignment and protein secondary and tertiary structure predictions

Clustal Omega was used to generate pairwise and multiple protein sequence alignments (Sievers et al., 2011). Protein secondary structures were predicted using Phyre2 (Kelley et al., 2015), I-TASSER (Roy et al., 2010; Yang et al., 2015), and Robetta (Kim et al., 2004) servers. The 3-D structural models of LHCX1 were generated using the Phyre2 web portal using intensive modeling mode (multiple templates and ab initio techniques, with 93% of residues modeled at  $> 90\%$  confidence). An additional 3-D structural model was predicted using the I-TASSER server (Roy et al., 2010; Yang et al., 2015) based on the crystal structure of the homodimeric FCP antenna complex of *P. tricornutum* (6A2W, LHCf3-LHCf4 genes) (Wang et al., 2019; C-score = 0.63, estimated TM-score =  $0.80 \pm 0.09$ , estimated RMSD =  $4.0 \pm 2.7$  Å).

### Statistical analyses

Statistical significance was assessed by two-tailed Student's *t* test. Linear regressions with 95% confidence values were plotted using SigmaPlot version 14. The relationship observed between LHCX1 accumulation and qE formation was tested between WT- and D95N-, E205Q-, and D95N E205Q-LHCX1 complementation lines. Because of the nonindependence between data points that arises from the use of more than one sample per gel and different gels during western blotting, the relationship between qE and LHCX1 was analyzed by fitting a mixed-effects model using the lme4 (Bates et al., 2015) and lmerTest packages (Kuznetsova et al., 2017) running in R version 4.02 (R Core Team, 2020). "Gel" was included as a random factor in a random slopes and intercepts model. "Replicate" was included as a fixed effect because there were only three levels making it unsuitable for use as a random factor, and both "cell line" and "LHCX1" were included as main effects and also as an interaction between the two. Initial data exploration found strong positive skew in both qE and LHCX1, thus for the model fitting both variables were log-transformed after the addition of a constant (0.1) to make the few negative values positive (Supplemental Figures S10 and S11). The significance of the main effects and the interaction between log(LHCX1 + 0.1)

and cell line was assessed using deletion tests with Satterthwaite's method. The interaction term was nonsignificant ( $F_{3,109.9} = 0.337$ ,  $P = 0.798$ ), as was the main effect of cell line once the interaction term was removed from the model ( $F_{3,116.6} = 0.7728$ ,  $P = 0.512$ ).

### Accession numbers

LHCX1 gene: UniProtKB - B7FYLO (B7FYLO\_PHATC).

### Supplemental data

The following materials are available in the online version of this article.

**Supplemental Table S1.** De-epoxidation state (diatoxanthin/(diadinoxanthin + diatoxanthin) of *P. tricornutum* cells acclimated to CL and IL.

**Supplemental Table S2.** Tandem LC-MS/MS analysis on *P. tricornutum* LHCX1 bands resolved by Tricine-SDS-PAGE.

**Supplemental Table S3.** Pigment content of bands and fractions isolated by sucrose density gradient and SEC.

**Supplemental Table S4.** List of primers used in this study

**Supplemental Table S5.** List of the constructions generated in this study, showing the insert, receiving vector and method used for insertion.

**Supplemental Figure S1.** *P. tricornutum* TM proteins resolved by 16% Tricine-SDS-PAGE.

**Supplemental Figure S2.** Generation and characterization of TALEN-induced LHCX1 KO mutant lines of *P. tricornutum* (Pt2).

**Supplemental Figure S3.** qE capacity in *P. tricornutum* WT cells and LHCX1 KO mutants.

**Supplemental Figure S4.** Nonphotochemical quenching, growth capacity and PSII quantum yield in the dark of *P. tricornutum* LHCX1 KO mutants.

**Supplemental Figure S5.** Generation and characterization of WT-LHCX1 complementation construct of *P. tricornutum* (Pt2).

**Supplemental Figure S6.** Protein composition of sucrose density gradient band and fraction samples resolved by 16% Tricine-SDS-PAGE.

**Supplemental Figure S7.** Low temperature (77 K) fluorescence emission of bands/fractions separated by sucrose density gradient of IL thylakoids.

**Supplemental Figure S8.** Changing solubilization of TMs from dark-adapted IL-grown *P. tricornutum* cells alters band pattern and separation, and LHCX1 distribution along the sucrose density gradient.

**Supplemental Figure S9.** Absorption spectra of bands separated by sucrose density gradient and fractions eluted by SEC for the IL acclimation.

**Supplemental Figure S10.** Protein composition of SEC fractions resolved by 16% Tricine-SDS-PAGE.

**Supplemental Figure S11.** (SEC of band B3 (FCP antenna; IL condition) and protein composition of eluted fractions.

**Supplemental Figure S12.** Predicted models of LHCX1 tertiary structure.

**Supplemental Figure S13.** Generation and characterization of WT-, D95N-, E205Q-, and D95N E205Q-LHCX1 complementation constructs of *P. tricornutum* (Pt2).



**Supplemental Figure S14.** Characterization of WT-, D95N-, E205Q-, and D95N E205Q-LHCX1 complementation constructs of *P. tricornutum* (Pt2).

**Supplemental Figure S15.** qE relaxation and sigmoidicity of WT-, D95N-, E205Q-, and D95N E205Q-LHCX1 complementation constructs of *P. tricornutum* (Pt2).

**Supplemental Figure S16.** Relationship between LHCX1 accumulation and qE formation observed in LHCX1 complementation constructs of *P. tricornutum* (Pt2).

**Supplemental Figure S17.** Log transformation of LHCX1/ATP-B signals and qE data.

**Supplemental Figure S18.** Pairwise sequence alignment of LHCX1 and LHCF4 proteins, and multiple sequence alignment of PsbS, Lhcb1 (LHCII), and Lhcb4 (CP29).

**Supplemental Figure S19.** Multiple sequence alignments of LHCX/LH818-like proteins related to *P. tricornutum* LHCX1.

## Acknowledgments

We thank Y. Tashiro for their initial assistance with CD spectroscopy. F. Saccon, S. Wilson, and B. Bailleul are acknowledged for fruitful discussions on the manuscript, as well as C.A.F. Haste for helping during initial experiments. R.J. Knell is acknowledged for advice and help with statistical analyses. K.K. Niyogi is acknowledged for kindly providing the anti-LHCSR antibody used in this study.

## Funding

This work was supported by grants funded by The Leverhulme Trust (RPG-2018-199), Biotechnology and Biological Sciences Research Council (BB/R015694/1), and The Royal Society Wolfson Research Merit Award (WM140084) to A.V.R. This work was also supported by funding from the European Assemble plus (Association of European Marine Biological Research Laboratories Expanded, H2020-INFRAIA-1-2016-2017), the Fondation Bettencourt-Schueller and the 'Initiative d'Excellence' program (Grant 'DYNAMO', ANR-11-LABX-0011-01), the ANR "BrownCut" (ANR-19-CR20-0020) to A.F., and the Gordon and Betty Moore Foundation (GBMF4981.01) to M.J.

**Conflict of interest statement.** The authors declare no competing interests.

## Data and materials availability

All data are available in the main text or [supplemental materials](#).

## References

- Allorent G, Lefebvre-Legendre L, Chappuis R, Kuntz M, Truong TB, Niyogi KK, Ulm R, Goldschmidt-Clermon M (2016) UV-B photoreceptor-mediated protection of the photosynthetic machinery in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci USA* **113**: 14864–14869
- Armbrust EV (2009) The life of diatoms in the world's oceans. *Nature* **459**: 185–192
- Bailleul B, Rogato A, de Martino A, Coesel S, Cardol P, Bowler C, Falcione A, Finazzi G (2010) An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proc Natl Acad Sci USA* **107**: 18214–18219
- Aspinall-O'Dea M, Wentworth M, Pascal A, Robert B, Ruban AV, Horton P (2002) *In vitro* reconstitution of the activated zeaxanthin state associated with energy dissipation in plants. *Proc Natl Acad Sci USA* **99**: 16331–16335
- Ballottari M, Truong TB, De Re E, Erickson E, Stella GR, Fleming GR, Bassi R, Niyogi KK (2016) Identification of pH-sensing sites in the light harvesting complex stress-related 3 protein essential for triggering non-photochemical quenching in *Chlamydomonas reinhardtii*. *J Biol Chem* **291**: 7334–7346
- Bates D, Maechler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* **67**: 1–48
- Bonente G, Howes BD, Caffarri S, Smulevich G, Bassi R (2008) Interactions between the photosystem II subunit PsbS and xanthophylls studied *in vivo* and *in vitro*. *J Biol Chem* **283**: 8434–8445
- Bonente G, Ballottari M, Truong TB, Morosinotto T, Ahn TK, Fleming GR, Niyogi KK, Bassi R (2011) Analysis of LHCSR3, a protein essential for feed-back de-excitation in the green alga *Chlamydomonas reinhardtii*. *PLoS Biol* **9**: e1000577
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248–254
- Büchel C (2020) Light harvesting complexes in chlorophyll *c*-containing algae. *Biochim Biophys Acta* **1861**: 148027
- Buck JM, Sherman J, Río Bártulos C, Serif M, Halder M, Henkel J, Falcione A, Lavaud J, Gorbunov MY, Kroth PG, et al. (2019) Lhc proteins provide photoprotection via thermal dissipation of absorbed light in the diatom *Phaeodactylum tricornutum*. *Nat Commun* **10**: 4167
- Calvaruso C, Rokka A, Aro EM, Büchel C (2020) Specific Lhc proteins are bound to PSI or PSII supercomplexes in the diatom *Thalassiosira pseudonana*. *Plant Physiol* **183**: 67–79
- Chukhutsina VU, Fristedt R, Morosinotto T, Croce R (2017) Photoprotection strategies of the alga *Nannochloropsis gaditana*. *Biochim Biophys Acta* **1858**: 544–552
- Correa-Galvis V, Poschmann G, Melzer M, Stühler K, Jahns P (2016a) PsbS interactions involved in the activation of energy dissipation in *Arabidopsis*. *Nat Plants* **2**: 15225
- Correa-Galvis V, Redekop P, Guan K, Griess A, Truong TB, Wakao S, Niyogi KK, Jahns P (2016b) Photosystem II subunit PsbS is involved in the induction of LHCSR protein-dependent energy dissipation in *Chlamydomonas reinhardtii*. *J Biol Chem* **291**: 17478–17487
- Croce R, Weiss S, Bassi R (1999) Carotenoid-binding sites of the major light-harvesting complex II of higher plants. *J Biol Chem* **274**: 29613–29623
- Croce R, Canino G, Ros F, Bassi R (2002) Chromophore organization in the higher-plant photosystem II antenna protein CP26. *Biochemistry* **41**: 7334–7343
- Dall'Osto L, Lico C, Alric J, Giuliano G, Havaux M, Bassi R (2006) Lutein is needed for efficient chlorophyll triplet quenching in the major LHCII antenna complex of higher plants and effective photoprotection *in vivo* under strong light. *BMC Plant Biol* **6**: 32
- Daboussi F, Leduc S, Maréchal A, Dubois G, Guyot V, Perez-Michaut C, Amato A, Falcione A, Juillerat A, Beurdeley M, et al. (2014) Genome engineering empowers the diatom *Phaeodactylum tricornutum* for biotechnology. *Nat Commun* **5**: 3831
- de la Cruz Valbuena G, Camargo FVA, Borrego-Varillas R, Perozeni F, D'Andrea C, Ballottari M, Cerullo G (2019) Molecular mechanisms of nonphotochemical quenching in the LHCSR3 protein of *Chlamydomonas reinhardtii*. *J Phys Chem Lett* **10**: 2500–2505



- De Martino A, Meichenin A, Shi J, Pan K, Bowler C (2007) Genetic and phenotypic characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) accessions. *J Phycol* **43**: 992–1009
- Dinc E, Tian L, Roy LM, Roth R, Goodenough U, Croce R (2016) LHCSR1 induces a fast and reversible pH-dependent fluorescence quenching in LHCI in *Chlamydomonas reinhardtii* cells. *Proc Natl Acad Sci USA* **113**: 7673–7678
- Dittami SM, Michel G, Collén J, Boyen C, Tonon T (2010) Chlorophyll-binding proteins revisited - a multigenic family of light-harvesting and stress proteins from a brown algal perspective. *BMC Evol Biol* **10**: 365
- Dominici P, Caffarri S, Armenante F, Ceoldo S, Crimi M, Bassi R (2002) Biochemical properties of the PsbS subunit of photosystem II either purified from chloroplast or recombinant. *J Biol Chem* **277**: 22750–22758
- Elrad D, Niyogi KK, Grossman AR (2002) A major light-harvesting polypeptide of photosystem II functions in thermal dissipation. *Plant Cell* **14**: 1801–1816
- Falcatore A, Jaubert M, Bouly JP, Bailleul B, Mock T (2020) Diatom molecular research comes of age: model species for studying phytoplankton biology and diversity. *Plant Cell* **32**: 547–572
- Falcatore A, Casotti R, Leblanc C, Abrescia C, Bowler C (1999) Transformation of nonselectable reporter genes in marine diatoms. *Mar Biotechnol* **1**: 239–251
- Falkowski PG, Katz ME, Knoll AH, Quigg A, Raven JA, Schofield O, Taylor FJR (2004) The evolution of modern eukaryotic phytoplankton. *Science* **305**: 354–360
- Fan M, Li M, Liu Z, Cao P, Pan X, Zhang H, Zhao X, Zhang J, Chang W (2015) Crystal structures of the PsbS protein essential for photoprotection in plants. *Nat Struct Mol Biol* **22**: 729–735
- Field CB, Behrenfeld MJ, Randerson JT, Falkowski PG (1998) Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**: 237–240
- Ghazaryan A, Akhtar P, Garab G, Lambrev PH, Büchel C (2016) Involvement of the LhcX protein Fcp6 of the diatom *Cyclotella meneghiniana* in the macro-organisation and structural flexibility of thylakoid membranes. *Biochim Biophys Acta* **1857**: 1373–1379
- Giovagnetti V, Ruban AV (2017) Detachment of the fucoxanthin chlorophyll *a/c* binding protein (FCP) antenna is not involved in the acclimative regulation of photoprotection in the pennate diatom *Phaeodactylum tricornutum*. *Biochim Biophys Acta* **1858**: 218–230
- Giovagnetti V, Ruban AV (2018) The evolution of the photoprotective antenna proteins in oxygenic photosynthetic eukaryotes. *Biochem Soc Trans* **46**: 1263–1277
- Guillard RR, Ryther JH (1962) Studies of marine planktonic diatoms. I. *Cyclotella nana* Husted and *Detonula confervacea* (Cleve) Gran. *Can J Microbiol* **8**: 229–239
- Gundermann K, Büchel C (2012) Factors determining the fluorescence yield of fucoxanthin-chlorophyll complexes (FCP) involved in non-photochemical quenching in diatoms. *Biochim Biophys Acta* **1817**: 1044–1052
- Gundermann K, Schmidt M, Weisheit W, Mittag M, Büchel C (2013) Identification of several sub-populations in the pool of light harvesting proteins in the pennate diatom *Phaeodactylum tricornutum*. *Biochim Biophys Acta* **1827**: 303–310
- Haniewicz P, Abram M, Nosek L, Kirkpatrick J, El-Mohsnawy E, Janna Olmos JD, Kouril R, Kargul JM (2018) Molecular mechanisms of photoadaptation of photosystem I supercomplex from an evolutionary cyanobacterial/algal intermediate. *Plant Physiol* **176**: 1433–1451
- Jeffrey SW, Humphrey GF (1975) New spectrophotometric equations for determining chlorophylls *a*, *b*, *c1* and *c2* in higher plants, algae and natural phytoplankton. *Biochem Physiol Pflanz* **167**: 191–194
- Keeling PJ (2013) The number, speed, and impact of plastid endosymbioses in eukaryotic evolution. *Annu Rev Plant Biol* **64**: 583–607
- Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* **10**: 845–858
- Kim DE, Chivian D, Baker D (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res* **32**: W526–W531
- Kosuge K, Tokutsu R, Kim E, Akimoto S, Yokono M, Ueno Y, Minagawa J (2018) LHCSR1-dependent fluorescence quenching is mediated by excitation energy transfer from LHCI to photosystem I in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci USA* **115**: 3722–3727
- Kuznetsova A, Brockhoff PB, Christensen RHB (2017) lmerTest Package: tests in linear mixed effects models. *J Stat Softw* **82**: 1–26
- Lavaud J, Rousseau B, van Gorkom HJ, Etienne AL (2002) Influence of the diadinoxanthin pool size on photoprotection in the marine planktonic diatom *Phaeodactylum tricornutum*. *Plant Physiol* **129**: 1398–1406
- Lavaud J, Rousseau B, Etienne AL (2003) Enrichment of the light-harvesting complex in diadinoxanthin and implications for the nonphotochemical fluorescence quenching in diatoms. *Biochemistry* **42**: 5802–5808
- Lepetit B, Volke D, Szabó M, Hoffmann R, Garab G, Wilhelm C, Goss R (2007) Spectroscopic and molecular characterization of the oligomeric antenna of the diatom *Phaeodactylum tricornutum*. *Biochemistry* **46**: 9813–9822
- Lepetit B, Volke D, Gilbert M, Wilhelm C, Goss R (2010) Evidence for the existence of one antenna-associated lipid-dissolved and two protein-bound pools of diadinoxanthin cycle pigments in diatoms. *Plant Physiol* **154**: 1905–1920
- Lepetit B, Goss R, Jakob T, Wilhelm C (2012) Molecular dynamics of the diatom thylakoid membrane under different light conditions. *Photosynth Res* **111**: 245–257
- Lepetit B, Sturm S, Rogato A, Gruber A, Sachse M, Falcatore A, Kroth PG, Lavaud J (2013) High light acclimation in the secondary plastids containing diatom *Phaeodactylum tricornutum* is triggered by the redox state of the plastoquinone pool. *Plant Physiol* **161**: 853–865
- Lepetit B, Gélén G, Lepetit M, Sturm S, Vugrinec S, Rogato A, Kroth PG, Falcatore A, Lavaud J (2017) The diatom *Phaeodactylum tricornutum* adjusts nonphotochemical fluorescence quenching capacity in response to dynamic light via fine-tuned LhcX and xanthophyll cycle pigment synthesis. *New Phytol* **214**: 205–218
- Li XP, Björkman O, Shih C, Grossman AR, Rosenquist M, Jansson S, Niyogi KK (2000) A pigment-binding protein essential for regulation of photosynthetic light harvesting. *Nature* **403**: 391–395
- Li XP, Gilmore AM, Caffarri S, Bassi R, Golan T, Kramer D, Niyogi KK (2004) Regulation of photosynthetic light harvesting involves intrathylakoid lumen pH sensing by the PsbS protein. *J Biol Chem* **279**: 22866–22874
- Liguori N, Roy LM, Opacic M, Durand G, Croce R (2013) Regulation of light harvesting in the green alga *Chlamydomonas reinhardtii*: the C-terminus of LHCSR is the knob of a dimmer switch. *J Am Chem Soc* **135**: 18339–18342
- 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
- Nagao R, Yokono M, Ueno Y, Suzuki T, Kumazawa M, Kato KH, Tsuboshita N, Dohmae N, Ifuku K, Shen JR, et al. (2021) Enhancement of excitation-energy quenching in fucoxanthin chlorophyll *a/c*-binding proteins isolated from a diatom *Phaeodactylum tricornutum* upon excess-light illumination. *Biochim Biophys Acta* **1862**: 148350

- Nawrocki WJ, Liu X, Croce R (2020) *Chlamydomonas reinhardtii* exhibits de facto constitutive NPQ capacity in physiologically relevant conditions. *Plant Physiol* **182**: 472–479
- Pan X, Li M, Wan T, Wang L, Jia C, Hou Z, Zhao X, Zhang J, Chang W (2011) Structural insights into energy regulation of light-harvesting complex CP29 from spinach. *Nat Struct Mol Biol* **18**: 309–315
- Park S, Steen CJ, Lyska D, Fischer AL, Endelman B, Iwai M, Niyogi KK, Fleming GR (2019) Chlorophyll-carotenoid excitation energy transfer and charge transfer in *Nannochloropsis oceanica* for the regulation of photosynthesis. *Proc Natl Acad Sci USA* **116**: 3385–3390
- Peers G, Truong TB, Ostendorf E, Busch A, Elrad D, Grossman AR, Hippler M, Niyogi KK (2009) An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462**: 518–521
- Peterman EJG, Hobe S, Calkoen F, van Grondelle R, Paulsen H, van Amerongen H (1996) Low-temperature spectroscopy of monomeric and trimeric forms of reconstituted light-harvesting chlorophyll *a/b* complex. *Biochim Biophys Acta* **1273**: 171–174
- Petroutsos D, Tokutsu R, Maruyama S, Flori S, Greiner A, Magneschi L, Cusant L, Kottke T, Mittag M, Hegemann P, et al. (2016) A blue-light photoreceptor mediates the feedback regulation of photosynthesis. *Nature* **537**: 563–566
- Pi X, Zhao S, Wang W, Liu D, Xu C, Han G, Kuang T, Sui SF, Shen JR (2019) The pigment-protein network of a diatom photosystem II–light harvesting antenna supercomplex. *Science* **365**: eaax4406
- Pollak B, Matute T, Nuñez I, Cerda A, Lopez C, Vargas V, Kan A, Bielinski V, von Dassow P, Dupont CL, et al. (2020) Universal loop assembly: open, efficient and cross-kingdom DNA fabrication. *Synth Biol* **5**: ysaa001
- Core Team R (2020) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
- Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* **5**: 725–738
- Ruban AV, Lavaud J, Rousseau B, Guglielmi G, Horton P, Etienne A-L (2004) The super-excess energy dissipation in diatom algae: comparative analysis with higher plants. *Photosynth Res* **82**: 165–175
- Ruban AV, Johnson MP, Duffy CDP (2012) The photoprotective molecular switch in the photosystem II antenna. *Biochim Biophys Acta* **1817**: 167–181
- Sacharz J, Giovagnetti V, Ungerer P, Mastroianni G, Ruban AV (2017) The xanthophyll cycle affects reversible interactions between PsbS and light-harvesting complex II to control non-photochemical quenching. *Nat Plants* **3**: 16225
- Schägger H (2006) Tricine–SDS–PAGE. *Nat Protoc* **1**: 16–22.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539
- Szabó M, Lepetit B, Goss R, Wilhelm C, Mustárdy L, Garab G (2008) Structurally flexible macro-organization of the pigment-protein complexes of the diatom *Phaeodactylum tricornutum*. *Photosynth Res* **95**: 237–245
- Su X, Ma J, Wei X, Cao P, Zhu D, Chang W, Liu Z, Zhang X, Li M (2017) Structure and assembly mechanism of plant C<sub>2</sub>S<sub>2</sub>M<sub>2</sub>-type PSII-LHCII supercomplex. *Science* **357**: 815–820
- Taddei L, Stella GR, Rogato A, Bailleul B, Fortunato AE, Annunziata R, Sanges R, Thaler M, Lepetit B, Lavaud J, et al. (2016) Multisignal control of expression of the LHCX protein family in the marine diatom *Phaeodactylum tricornutum*. *J Exp Bot* **67**: 3939–3951
- Taddei L, Chukhutsina VU, Lepetit B, Stella GR, Bassi R, van Amerongen H, Bouly JP, Jaubert M, Finazzi G, Falcietore A (2018) Dynamic changes between two LHCX-related energy quenching sites control diatom photoacclimation. *Plant Physiol* **177**: 953–965
- Tibiletti T, Auroy P, Peltier G, Caffarri S (2016) *Chlamydomonas reinhardtii* PsbS protein is functional and accumulates rapidly and transiently under high light. *Plant Physiol* **171**: 2717–2730
- 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. *Biochim Biophys Acta* **1767**: 1428–1435
- Wang W, Yu LJ, Xu C, Tomizaki T, Zhao S, Umena Y, Chen X, Qin X, Xin Y, Suga M, et al. (2019) Structural basis for blue-green light harvesting and energy dissipation in diatoms. *Science* **363**: eaav0365
- Wilk L, Grunwald M, Liao PN, Walla PJ, Kühlbrandt W (2013) Direct interaction of the major light-harvesting complex II and PsbS in nonphotochemical quenching. *Proc Natl Acad Sci USA* **110**: 5452–5456
- Xiang T, Nelson W, Rodriguez J, Tolleter D, Grossman AR (2015) *Symbiodinium* transcriptome and global responses of cells to immediate changes in light intensity when grown under autotrophic or mixotrophic conditions. *Plant J* **82**: 67–80
- Xu C, Pi X, Huang Y, Han G, Chen X, Qin X, Huang G, Zhao S, Yang Y, Kuang T, et al. (2020) Structural basis for energy transfer in a huge diatom PSI-FCPI supercomplex. *Nat Commun* **11**: 5081
- Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y (2015) The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**: 7–8