

# The intracellular distribution of inorganic carbon fixing enzymes does not support the presence of a C4 pathway in the diatom *Phaeodactylum tricornerutum*

Daniela Ewe<sup>1,4</sup>  · Masaaki Tachibana<sup>2,5</sup> · Sae Kikutani<sup>2,6</sup> · Ansgar Gruber<sup>1,7</sup> · Carolina Río Bártulos<sup>1</sup> · Grzegorz Konert<sup>4</sup> · Aaron Kaplan<sup>3</sup> · Yusuke Matsuda<sup>2</sup>  · Peter G. Kroth<sup>1</sup>

## Abstract

Diatoms are unicellular algae and important primary producers. The process of carbon fixation in diatoms is very efficient even though the availability of dissolved CO<sub>2</sub> in sea water is very low. The operation of a carbon concentrating mechanism (CCM) also makes the more abundant bicarbonate accessible for photosynthetic carbon fixation. Diatoms possess carbonic anhydrases as well as metabolic enzymes potentially involved in C4 pathways; however, the question as to whether a C4 pathway plays a general role in diatoms is not yet solved. While genome analyses indicate that the diatom *Phaeodactylum tricornerutum* possesses all the enzymes required to operate a C4 pathway, silencing of the pyruvate orthophosphate dikinase (PPDK) in a genetically transformed cell line does not lead to reduced photosynthetic carbon fixation. In this study, we have determined the intracellular location of all enzymes potentially involved in C4-like carbon fixing pathways in *P. tricornerutum* by expression of the respective proteins fused to green fluorescent protein (GFP), followed by fluorescence microscopy. Furthermore, we compared the results to known pathways and locations of enzymes in higher plants performing C3 or C4 photosynthesis. This approach revealed that the intracellular distribution of the investigated enzymes is quite different from the one observed in higher plants. In particular, the apparent lack of a plastidic decarboxylase in *P. tricornerutum* indicates that this diatom does not perform a C4-like CCM.

**Keywords** C4 photosynthesis · Chloroplast · Green fluorescent protein (GFP) · Carboxylation · Decarboxylation

## Introduction

Diatoms (Bacillariophyceae), unicellular algae belonging to the Stramenopiles, operate very efficient photosynthesis and are thought to contribute substantially to global carbon fixation (Nelson et al. 1995; Granum et al. 2005; Shen et al. 2017). CO<sub>2</sub>, the substrate of ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), is of limited supply for

Daniela Ewe  
ewe@alga.cz

<sup>1</sup> Fachbereich Biologie, Universität Konstanz,  
78457 Konstanz, Germany

<sup>2</sup> Department of Bioscience, School of Science  
and Technology, Kwansei Gakuin University, Sanda,  
Hyogo 669-1337, Japan

<sup>3</sup> Department of Plant and Environmental Sciences,  
Edmond J. Safra Campus–Givat Ram, Hebrew University  
of Jerusalem, 91904 Jerusalem, Israel

<sup>4</sup> Centre Algatech, Institute of Microbiology of the Czech  
Academy of Sciences, Třeboň, Czech Republic

<sup>5</sup> Lion Corporation Pharmaceutical Laboratories No.1,  
Odawara, Kanagawa 256-0811, Japan

<sup>6</sup> Tech Manage Corp., Tokyo 160-0023, Japan

<sup>7</sup> Biology Centre, Institute of Parasitology, Czech Academy  
of Sciences, České Budějovice, Czech Republic

marine phytoplankton due to (i) the low partial pressure of  $\text{CO}_2$  in the atmosphere, (ii) the much faster  $\text{CO}_2$  hydration rate in the water relative to the  $\text{HCO}_3^-$  dehydration rate, and (iii) the high salinity and slight alkalinity in seawater. The RubisCO enzymes found in diatoms have a  $\text{CO}_2$  half-saturation constant of about 23–68  $\mu\text{M}$  (Young et al. 2016). However, the total amount of dissolved inorganic carbon (DIC) in seawater equals a  $\text{CO}_2$  concentration of about 10–15  $\mu\text{M}$  at pH 8.2 (Riebesell et al. 1993). This discrepancy indicates that diatoms may operate a  $\text{CO}_2$ -concentrating mechanism (CCM) to improve the efficiency of carbon fixation (Roberts et al. 2007a; Reinfelder 2011; Clement et al. 2016). CCMs increase the cellular  $\text{CO}_2$  concentration in spatial proximity to RubisCO, leading to a higher carbon fixation rate in a  $\text{CO}_2$ -limited environment. Indeed, calculations based on oxygen evolution rates of *Phaeodactylum tricorutum* cells indicate that the intracellular concentration of DIC is 4–6 times higher than outside the cell (Badger et al. 1998).

Two main CCM mechanisms are known to operate in diatoms, biophysical and/or biochemical (Reinfelder 2011).

Biophysical CCMs are characterized by active inorganic carbon ( $\text{C}_i$ ) accumulation within the plastid via specific bicarbonate transporters at the plasma membrane (Nakajima et al. 2013), and most probably also at the plastid envelopes, by carbonic anhydrases (CA), which catalyze the inter-conversion of  $\text{CO}_2$  and  $\text{H}_2\text{O}$  to  $\text{HCO}_3^-$  and  $\text{H}^+$  (or vice versa) (Tachibana et al. 2011). A further characteristic is the presence of a pyrenoid that, similar to the bacterial carboxysomes, contains high concentrations of  $\text{CO}_2$ , RubisCO enzyme and in *P. tricorutum* three types of CAs (Badger et al. 1998; Kaplan and Reinhold 1999; Kikutani et al. 2016; Tsuji et al. 2017b). In combination, the transporters, the CAs and the pyrenoid, may control the intracellular flux of DIC to optimize the CCM efficiency (Giordano et al. 2005; Roberts et al. 2007b; Kroth et al. 2008; Matsuda et al. 2011; Hopkinson et al. 2013; Samukawa et al. 2014; Young and Hopkinson 2017). Biochemical CCMs include the formation of intermediate organic molecules, typically consisting of four carbon atoms, which are subsequently decarboxylated in close proximity to RubisCO to increase the local  $\text{CO}_2$  concentration. Accordingly, this mechanism is also referred to as the C4 pathway or C4-type CCM (Reinfelder 2011). The carbon source ( $\text{CO}_2$  or  $\text{HCO}_3^-$ ), the  $\text{C}_i$  uptake system (solute carrier 4 transporter or external CAs) and the extent to which the  $\text{C}_i$  is actively transported and subsequently accumulated inside the cells can be species-specific and might differ in the same species under different environmental conditions such as high and low  $\text{CO}_2$  concentrations (Tortell 2000; Clement et al. 2017; Tsuji et al. 2017a).

In higher plants, C4 and crassulacean acid metabolism (CAM) photosynthesis evolved independently at least several dozens of times, with adaptations of leaf anatomy, protein function, and photosynthetic efficiency (Leegood 2002;

Sage et al. 2012). The leaf anatomy of the C4 model plant *Zea mays* shows the so-called Kranz anatomy, consisting of large bundle sheath cells which are tightly positioned around the vascular bundle in a wreath-like manner, which is itself embedded in tissue comprising more loosely arranged mesophyll cells (Ray and Black 1976). The two different cell types are specialized to perform different aspects of the C4 pathway. In the cytosol of mesophyll cells,  $\text{C}_i$  is chemically pre-fixed as  $\text{HCO}_3^-$  to phosphoenolpyruvate (PEP) by PEP carboxylase (PEPC) which leads to the formation of oxaloacetic acid (OAA). In the chloroplast of the same cells, the previously formed and highly reactive OAA is reduced to malate which, in turn, is transported out of the mesophyll and into the chloroplasts of the bundle sheath cells, where it is further decarboxylated to release  $\text{CO}_2$ , the substrate for carbon fixation by RubisCO (Leegood 2002). The convergent evolution of C4/CAM photosynthesis results in a huge variety of species-specific pathways, which share some common features like the spatial or temporal separation of carboxylation and decarboxylation, but which also differ in some details such as in the nature of the C4 storage metabolite or in their enzymatic features, e.g., regarding their decarboxylases or C4 molecules. For example, the transient and storage C4 metabolite can be aspartate instead of malate, and the decarboxylase for  $\text{CO}_2$  release can either be a chloroplastic NADP-dependent malic enzyme (ME), a mitochondrial NAD-dependent ME, or a cytosolic PEP carboxykinase (PEPCK) (Leegood 2002; Sage et al. 2012). A curiosity among the C4 photosynthesis plants seems to be *Bienertia cycloptera*, a member of the Chenopodiaceae family, which performs a biochemical CCM, although it is lacking a Kranz-like anatomy. *B. cycloptera* performs C4 photosynthesis within single cells, which consists of two compartments: a central compartment with a similar function like the bundle sheath cells, containing a high density of chloroplasts with well-developed grana, RubisCO enzymes, and starch, and a peripheral compartment, which is functionally similar to the mesophyll cells, containing peripheral chloroplasts with less developed grana, less RubisCO enzymes, and no starch (Voznesenskaya et al. 2002; Edwards and Voznesenskaya 2011).

As mentioned before, the operation of a CCM in phytoplankton, in an aquatic environment, is very likely (Leegood 2002; Giordano et al. 2005). Only a few aquatic plants, e.g., *Egeria densa*, *Elodea Canadensis*, and *Ottelia alismoides*, which display features of a biochemical CCM, and two marine algal species, the macroalga *Udotea flabellum* and the diatom *Thalassiosira weissflogii*, are reported to perform C4 photosynthesis (Leegood 2002; Zhang et al. 2014; Shao et al. 2017). Moreover, based on the missing Kranz anatomy in the filaments of *U. flabellum* and on the fact that most phytoplankton species are single celled, the performance of a C4-like photosynthesis without the spatial separation of  $\text{C}_i$

fixation and release by two different cell types is conceivable for diatoms and other phytoplankton as well (Reiskind and Bowes 1991; Reinfelder et al. 2000). To avoid futile carboxylation and decarboxylation and to guarantee a compartment in which CO<sub>2</sub> can be concentrated for fixation by RubisCO, diatoms might take advantage of their pyrenoid as well as of their four-layered plastids which provide additional compartmentation (Kroth et al. 2008; Tachibana et al. 2011).

As well as the structure of the diatom cells, the localization of their putative C<sub>4</sub> enzymes is essential for this metabolic pathway. The genomic sequences of the multipolar centric diatom *Thalassiosira pseudonana* (Armbrust et al. 2004) and the raphid pennate diatom *P. tricornutum* (Bowler et al. 2008) reveal the presence of several carboxylases and decarboxylases, which potentially could be utilized for a C<sub>4</sub>-like CCM (Kroth et al. 2008). Furthermore, both diatoms possess a pyruvate phosphate dikinase (PPDK), which converts pyruvate into PEP, the initial three carbon molecule that accepts HCO<sub>3</sub><sup>-</sup> in the C<sub>4</sub> CCM of higher plants (Kroth et al. 2008; Sage et al. 2012). However, neither <sup>14</sup>C labeling experiments designed to identify the primary product of the C<sub>3</sub> fixation (Holdsworth and Colbeck 1976), nor PPDK silencing experiments via RNA interference (RNAi) (Haimovich-Dayana et al. 2013) support the presence of a C<sub>4</sub>-like CCM in *P. tricornutum*. However, Kustka et al. (2014) recently suggested that in low CO<sub>2</sub> concentrations, *T. pseudonana* uses a “closed-loop biochemical model,” an atypical C<sub>4</sub>-type CCM, in which the back-reaction of the pyruvate carboxylase (PYC) is assumed to be responsible for the release of CO<sub>2</sub> in the plastid. Young et al. (2016) report a large variation of RubisCO kinetics in different diatoms that may affect the respective CCMs of the algae.

In this work, we present an overview of the presence of enzymes that might be involved in a possible C<sub>4</sub> metabolism in *P. tricornutum*. We investigated the intracellular location of the relevant gene products via expression of GFP-fusion proteins, and discuss the results in comparison to the location of the respective enzymes in the C<sub>4</sub> model plant *Zea mays* and in the C<sub>3</sub> model plant *Arabidopsis thaliana*.

## Materials and methods

### Strains and culture conditions

*Phaeodactylum tricornutum* strain UTEX 642 was cultivated in 30 PSU artificial seawater (ASW) (Harrison et al. 1980), enriched with f/2 nutrition (Guillard and Ryther 1962) under continuous illumination (photosynthetic photon flux density of 50 μmol photons m<sup>-2</sup> s<sup>-1</sup>) at 20 °C and constant aeration with atmospheric air. Solid media contained 1.2% Agar (Wako, Osaka, Japan). Transformation of this strain has been conducted in the Matsuda laboratory in

Sanda, Japan. *P. tricornutum* strain UTEX 646 was grown in 16.6 PSU ASW (Tropic Marine, Dr. Biener GMBH, Wartenberg/Angersbach, Germany) enriched with modified f/2 nutrition (Guillard and Ryther 1962) as described in (Rottberger et al. 2013). Solid media contained 1.2% Bacto Agar (Becton, Dickinson and Company, Le Pont de Claix, France). Cells were cultivated at 20 °C and at 75 μmol photons m<sup>-2</sup> s<sup>-1</sup> continuous light in Erlenmeyer flasks on a shaker. Transformation of this strain was conducted in the Kroth laboratory in Konstanz, Germany.

### Sequence analyses

Protein sequences of *P. tricornutum* have been obtained from the Joint Genome Institute database (*P. tricornutum* 2.0) [<http://genome.jgi-psf.org/Phatr2/Phatr2.home.html>, (Bowler et al. 2008; Grigoriev et al. 2012)] based on their respective protein IDs or on BLAST searches (Altschul et al. 1997). Gene structures, annotations, and putative enzyme localizations have furthermore been checked with help of the programs TargetP 1.1 (Emanuelsson et al. 2000), SignalP 3.0 (Bendtsen et al. 2004), SignalP 4.1 Version (Petersen et al. 2011), TMHMM 2.0 (Sonnhammer et al. 1998), NetStart 1.0 (Pedersen and Nielsen 1997) (all available via the Center for Biological Sequence Analysis (CBS) webserver of the Technical University of Denmark, <http://www.cbs.dtu.dk/index.shtml>) and ASAFind (Gruber et al. 2015) (available at <http://rocaplab.ocean.washington.edu/tools/asafind>), following the methods described in Gruber and Kroth (2014) and Gruber and Kroth (2017).

*Arabidopsis thaliana* sequences were screened in the “tair” (The Arabidopsis Information Resource) database (<http://www.arabidopsis.org>). *Zea mays* proteins were identified in <http://www.uniprot.org> (Consortium 2014), <http://www.maizegdb.org> (Schaeffer et al. 2011) and NCBI (<http://www.ncbi.nlm.nih.gov>) (Altschul et al. 1997).

### Transformation vector and plasmid constructions

Standard cloning procedures were used for plasmid construction (Sambrook et al. 1989). The pPha-T1 *P. tricornutum* transformation vector (GenBank accession AF219942.1) (Zaslavskaia et al. 2000), equipped with a bleomycin-resistance cassette has been used for the gene localization studies. The gene encoding the enhanced green fluorescence protein (egfp) was fused to the 3'-ends of the studied gene sequences according to the cloning strategy described for the PtHLIP2:GFP (and other GFP fusion protein constructs) by Gruber et al. (2007). In case of full length constructs, the stop codons of the target genes were omitted.

## Biolistic transformation

Cells were genetically transformed using the Biolistic PDS-1000/He Particle Delivery System (Bio-Rad, Hercules, CA, USA) fitted with 1350 and 1550 psi rupture disks as described in Zaslavskaja et al. (2000). After transformation, cells were allowed to recover for 24 h before being plated onto *f*/2 medium containing 75–100  $\mu\text{g ml}^{-1}$  Zeocin (Invitrogen, Carlsbad, CA, USA). The plates were incubated at 20–22 °C under constant illumination (50–75  $\mu\text{mol m}^{-2} \text{s}^{-1}$ ). The resulting Zeocin resistant cell lines were screened for GFP fluorescence and the positive ones were further studied. *P. tricornutum* strain UTEX 642 was used to generate the PEPC1full, PEPC2full, PEPCKfull, PPDKfull, RP1full, ME1full, and ME2full transformed cell lines and *P. tricornutum* strain UTEX 646 was used to generate the PYC1pre, PYC2pre, PEPCKpre, PPDKpre, ME1pre, MDH1full, and MDH2full transformed cell lines. The suffixes “pre” and “full” refer to whether the presequence or the full length of the corresponding gene was cloned upstream of the *egfp* gene, respectively.

## MitoTracker labeling

To label mitochondria, the fluorescent stain MitoTracker® OrangeCMTMRos (Molecular probes, Eugene, OR, USA) was added to the cell cultures PEPCKfull, ME1full, and ME2full to a final concentration of 500 nM and to the cultures PYC1pre, MDH1full and MDH2full to a final concentration of 5  $\mu\text{M}$ . (Different MitoTracker concentrations were used because the experiment was performed in two different laboratories.) After the incubation at 20 °C for 60 min, the cells were washed with *f*/2 medium three times and were re-suspended in *f*/2 medium.

## Fluorescence microscopy

Analyses of the transformed cell lines were conducted with laser-scanning confocal microscopes: A1Rsi (Nikon, Japan); LSM 510 META (Zeiss, Germany); LSM 880 (Zeiss, Germany), or with an epifluorescence microscope: Olympus BX51 (Olympus Europe, Germany).

The A1Rsi was equipped with a Plan Apo VC 60x/1.4 Oil DIC objective (Nikon). Using the A1Rsi microscope, the following settings were used: chlorophyll (Chl) *a/c* auto-fluorescence was excited with a 638 nm laser and emission was detected from 662 to 737 nm; excitation of GFP was at 488 nm and fluorescence detection was from 500 to 550 nm emission and the MitoTracker was excited at 561 nm and detected from 570 to 620 nm.

The LSM 510 META was operated using a Plan-Apochromat 63  $\times$ /1.4 Oil DIC objective. GFP and Chl fluorescence was excited at 488 nm, filtered with a beam

splitter (HFT 488/543), and detected by two different photomultiplier tubes with a band pass filter (BP 505–530) for GFP fluorescence and a low pass filter (LP 650) for Chl auto-fluorescence. Maximum intensity *z*-projections were calculated from slices of image stacks to ensure complete detection of fluorochromes within a cell.

The LSM 880 was operated using a Plan-Apochromat 63  $\times$ /1.4 Oil DIC M27 objective; GaAsP photomultiplier. GFP and Chl were excited at 488 nm (MBS 488/543/633) and detected at 500–553 nm and 695–760 nm, respectively. MitoTracker was excited at 514 nm (MBS 458/514), fluorescence was detected at 570–624 nm.

The Olympus BX51 was equipped with a Zeiss AxioCam MRm digital camera system (Zeiss). To view transmitted light images (100xUplanFL objective, Olympus), Normarski's differential interference contrast illumination was used; for Chl auto-fluorescence, the mirror unit U-MWSG2 (Olympus) was used; and for GFP detection, the filter set 41020 (Chroma Technology Corp, Bellows Falls, VT, USA).

## Determination of photosynthetic parameters

Cells grown under atmospheric air conditions were harvested by centrifugation at the mid-logarithmic phase and washed with DIC-free *f*/2 medium. Cells were suspended in the DIC-free buffer at a Chl *a* concentration of 10  $\mu\text{g ml}^{-1}$ . The rate of photosynthetic  $\text{O}_2$  evolution at various concentration of DIC was measured with a Clark-type oxygen electrode (Birmingham and Colman 1979; Matsuda et al. 2001). The concentration of DIC at the  $\text{CO}_2$  compensation point was measured by gas chromatography according to the method described by Birmingham and Colman (1979). The maximum photosynthetic capacity ( $P_{\text{max}}$ ), the DIC concentration which gives half  $P_{\text{max}}$  ( $K_{0.5}[\text{DIC}]$ ), and the apparent photosynthetic conductance (APC) values were determined by curve fitting with the non-linear least square method.

For the PEPCK inhibition, 3-mercaptopicolinic acid (3-MPA; Toronto Research Chemicals Inc., ON, Canada) was dissolved in *f*/2 medium and added at various concentrations 10 min before the  $\text{O}_2$  electrode assay. For the PEPC inhibition, a PEP analog 3,3-dichloro-2-(dihydroxyphosphinoylmethyl)-propionate (DCDP) was dissolved in 10 mM Tris-HCl (pH 7.5), and various concentrations of DCDP were added to cells in the presence of 0.0075% (v/w) Triton X-100. Triton X-100 was used to insure the permeation of DCDP into the cells. To minimize the toxicity of Triton X-100 to the cells, cells were immediately centrifuged at 5000 $\times g$  for 5 min and resuspended in the DIC-free *f*/2 medium. This washing step was repeated three times and washed cells in DIC-free *f*/2 medium were used for the  $\text{O}_2$  electrode assay. The rate of  $\text{O}_2$  evolution at various DIC concentrations was measured by injecting a known amount of  $\text{NaHCO}_3$  as substrate. The maximum

photosynthetic O<sub>2</sub> evolution rate, the half-saturation constant for DIC and apparent photosynthetic conductance were determined by rectangular-hyperbola fitting to data with non-linear least squares regression.

## Results and discussion

### From in silico to in vivo—the subcellular localization of putative C4 photosynthesis enzymes

Before experimentally confirming the subcellular locations of the enzymes that might be involved in a biochemical CCM pathway, we checked the existing gene models and updated them where necessary. This is a crucial part of the work because the targeting of the protein depends on N-terminal pre-sequences, which are not always part of an automatically created gene model. Proteins which carry a mitochondrial targeting sequence are transported into the mitochondrion, proteins with a bipartite pre-sequence enter the secretory pathway where they are then transferred into either the endoplasmic reticulum (ER), the plastid or out of the cell. Proteins without a special pre-sequence stay in the cytosol (Kilian and Kroth 2005; Bolender et al. 2008; Gruber and Kroth 2017). Based on our thorough examination of the gene models, we fused gene fragments encoding either the predicted pre-sequences or the complete genes to the egfp gene, and expressed the constructs in *P. tricornutum* (see Table 1). In many proteins, the pre-sequence alone is sufficient for targeting of the protein of interest (Apt et al.

2002) and we benefit from easier handling of shorter gene fragments during the process of creating transformants. Using full sequences has the advantage that internal targeting signals, which are needed e.g., to anchor the protein, are considered (Bolender et al. 2008).

Experiments, including cloning, transformation, and microscopy work, were performed on two different *P. tricornutum* strains in two different laboratories. Even under independent experimental setup and regardless of the cloned fragment length, the localization of PEPCK, PPK, and ME1 gave identical results (for images see Fig. 1 & Online Resource Fig. 5). To activate the light-dependent fcpA promoter which drives egfp expression, the transformed cell lines were cultivated at 75 µE continuous light. Fluorescence microscopy allowed visual allocation of the proteins to the intracellular compartments including mitochondria (PEPC2, PYC1, PEPCK, ME1, ME2, MDH1, MDH2), plastids (PEPC1, PYC2, RP1) and the cytosol (PPDK) according to the observed fluorescence in comparison with the plastid auto-fluorescence, transmitted light views, and MitoTracker stainings (Fig. 1 & Online Resource Fig. 5).

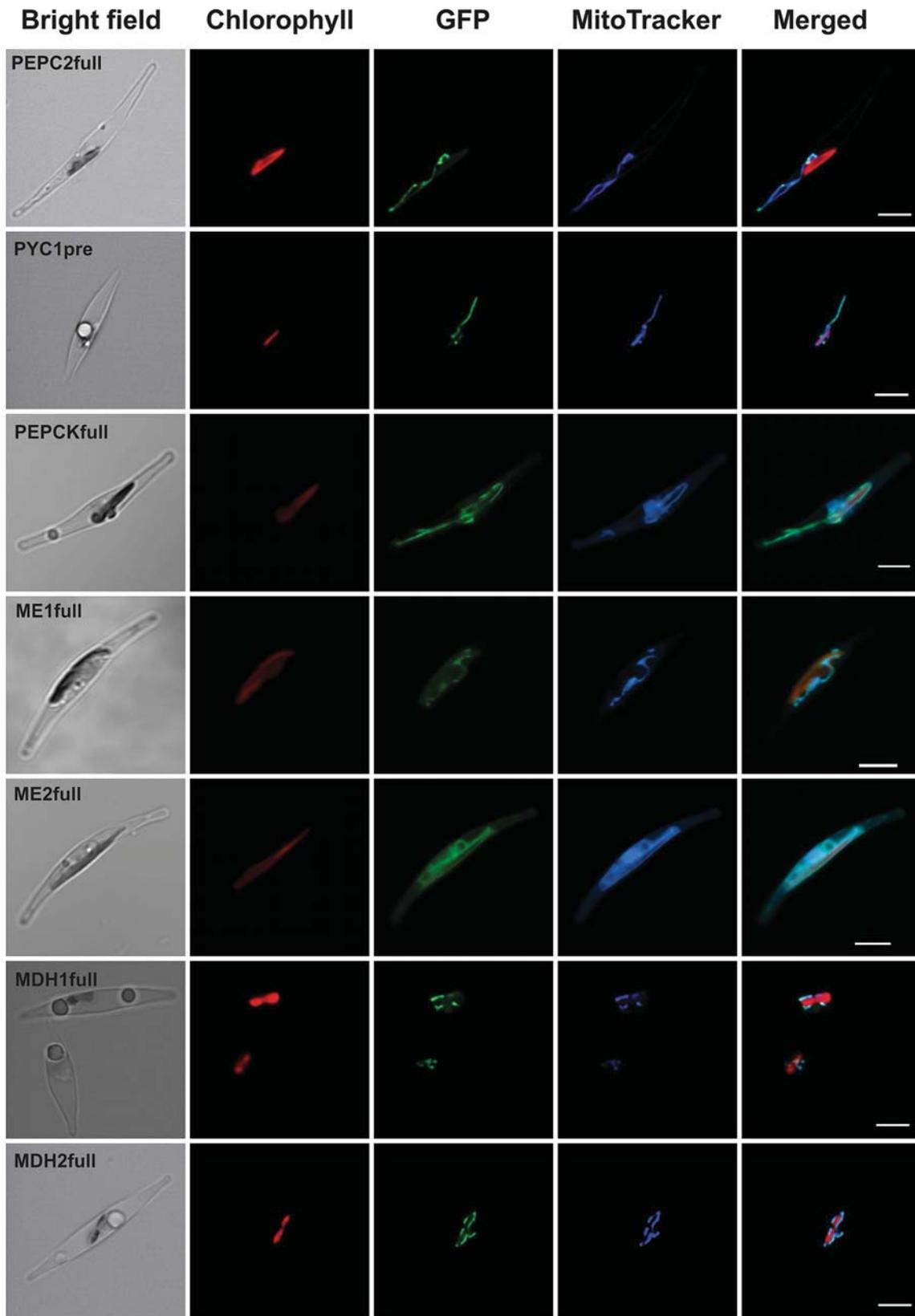
### Single cell C4 photosynthesis—advantages and disadvantages of diatoms

With 22 described types of Kranz anatomy, this structure seems to be a specific and common feature of C4 photosynthesis in higher plants (Leegood 2002; Sage et al. 2011). In unicellular organisms like diatoms, such a differentiation in two cell types is obviously not possible. Also a system

**Table 1** Putative C4 CCM proteins and their localizations in *P. tricornutum*

Protein name	Protein ID (JGI)	Cloned amino acid sequence	Transformed strain	GFP localization
PEPCK (phosphoenolpyruvate carboxykinase)	55018	41 aa	UTEX 646	Mitochondrion
		Full length	UTEX 642	Mitochondrion
ME1 (NAD-dependent malic enzyme 1)	56501	94 aa	UTEX 646	Mitochondrion
		Full length	UTEX 642	Mitochondrion
ME2 (NADP-dependent malic enzyme 2)	27477	Full length	UTEX 642	Mitochondrion
PEPC1 (phosphoenolpyruvate carboxylase 1)	56768	Full length	UTEX 642	Plastidic membranes
PEPC2 (phosphoenolpyruvate carboxylase 2)	56769	Full length	UTEX 642	Mitochondrion
PYC1 (pyruvate carboxylase 1)	30519	55 aa	UTEX 646	Mitochondrion
PYC2 (pyruvate carboxylase 2)	49339	67 aa	UTEX 646	Stroma
PPDK (pyruvate phosphate dikinase)	21988	112 aa	UTEX 646	Cytosol
		Full length	UTEX 642	Cytosol
RP1-PPDK (PPDK regulator protein 1)	49027	Full length	UTEX 642	Stroma
MDH1 (malate dehydrogenase 1)	51297	Full length	UTEX 646	Mitochondrion
MDH2 (malate dehydrogenase 2)	54834	Full length	UTEX 646	Mitochondrion

Each gene which was experimentally localized by GFP fusion experiments is listed above. The column “cloned amino acid sequence” indicates whether only a fragment encoding an N-terminal part (given as amino acid length) or the full length gene had been cloned to the 5'-end of the egfp gene. Column “transformed strain” indicates whether strain UTEX 642 (in the Matsuda Laboratory) or UTEX 646 (in the Kroth Laboratory) was transformed. (Images showing GFP localization experiments can be found in Fig. 1 and in Online Resource Fig. 5.)



◀ **Fig. 1** Images of proteins which have been fused to GFP for localization analyses in this study. Presented are fluorescence microscopy images showing transformed *P. tricornutum* cells with expressed protein::GFP fusion proteins of all investigated proteins which might be involved in a putative C<sub>4</sub> CCM. Bright field, chlorophyll auto-fluorescence (red), GFP fluorescence (green), and a merged image, showing an overlap of the chlorophyll auto- and GFP fluorescence, are shown from left to right. The transformed cell lines which express PYC1pre::GFP, PEPCKfull::GFP, ME1full::GFP, ME2full::GFP, MDH1full::GFP, and MDH2full::GFP constructs have additionally been stained with MitoTracker (blue) and their according merged images are an overlap of chlorophyll auto-fluorescence, GFP fluorescence, and MitoTracker fluorescence. Scale bars of each image represent 5  $\mu$ m. “full” or “pre” suffixes after the enzyme’s name indicate whether a full length or pre-sequence has been used for plasmid construction. PEPC2, PYC1, PEPCK, ME1, ME2, MDH1, and MDH2 are located in the mitochondrion; PYC2 and RP1 are located in the plastid’s stroma and PPDK was detected in the cytosol. *P. tricornutum* strain UTEX 642 has been used to generate the PEPC1full, PEP-C2full, PEPCKfull, RP1full, ME1full, and ME2full transformed cell lines and *P. tricornutum* strain UTEX 646 has been used to generate the PYC1pre, PYC2pre, PPDKpre, MDH1full, and MDH2full transformed cell lines. *PEPC* phosphoenolpyruvate carboxylase, *PYC* pyruvate carboxylase, *PEPCK* phosphoenolpyruvate carboxykinase, *ME* malic enzyme, *MDH* malate dehydrogenase, *PPDK* pyruvate phosphate dikinase, *RP* PPDK regulator protein

like in the flowering plant *B. cycloptera*, which misses the Kranz anatomy and compensates for it with two specialized types of chloroplasts (Voznesenskaya et al. 2002), cannot be expected in diatoms, because diatoms usually only possess one type of plastid (and in many species just one plastid per cell). In order to separate the process of C<sub>i</sub> pre-fixation by PEPC and/or PEPCK from the decarboxylation of the C<sub>4</sub> acid, so that the released CO<sub>2</sub> is available to RubisCO, diatoms might utilize the additional compartments provided by their secondary plastids and pyrenoids (McFadden 2001; Kroth et al. 2008; Tachibana et al. 2011). The plastid in diatoms is surrounded by four membranes which separate the endoplasmic reticulum (ER) lumen, periplastidic space (PPS), inter envelope space (IES), and stroma from each other (Kilian and Kroth 2005) (Fig. 3). The stroma itself hosts the pyrenoid which again might contain a sub-pyrenoid structure as has been shown in *P. tricornutum* (Tachibana et al. 2011). If these additional compartments are used for C<sub>4</sub> photosynthesis, the localization of all enzymes which are involved in this pathway will help to answer the question how it can be performed successfully without futile cycling of carboxylation and decarboxylation reactions.

Our analyses revealed that the intracellular distribution of the carboxylases and decarboxylases differs substantially between *P. tricornutum* and the two land plants *A. thaliana* (C<sub>3</sub> model plant) and *Z. mays* (C<sub>4</sub> model plant) (Fig. 2 & Online Resource Table IV). C<sub>i</sub> pre-fixation in the diatom occurs exclusively in the organelles whereas in higher plants it takes place solely in the cytosol. The decarboxylases in *Z. mays* have cytosolic and chloroplastic decarboxylases whereas

the putative C<sub>4</sub> decarboxylases of *P. tricornutum* are located in the mitochondria. In the following, we discuss these differences for the various enzymes in more detail.

### Phosphoenolpyruvate carboxylases (PEPCs)—neither PEPC1 nor PEPC2 pre-fix C<sub>i</sub> in the cytosol of *P. tricornutum* cells

The analysis of pre-sequences predicts that the two isoenzymes of PEPC in *P. tricornutum* are targeted to different compartments (Kroth et al. 2008). While the expression of the respective GFP fusion proteins confirmed the location of PEPC1 to be associated to the plastid membranes, the localization of PEPC2 is mitochondrial and not cytosolic as expected from the sequence analysis (Fig. 1 & Online Resource Table V). For PEPC1, the exact location within the four-layered plastid envelope system could not be resolved by microscopy; however, the GFP expressing cell lines strongly resembled a phenotype that has previously been described as “blob”-like structures (BLS) (Kilian and Kroth 2005; Gruber et al. 2007) and is proposed to represent the periplastidic space in diatoms (Gould et al. 2006; Sommer et al. 2007; Gruber et al. 2009; Weber et al. 2009; Moog et al. 2011).

The PEPCs in *A. thaliana* and *Z. mays* are located in the cytosol (Fig. 2 & Online Resource Table IV). In *A. thaliana*, PEPCs are involved in amino acid synthesis, generating OAA which enters the tricarboxylic acid (TCA) cycle forming  $\alpha$ -ketoglutarate, the precursor of glutamate (Taylor et al. 2010). Glutamate and its derivative glutamine are synthesized in the chloroplasts of *A. thaliana* and are known as transport compounds for nitrogen (Taylor et al. 2010). Based on its mitochondrial localization, Pt\_PEPC2 could be involved in a similar pathway because the diatom’s glutamate synthase (GLU) (ID Pt\_56605) as well as its glutamine synthetase (GLN, ID Pt\_51092), required for the conversion of glutamate to glutamine, are predicted to be located in the diatom plastid, while in *A. thaliana* the glutamine synthetase is most likely cytosolic (Online Resource Table V). If  $\alpha$ -ketoglutarate from the TCA cycle should be required as a precursor for amino acid synthesis in *P. tricornutum*, it would have to be transported from the mitochondria to the plastid stroma across six membranes.

Therefore, we propose an alternative pathway for *P. tricornutum*: pyruvate, deriving from cytosolic glycolysis, may be transported into the plastid, where it could be carboxylated to OAA via the plastidic PYC2 (Fig. 3). OAA may then be converted into  $\alpha$ -ketoglutarate, a substrate for the glutamate synthase, by the aspartate aminotransferase 4 (AAT4; please see predictions of AAT4 in the Online Resource Table V & Fig. 3).

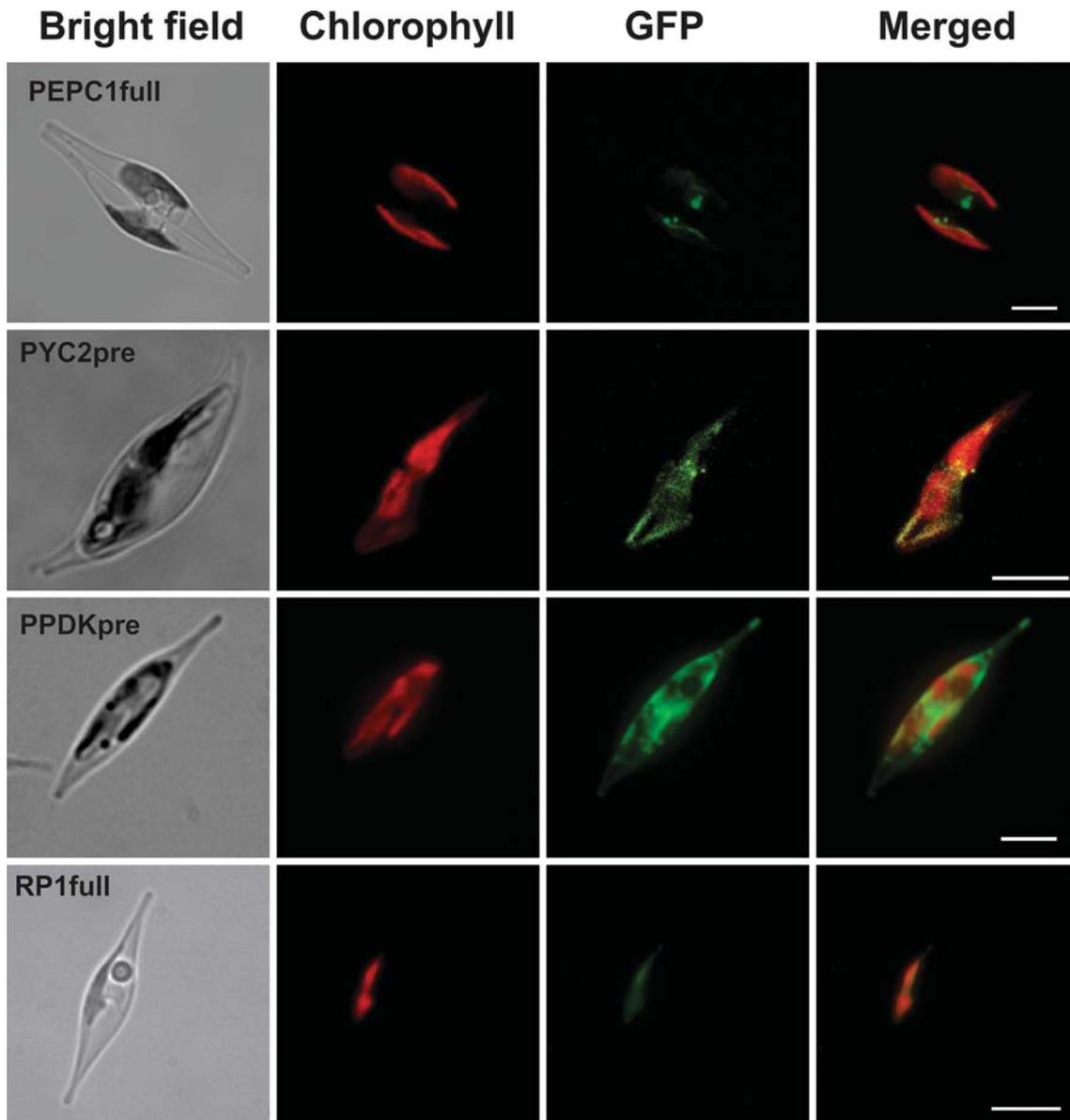
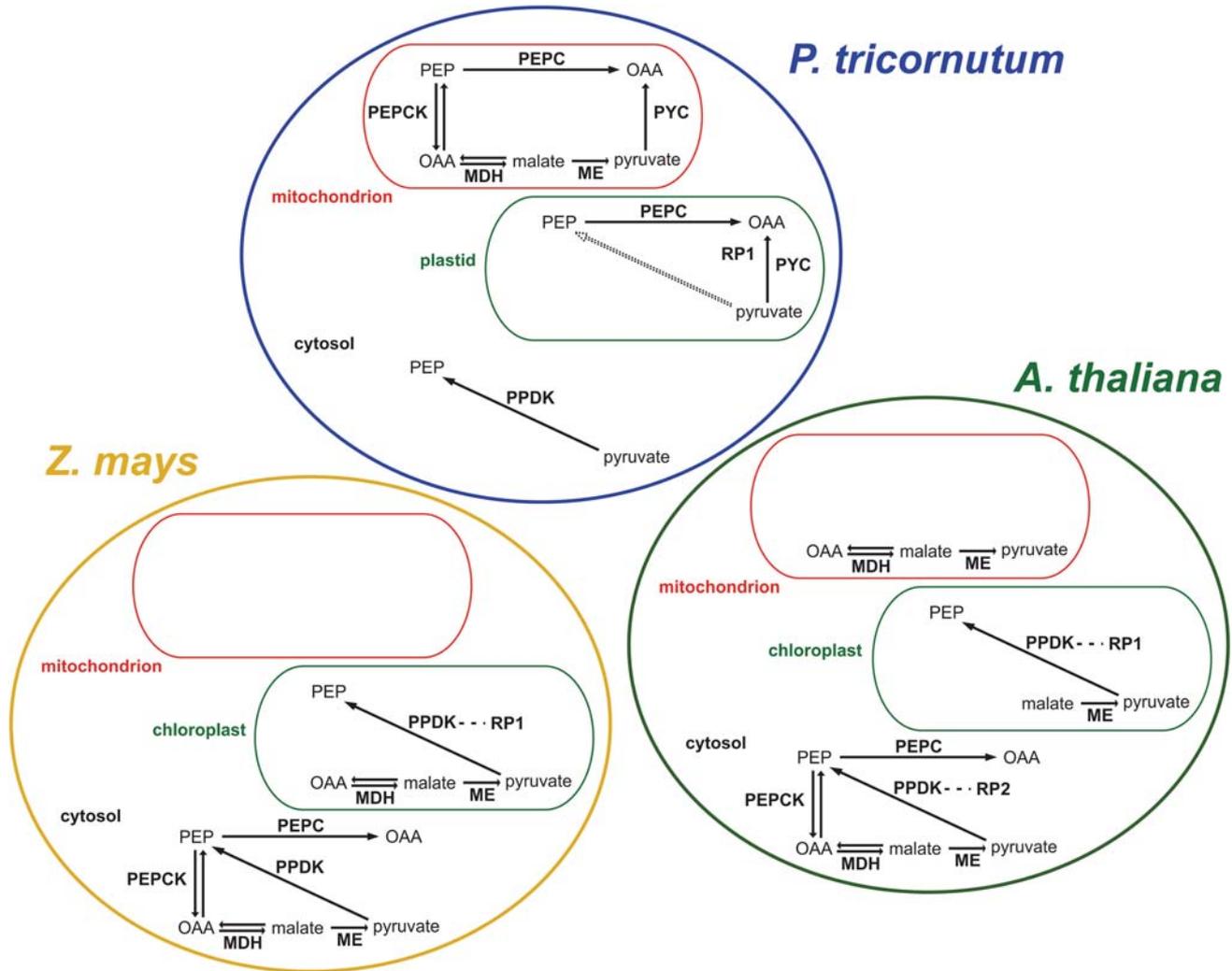


Fig. 1 (continued)

### Pyruvate carboxylases (PYCs)—the mitochondrial PYC1 and plastidic PYC2 might supply the TCA cycle and amino acid synthesis with OAA, respectively

PYCs catalyze the carboxylation of pyruvate and bicarbonate under the consumption of ATP to form OAA. Two isoforms of the PYCs have been identified in *P. tricornutum*, PYC1 is located in the mitochondria, while PYC2 is apparently a stromal protein (Fig. 1). PYCs are neither found in cyanobacteria nor in higher plants (Tsuji et al. 2012), so their presence in diatoms is remarkable. Genes encoding PYC enzymes have been detected in different diatoms and in related algae (Heterokonts, Haptophytes) (Appleby et al.

1980; Descolas-Gros and Oriol 1992; Tsuji et al. 2012). Studies in the yeast *Saccharomyces cerevisiae* (Stucka et al. 1991) and the Haptophyte *Emiliania huxleyi* (Tsuji et al. 2012) revealed an anaplerotic function of PYC because it supplies the TCA cycle with OAA. Tsuji et al. (2012) report an inhibition of the plastidic Eh\_PYC1 activity by L-malate, the reduced form of the PYC-product OAA, indicating a negative feedback loop of the enzyme activity. As aspartate, a C4 molecule derived from OAA by aspartate aminotransferase (AAT), has a much weaker inhibitory influence on PYC, it might more likely be involved in the accumulation of C4 molecules (Tsuji et al. 2012). This indicates that in



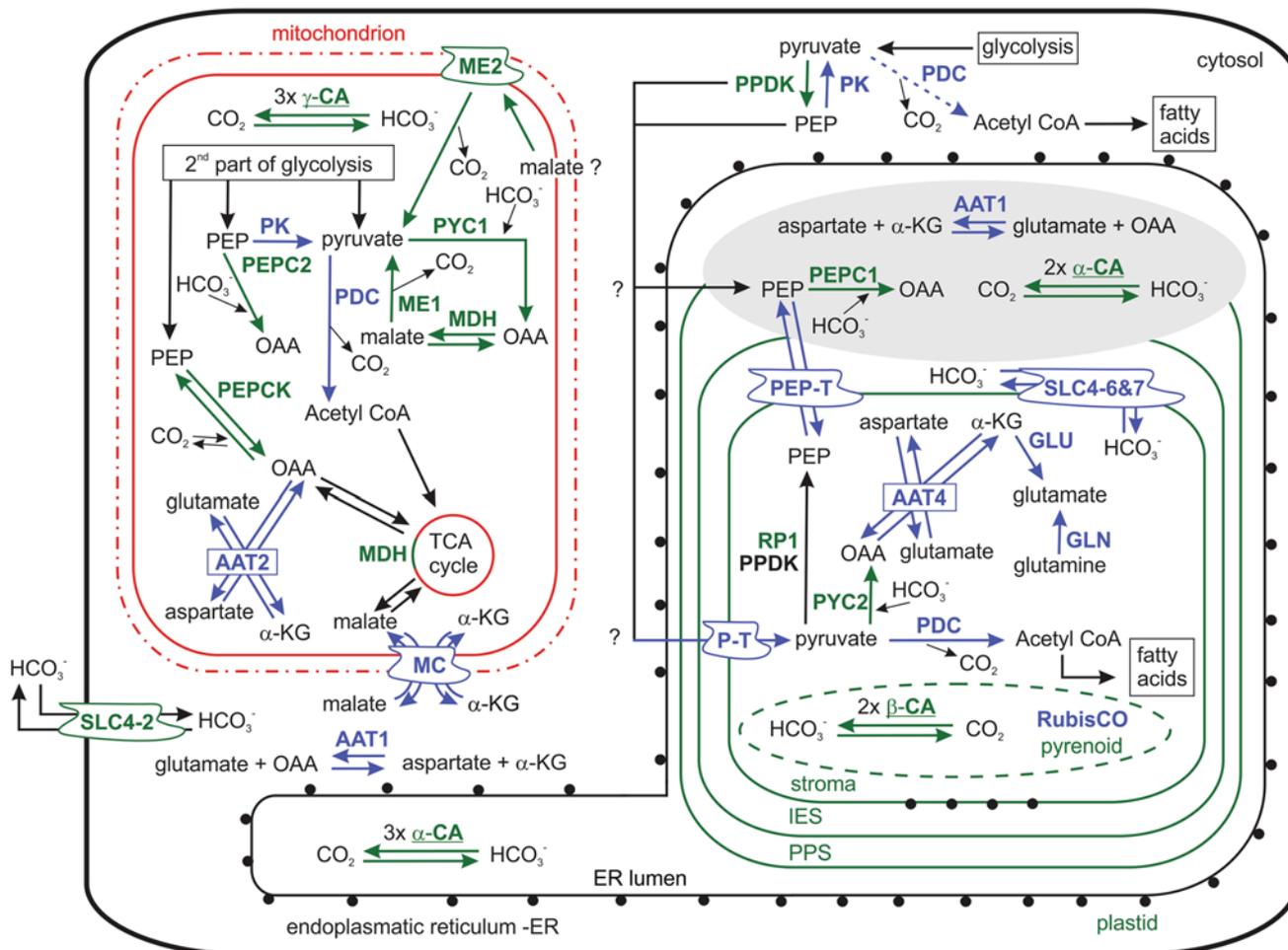
**Fig. 2** Localization of putative C4 proteins in *P. tricornutum*, *A. thaliana*, and *Z. mays*. The diagram shows a simplified cellular distribution of the investigated C4 enzymes in the C4 plant *Z. mays*, the C3 plant *A. thaliana* and in the diatom *P. tricornutum*, respectively. The distribution of the decarboxylases ME and PEPCCK differs between the plants and the diatom. While *Z. mays* and *A. thaliana* each contain a chloroplastic decarboxylase, the plastid of *P. tricornutum* con-

tains no decarboxylase. Enzymes (in bold): MDH: malate dehydrogenase; ME: malic enzyme; PEPC: phosphoenolpyruvate carboxylase; PPK: pyruvate phosphate dikinase; PYC: pyruvate carboxylase; RP1(2): PPK regulator protein 1(2). Substrates and products: PEP phosphoenolpyruvate, OAA oxaloacetic acid. A detailed overview of the protein IDs and their experimentally demonstrated localizations can be found in Table 1 and in Online Resource Table IV

*E. huxleyi* PYC might be involved in amino acid synthesis rather than in a potential C4 metabolism.

Screening the *P. tricornutum* genome, we could not identify any gene encoding a plastidic malate dehydrogenase (MDH) isoenzyme (Online Resource Table V), suggesting that OAA cannot be converted into malate via this reaction within *P. tricornutum* plastids. Accordingly, there might be no or little malate in the plastid to inhibit Pt\_PYC2 (Fig. 3). This makes Pt\_PYC2 a good candidate for the production of OAA, a precursor for amino acid synthesis. In C3 plants, MDHs play a major role in the redox regulation between the chloroplast and the cytoplasm (Heber 1974; Scheibe

1987; Buchanan 1991). Considering the apparent absence of MDHs in the plastid and in the cytosol of *P. tricornutum*, it is not clear yet how this essential reaction, the interconversion from OAA and  $\text{NAD(P)H} + \text{H}^+$  into malate and  $\text{NAD(P)}^+$  and vice versa, can be performed or whether it is required at all. According to Bowes (2011) the submerged and facultative C4 plant *Hydrilla verticillata*, which performs the complete C4 photosynthesis in a single cell, possesses two MDH isoforms, one is located in the chloroplast and one in the cytosol. Whether or not the enzymes in the C3 photosynthesis leaves contribute to redox regulation remains elusive.



**Fig. 3** Enzyme composition in *P. tricornutum*. The scheme describes the putative intracellular localization of metabolic enzymes in the diatom *P. tricornutum*. The red box represents the mitochondrion; green lines surround the plastid, and the black line represents the endoplasmic reticulum which is studded with ribosomes (black dots). Enzymes shown in green: the enzyme location has been confirmed by localization experiments; enzymes shown in blue: the enzyme location has been predicted; underlined enzyme names: some of the isoforms have been located experimentally, while others are only predicted; dotted arrow: the gene of one of the subunits cannot be found in the genome so far; gray oval: the enzymes in this oval are located either in the IES, PPS, or ER lumen. Enzymes: AAT aspartate aminotransferase, CA carbonic anhydrase, GLN glutamine synthetase, GLU glutamate synthase, MDH malate dehydrogenase, ME1 malic

enzyme (NAD dependent), ME2 malic enzyme (NADP dependent), PEPC PEP carboxylase, PDC pyruvate dehydrogenase complex, PK pyruvate kinase, PPK pyruvate phosphate dikinase, PYC pyruvate carboxylase, RP1 PPK regulator protein 1. Substrates and products:  $\alpha$ -KG  $\alpha$ -ketoglutarate, PEP phosphoenolpyruvate, OAA oxaloacetic acid. Transporter: MC malate/ $\alpha$ -ketoglutarate carrier, P-T pyruvate transporter, PEP-T PEP transporter, SLC4 solute carrier 4 family  $\text{HCO}_3^-$  transporter. The direction of the transported substrates, indicated by arrows, is proposed but not proven yet. Compartments: IEP inter envelope space, PPS periplastidic space. All enzymes and transporters including protein IDs and their references are listed in Online Resource Table IV. Please note that in order to keep a clear arrangement of the figure, cofactors or reducing agents are not shown

Another aspect supporting a further role of PYC is based on  $^{14}\text{C}$  labeling experiments in *E. huxleyi* (Tsuji et al. 2009). Here like in *P. tricornutum* (Holdsworth and Colbeck 1976), the first initial products of  $^{14}\text{C}$  labeled carbon metabolites are phosphate esters, followed by lipids, aspartate, alanine, and glutamate. Aspartate concentrations, however, remain constant over time, which can be interpreted as an indication for a non-C4 CCM fixation mechanism. Since Eh\_PYC1 (GenBank BAH22705) and Pt\_PYC2 are phylogenetically related (Tanaka et al. 2014) and as both are located in the

plastid, we assume that the PYC2 enzyme in *P. tricornutum* may also be involved in amino acid synthesis.

In addition to the replenishment of the TCA cycle and amino acid synthesis, PYCs are also known to be involved in gluconeogenesis. In *P. tricornutum*, gluconeogenesis might be used to shuttle fixed  $\text{C}_i$  to produce chrysolaminarin via the nucleotide sugar uracil-diphosphate glucose (UDP-Glc) (Chauton et al. 2013). So far, the intracellular location of the different steps of chrysolaminarin biosynthesis in diatoms is unknown.

A different hypothesis has been put forward recently by Kustka et al. (2014). They investigated the carbon metabolism of *T. pseudonana* under low CO<sub>2</sub> conditions, demonstrating a 1.8-fold increase in PYC abundance via proteomic analysis. They furthermore postulate that PYC in *T. pseudonana* might predominantly catalyze the decarboxylation of OAA, which however according to their calculation would require maintenance of a neutral pH, an ATP/ADP ratio of 2.5 and an OAA concentration > 1 mM in the stroma. Kustka et al. (2014) do not provide details on the calculation and especially do not mention the maximum tolerated concentration of pyruvate which would still allow the proposed “backwards” reaction. We think that the three requirements mentioned by Kustka et al. (2014) are pretty unrealistic to be assumed for a plastid stroma. Furthermore, pyruvate accumulation in the stroma would need to be prevented by consecutive reaction or transport steps, but there is no evidence for such processes so far. Therefore, we think that it is highly unlikely that decarboxylation of OAA by PYC contributes to photosynthetic carbon fixation in *T. pseudonana*. This is also in line with various studies on carbon fixation in *T. pseudonana*, including isotope labeling experiments (Roberts et al. 2007a), studies of kinetic properties of C4 enzymes (Clement et al. 2016) and localization of C4 related enzymes (Tanaka et al. 2014), which all do not support the presence of a biochemical CCM in this organism.

### **Pyruvate phosphate dikinase (PPDK) and its regulator protein 1 (RP1)—a plastid pre-sequence but cytosolic GFP fluorescence—is PPDK dual targeted?**

PPDK converts pyruvate into PEP (Parsley and Hibberd 2006), which can either be provided to gluconeogenesis or to bicarbonate pre-fixation as a part of a biochemical CCM. PEP is the essential receptor for bicarbonate in the C4 metabolism, and therefore, the discovery of a PPDK in *P. tricornutum* initiated speculations about C4 CCM in this alga (Kroth et al. 2008). PPDK RNAi silencing transformants of *P. tricornutum* showed a clearly reduced PPDK activity while photosynthetic activity under low HCO<sub>3</sub><sup>-</sup> conditions was not affected (Haimovich-Dayana et al. 2013).

Both *A. thaliana* and *Z. mays* possess gene copies of the *ppdk* gene (Online Resource Table IV) which contain two alternative start codons, separated by an intron, resulting in the formation of two different mRNAs (Parsley and Hibberd 2006). One mRNA encodes a protein which carries a transit sequence for chloroplast targeting, while the second protein has no signal sequence and therefore is located in the cytosol (Parsley and Hibberd 2006). This dual targeting mechanism is also known for the *ppdk* genes of *Oryza sativa* (accession number D87745) and *Flaveria trinervia* (accession number X79095) (Parsley and Hibberd 2006). The location of

the PPDK in *P. tricornutum* is unclear; the full length protein has a plastid targeting pre-sequence, while expression of PPDK::GFP fusion proteins indicate a cytosolic and not plastidic location of the enzyme (Fig. 1 & Online Resource Fig. 5; please notice the auto-fluorescence of the chlorophyll which represents the plastid in the “Merged” column). This ambiguous result could be explained by a similar effect in *A. thaliana*. The regular start codon is preceded by additional in-frame ATGs (pos. 71, 77, 100), possibly initiating an alternative translation of the proteins that are not imported into the plastid. However, the absence of any GFP fluorescence within the plastids rather supports the notion that the PPDK in *P. tricornutum* is a cytosolic protein that could be involved in gluconeogenesis. Interestingly, we could identify a single gene in *P. tricornutum* encoding a protein related to a PPDK regulator protein (RP1). Expression of an RP1::GFP fusion protein indicated that this protein is located in the plastid (Fig. 1 & Online Resource Table V). The light-dependent PPDK-RP in *A. thaliana* activates the PPDK in the light and inactivates it in the dark by reversible phosphorylation of PPDK (Chastain et al. 2011). Here, two isoforms of PPDK regulator proteins have been found, while RP1 is located in the chloroplast, RP2 is found in the cytosol (Chastain et al. 2008). Accordingly, in the C4 plant *Z. mays*, only a chloroplastic regulator protein has been found (Smith et al. 1994), even though the second *ppdk* splice variant is leading to a cytosolic isoform (Sheen and Bogorad 1987; Sheen 1991; Parsley and Hibberd 2006). There are two putative explanations for this: (i) cytosolic PPDK regulator proteins may not have been identified yet in *Z. mays* and *P. tricornutum* or (ii) the cytosolic PPDK activity is not modulated by a RP. In the latter case, however, it is unclear how the PPDK is regulated, how the direction of the reversible reaction is controlled, and how the activity rate is adjusted to changing conditions. According to Jenkins and Hatch (1985), the direction of the reversible reaction of C4 PPDK from maize is driven by pH. The assumed pH in the cytosol of single-celled phytoplankton, which, according to Chastain et al. (2011), drives the reaction of PPDK into pyruvate formation, is about 7.2 [*T. weissflogii*: pH 7.35 (Herve et al. 2012); *Chlamydomonas reinhardtii*: pH 7.4 ± 0.1 (Braun and Hegemann 1999); *E. huxleyi*: pH 6.9 ± 0.3 and *Coccolithus pelagicus*: pH 7.0 ± 0.3 (Anning et al. 1996)]. It is not mentioned whether the pH is driving the PPDK activity directly or via the regulator protein.

### **Phosphoenolpyruvate carboxykinase (PEPCK) might be involved in gluconeogenesis**

PEPCK utilizes CO<sub>2</sub> as a substrate for carboxylation, while in CAM plants, the enzyme may also function as a decarboxylase, converting OAA to PEP (Dittrich et al. 1973; Häusler et al. 2002). A high PEPCK activity in *A. thaliana*

might cause CO<sub>2</sub> provision to photosynthesis (Brown et al. 2010). While the PEPCKs of *Z. mays* and *A. thaliana* are located in the cytosol (Fig. 2 & Online Resource Table IV), the respective enzyme of *P. tricornutum* is located in the mitochondria. GFP localization studies based on two different *P. tricornutum* strains and two independently cloned plasmids confirm the result of the sequence analysis and show a mitochondrial localization (Fig. 1, Online Resource Fig. 5 & Table V). This interesting finding made us curious so we also identified all enzymes needed for the first part of gluconeogenesis (PYC, PEPCK, enolase 2, phosphoglycerate mutase 2, phosphoglycerate kinase 3, glyceraldehyde 3 phosphate dehydrogenase/triosephosphate isomerase fusion protein). In *P. tricornutum*, all of these proteins, including the PEPCK, are predicted to be located in the mitochondria (Online Resource Table V). The subsequent enzymatic reactions of gluconeogenesis, namely the formation of the C6 sugar fructose-1,6-bisphosphate, are predicted to occur either in the cytosol or in the plastid (Online Resource Table V). These results might indicate that the Pt\_PEPCK contributes rather to the initial steps of gluconeogenesis than to a potential CCM.

### Malic enzymes (ME) and the co-localized malate dehydrogenases (MDH) might serve to replenish the TCA cycle

MEs decarboxylate malate into pyruvate and CO<sub>2</sub> using NAD<sup>+</sup> or NADP<sup>+</sup> as a cofactor and they are often located in the same compartment as the malate dehydrogenases (Fig. 2). Both types of MEs catalyze the formation of pyruvate from OAA. NADP-dependent MEs of eukaryotes and prokaryotes are involved in at least two metabolic pathways: as providers of reduction equivalents (NADPH) in lipid biosynthesis (Lai et al. 2002; Hao et al. 2014) and as providers of CO<sub>2</sub> for RubisCO in C4 CCMs (Jenkins et al. 1987a).

According to the localization of GFP fusion proteins, MEs and MDHs of *P. tricornutum* are strictly mitochondrial proteins (Fig. 1 & Online Resource Fig. 5). Considering that CO<sub>2</sub> released by these two enzymes would have to diffuse across six membranes to reach RubisCO in the plastid, with a number of carbonic anhydrases converting it into bicarbonate, it seems rather unlikely that these enzymes might be involved in a biochemical CCM (Fig. 3). Wang et al. (2014) state that even in land plants, having chloroplasts with two envelope membranes, a functional C4 pathway strictly requires a chloroplastic decarboxylase.

Another role of the MEs in *P. tricornutum* could be the provision of NADPH for the elongation of the fatty acid acyl chain in lipid biosynthesis, as has already been shown in microorganisms (Wynn and Ratledge 1997; Ratledge and Wynn 2002; Zhang et al. 2007; Beopoulos et al. 2009). Wynn et al. (1997) revealed that the inhibition of

ME activity in *Mucor circinelloides* (oleaginous fungus) reduces the lipid content from 24 to 2% w/w dry weight. This has been supported by the experiments of Zhang et al. (2007) who showed that the overexpression of the ME gene in *M. circinelloides* led to a two- to threefold increase of ME activity as well as a 2.5-fold increase in lipid content. Both experiments show that ME is a key player in lipid synthesis in fungi. In plants, lipid biosynthesis is initiated by the chloroplastic Acetyl CoA Carboxylase (ACC), and elongation occurs via the cytosolic ACC (Nikolau et al. 2003). We found that *P. tricornutum* possesses both isoforms of ACCs, which are localized in the same subcellular compartments as in plants (Online Resource Table V). However, in *P. tricornutum* the MEs are strictly localized in the mitochondria, making it unlikely that they provide NADPH for the plastidic or cytosolic ACCs.

Of course, we cannot exclude the possibility that there might be a plastidic and/or cytosolic ME isoform in *P. tricornutum*, but we can confirm that the gene models we used (Pt\_56501 & Pt\_27477) are correct to our best knowledge. In previous studies, the presence of a cytosolic (Pt\_54082) as well as plastidic (Pt\_51970) ME in *P. tricornutum* has been reported (Yang et al. 2013; Davis et al. 2017); however, comparison of the gene models revealed that the predicted locations of both MEs which are mentioned in Davis et al. (2017) and Yang et al. (2013) are most likely incorrect because the gene models they used were truncated. Pt\_54082 & Pt\_56501 as well as Pt\_51970 & Pt\_27477 refer to the same loci on the chromosomes and hence describe the same proteins. Pt\_54082 and Pt\_51970 are truncated at the N-terminus by 78 and 67 amino acids, respectively. In diatoms, the N-terminus of a protein may carry a bipartite pre-sequence to target a nucleus encoded protein into the plastid (Kilian and Kroth 2005) and such a targeting pre-sequence is not always included in homology-based gene models.

After excluding the two possibilities that MEs in *P. tricornutum* contribute substantially to a biochemical CCM or to NADPH provision for lipid synthesis, we propose a third option: the main task of MEs together with MDHs, pyruvate dehydrogenase complex (PDC) and AAT, which provide  $\alpha$ -ketoglutarate and aspartate from OAA and glutamate, might be the replenishment of the TCA cycle with malate, OAA, and Acetyl CoA (Fig. 3).

### The improbability of a biochemical C4-type CCM pathway in *P. tricornutum*

A biochemical C4 CCM pathway requires spatial compartmentation which is achieved in higher plants by a specialized leaf anatomy involving two different cell types (bundle sheath and mesophyll cells) or in single-celled photosynthetic cells by dimorphic chloroplasts (containing

a central compartment and peripheral chloroplasts) (Sheen and Bogorad 1987; Voznesenskaya et al. 2002; Offermann et al. 2011). In C4 plants like *Zea mays*, the carboxylases PEPC and PEPCK are fixing  $C_i$  in the cytosol of mesophyll cells. After the resulting C4 molecules have been transported into the chloroplasts of the bundle sheath cells, the  $CO_2$  again is released via decarboxylases such as ME or PEPCK (Sage et al. 2012) (Fig. 2 & Online Resource Table IV). The bundle sheath cells lack cytosolic CAs, so that  $CO_2$ , which diffuses into the cytosol, does not lead to cytosolic accumulation of bicarbonate. Therefore, maintenance of a cytosolic  $CO_2$  concentration close to the plastidic  $CO_2$  level may prevent loss of  $CO_2$  by diffusion. However, in a single cell C4 pathway, the release of  $CO_2$  in the direct vicinity of RubisCO is essential for the operation of an efficient C4 photosynthesis (Wang et al. 2014).

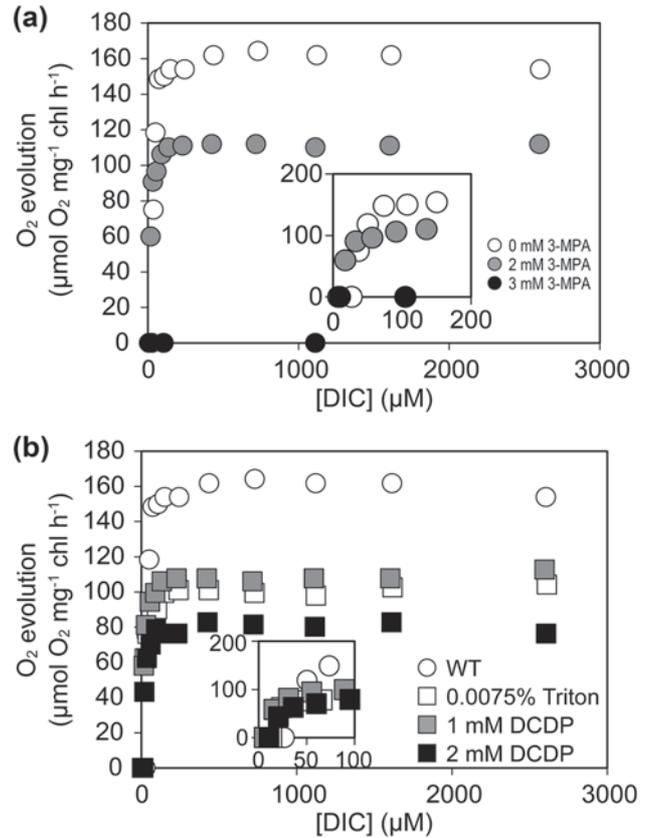
C3 plants like *A. thaliana* generally possess the genes for enzymes that are necessary for C4 pathways in other organisms (Fig. 2 & Online Resource Table IV), without actually performing C4 photosynthesis (Leegood 2002; Sage et al. 2012; Li et al. 2014). The respective enzymes are required for other processes, like the shikimate pathway or the anaplerotic link between the TCA cycle and glycolysis (Brown et al. 2010). In order to examine whether PEPCK and PEPC are involved in a CCM in *P. tricornutum*, we measured photosynthetic activities in the presence of inhibitors for PEPCK and PEPC using 3-mercaptopycolinic acid (3-MPA) and 3,3-dichloro-2-(dihydroxyphosphinoylmethyl)-propenoate (DCDP), which specifically inhibit respectively the oxaloacetate decarboxylation activity of the PEPCK enzyme and the PEP carboxylation activity of the PEPC enzyme in C4 plant leaves. RubisCO activity is not inhibited by these drugs (Ray and Black 1976). If the mitochondrial PEPCK decarboxylation contributes to the significant fraction of the diatom CCM, 3-MPA should show some negative effects on the photosynthetic affinity for DIC in *P. tricornutum*. The photosynthetic parameters in wild-type cells have been determined in the presence of various concentrations of 3-MPA (Table 2; Fig. 4a). The initial slope, described by the relationship between photosynthetic rate and DIC concentration

**Table 2** Evaluation of the toxic effect of various concentrations of the PEPCK inhibitor 3-MPA on the photosynthetic affinity for  $C_i$  and on the maximum photosynthetic rate in wild-type cells

3-MPA (mM)	$P_{max}$ ( $\mu\text{mol O}_2 \text{ mg}^{-1} \text{ chl h}^{-1}$ )	$K_{0.5}$ ( $\mu\text{M}$ )	APC ( $\mu\text{mol O}_2 \text{ mg}^{-1} \text{ chl h}^{-1} \mu\text{M}^{-1}$ )
0	$163.0 \pm 0.9^a$	$36.0 \pm 2.7^a$	$4.2 \pm 0.7^a$
2	$124.0 \pm 25.1^a$	$17.6 \pm 0.9^a$	$3.7 \pm 0.3^a$
3	N.D. <sup>b</sup>	N.D.	N.D.

<sup>a</sup>Values are the means  $\pm$  SD of three separate measurements

<sup>b</sup>Not detected



**Fig. 4** The kinetic plots of photosynthetic rate in air-grown cells as a function of [DIC] in the absence or presence of various concentrations of 3-MPA or DCDP. **a** The kinetic plots of the photosynthetic rates in air-grown cells as a function of dissolved inorganic carbon concentration ([DIC]) in the absence or presence of 3-MPA. Inset: Plots at DIC concentrations below 200  $\mu\text{M}$ . White circles: control (non-treated wild-type cells). Gray circles: 2 mM 3-MPA. Black circles: 3 mM 3-MPA. **b** The kinetic plots of the photosynthetic rates in air-grown cells as a function of [DIC] in the absence or presence of DCDP. Inset: Plots at DIC concentrations below 100  $\mu\text{M}$ . White circles: control (non-treated wild-type cells). White squares: 0.0075% Triton-X 100. Gray squares: 1 mM DCDP. Black squares: 2 mM DCDP

(apparent photosynthetic conductance: APC), did not decrease by addition of 2 mM 3-MPA. This implies that the DIC conductance in whole cell photosynthesis, i.e. the DIC uptake and transport to RubisCO, was not reduced even though the PEPCK activity was suppressed. This further indicates that CCM efficiency is not affected by a reduced PEPCK activity. Therefore, we conclude that the PEPCK in *P. tricornutum* does not contribute to the CCM.

On the other hand, photosynthesis at the saturating concentration of DIC was clearly suppressed by 2 mM 3-MPA (Fig. 4a). This allows the assumption that the efficiency of the electron transport chain, the strength of the proton gradient, the actual  $CO_2$  fixation via RubisCO in the Calvin Cycle or any other downstream reaction after the  $H_2O$  oxidation

**Table 3** Evaluation of the toxic effect of various concentrations of the PEPC inhibitor DCDP on the photosynthetic affinity for  $C_i$  and on the maximum photosynthetic rate in wild-type cells

	$P_{max}$ ( $\mu\text{mol O}_2 \text{ mg}^{-1} \text{ chl h}^{-1}$ )	$K_{0.5}$ ( $\mu\text{M}$ )	APC ( $\mu\text{mol O}_2 \text{ mg}^{-1} \text{ chl h}^{-1} \mu\text{M}^{-1}$ )
No treatment	163.0 $\pm$ 0.9 <sup>a</sup>	36.0 $\pm$ 2.7 <sup>a</sup>	4.2 $\pm$ 0.7 <sup>a</sup>
Control (+0.0075% Triton)	126.7 $\pm$ 22.7 <sup>a</sup>	55.3 $\pm$ 41.4 <sup>a</sup>	2.6 $\pm$ 1.4 <sup>a</sup>
1 mM DCDP (+0.0075% Triton)	114.1 $\pm$ 6.5 <sup>a</sup>	26.6 $\pm$ 11.5 <sup>a</sup>	2.9 $\pm$ 0.2 <sup>a</sup>
2 mM DCDP (+0.0075% Triton)	93.7 $\pm$ 16.7 <sup>a</sup>	32.5 $\pm$ 18.5 <sup>a</sup>	2.4 $\pm$ 0.03 <sup>a</sup>

<sup>a</sup>Values are the means  $\pm$  SD of three separate measurements

might be suppressed by 3-MPA. 3 mM 3-MPA completely abolished photosynthesis revealing that 3-MPA has strong toxicity to the photosynthesis machinery but has no effect on the photosynthetic efficiency under limited DIC concentrations (Fig. 4a; Table 2).

A similar inhibitory profile on photosynthetic parameters of *P. tricornutum* WT cells was also observed in the DCDP treatment. DCDP is a PEP analog and known as a specific inhibitor for PEPC (Jenkins et al. 1987b; Izui et al. 2004).

2 mM DCDP resulted in a decline of  $P_{max}$  by 74%, from a control value of 126.7  $\mu\text{mol O}_2 \text{ mg}^{-1} \text{ chl h}^{-1}$  down to 93.7  $\mu\text{mol O}_2 \text{ mg}^{-1} \text{ chl h}^{-1}$ , while APC was not affected (Table 3; Fig. 4b).  $K_{0.5}[\text{DIC}]$  was unaffected by the DCDP treatment (Table 3), suggesting that the inhibitor shows a toxic effect on photosynthesis but not on the CCM. These findings strongly indicate that neither PEPC nor PEPCK are involved in an active CCM. Moreover, the work by Clement et al. (2017) revealed that even though the PEPC activity in *P. tricornutum* increased under low  $\text{CO}_2$  concentrations, the overall PEPC:RubisCO ratio remains below one. Furthermore, they show that neither the PPDK nor the NAD-ME activity are affected by changing  $\text{CO}_2$  concentrations which leads to the conclusion that *P. tricornutum* does not perform a biochemical CCM. Instead, the core C4 enzymes seem to participate in other pathways e.g., gluconeogenesis, amino acid synthesis or replenishment of the TCA cycle.

## Conclusion

In this study, the localizations of the core C4 CCM proteins in the diatom *P. tricornutum* have been determined experimentally. None of the known, investigated decarboxylases, either the MEs or the PEPCK, are located in the plastid, which we consider as a prerequisite for a single cell C4-type CCM pathway. The intracellular distribution of the investigated enzymes is strikingly similar to those of *T. pseudonana*, determined by Tanaka et al. (2014). In contrast, the localization of other critical players for a biophysical CCM, the CAs, is clearly different between the two diatoms (Samukawa et al. 2014). Supporting the <sup>14</sup>C labeling

tests by Holdsworth and Colbeck (1976), the enzyme activity tests by Clement et al. (2016) and the PPDK silencing experiments by Haimovich-Dayan et al. (2013), our results confirm that a C4-like CCM does not play an essential role in  $\text{CO}_2$  fixation in *P. tricornutum* and that the investigated enzymes are more likely to fulfill similar functions as seen in C3 plants.

**Acknowledgements** We are grateful to D. Ballert (Universität Konstanz) for the genetic transformation of *P. tricornutum* and the cultivation of the transformed cell lines; to the Bioimaging Center at the University of Konstanz for providing the LSM 510 META; to Dr. Katsura Izui for providing the DCDP and to Dr. Pavel Hrouzek (Centre Algatech) for providing the MitoTracker. This study was supported by the German-Israel Foundation for Scientific Research and Development (GIF; Project No. G-1011-199.12/2008 to A.K. and P.G.K.); the University of Konstanz, the Grant-in-Aid for Scientific Research B (Grant No. 24310015 to Y. M.); the Grant-in-Aid for Scientific Research on Innovative Areas (Grant No. 16H06557 to Y. M.); the Japan Society for the Promotion of Science (JSPS; Grant No. 26870750 to S. K.) and by a 2 years Postdoctoral Researcher Fellowship (Mzdová podpora postdoktorandů) from the Academy of Sciences of the Czech Republic (to D.E.).

## References

- Allen AE, Moustafa A, Montsant A, Eckert A, Kroth PG, Bowler C (2012) Evolution and functional diversification of fructose biphosphate aldolase genes in photosynthetic marine diatoms. *Mol Biol Evol* 29:367–379
- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Alvarez CE, Saigo M, Margarit E, Andreo CS, Drincovich MF (2013) Kinetics and functional diversity among the five members of the NADP-malic enzyme family from *Zea mays*, a C4 species. *Photosynth Res* 115:65–80
- Anning T, Nimer N, Merrett MJ, Brownlee C (1996) Costs and benefits of calcification in coccolithophorids. *J Mar Syst* 9:45–56
- Appleby G, Colbeck J, Holdsworth ES (1980) Beta carboxylation enzymes in marine phytoplankton and isolation and purification of pyruvate carboxylase from *Amphidinium carterae* (Dinophyceae). *J Phycol* 16:290–295
- Apt KE, Zaslavkaia L, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR, Kroth PG (2002) In vivo characterization

- of diatom multipartite plastid targeting signals. *J Cell Sci* 115:4061–4069
- Armbrust EV, Berges JA, Bowler C, Green BR, Martinez D, Putnam NH, Zhou SG, Allen AE, Apt KE, Bechner M, Brzezinski MA, Chaal BK, Chiovitti A, Davis AK, Demarest MS, Detter JC, Glavina T, Goodstein D, Hadi MZ, Hellsten U, Hildebrand M, Jenkins BD, Jurka J, Kapitonov VV, Kröger N, Lau WWY, Lane TW, Larimer FW, Lippmeier JC, Lucas S, Medina M, Montsant A, Obornik M, Parker MS, Palenik B, Pazour GJ, Richardson PM, Rynearson TA, Saito MA, Schwartz DC, Thamatrakoln K, Valentin K, Vardi A, Wilkerson FP, Rokhsar DS (2004) The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. *Science* 306:79–86
- Badger MR, Andrews TJ, Whitney SM, Ludwig M, Yellowlees DC, Leggat W, Price GD (1998) The diversity and coevolution of Rubisco, plastids, pyrenoids, and chloroplast-based CO<sub>2</sub>-concentrating mechanisms in algae. *Can J Bot (Revue Canadienne De Botanique)* 76:1052–1071
- Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* 340:783–795
- Beopoulos A, Cescut J, Haddouche R, Uribelarra JL, Molina-Jouve C, Nicaud JM (2009) *Yarrowia lipolytica* as a model for bio-oil production. *Prog Lipid Res* 48:375–387
- Birmingham BC, Colman B (1979) Measurement of carbon dioxide compensation points of freshwater algae. *Plant Physiol* 64:892–895
- Bolender N, Sickmann A, Wagner R, Meisinger C, Pfanner N (2008) Multiple pathways for sorting mitochondrial precursor proteins. *Embo Rep* 9:42–49
- Bowes G (2011) Single-cell C<sub>4</sub> photosynthesis in aquatic plants. In: Raghavendra AS, Sage RF (eds) C<sub>4</sub> photosynthesis and related CO<sub>2</sub> concentrating mechanisms, vol 32, Springer, Dordrecht, pp 63–80
- Bowler C, Allen AE, Badger JH, Grimwood J, Jabbari K, Kuo A, Maheswari U, Martens C, Maumus F, O'tillar RP, Rayko E, Salamov A, Vandepoele K, Beszteri B, Gruber A, Heijde M, Katinka M, Mock T, Valentin K, Verret F, Berges JA, Brownlee C, Cadoret JP, Chiovitti A, Choi CJ, Coesel S, De Martino A, Detter JC, Durkin C, Falciatore A, Fournet J, Haruta M, Huysman MJJ, Jenkins BD, Jiroutova K, Jorgensen RE, Joubert Y, Kaplan A, Kröger N, Kroth PG, La Roche J, Lindquist E, Lommer M, Martin-Jézéquel V, Lopez PJ, Lucas S, Mangogna M, McGinnis K, Medlin LK, Montsant A, Oudot-Le Secq MP, Napoli C, Obornik M, Parker MS, Petit JL, Porcel BM, Poulsen N, Robison M, Rychlewski L, Rynearson TA, Schmutz J, Shapiro H, Siant M, Stanley M, Sussman MR, Taylor AR, Vardi A, von Dassow P, Vyverman W, Willis A, Wyrwicz LS, Rokhsar DS, Weissenbach J, Armbrust EV, Green BR, Van De Peer Y, Grigoriev IV (2008) The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature* 456:239–244
- Braun FJ, Hegemann P (1999) Direct measurement of cytosolic calcium and pH in living *Chlamydomonas reinhardtii* cells. *Eur J Cell Biol* 78:199–208
- Brown NJ, Palmer BG, Stanley S, Hajaji H, Janacek SH, Astley HM, Parsley K, Kajala K, Quick WP, Trenkamp S, Fernie AR, Maurino VG, Hibberd JM (2010) C<sub>4</sub> acid decarboxylases required for C<sub>4</sub> photosynthesis are active in the mid-vein of the C<sub>3</sub> species *Arabidopsis thaliana*, and are important in sugar and amino acid metabolism. *Plant J* 61:122–133
- Buchanan BB (1991) Regulation of CO<sub>2</sub> assimilation in oxygenic photosynthesis: the ferredoxin/thioredoxin system Perspective on its discovery, present status, and future development. *Arch Biochem Biophys* 288:1–9
- Chastain CJ, Xu WX, Parsley K, Sarath G, Hibberd JM, Chollet R (2008) The pyruvate, orthophosphate dikinase regulatory proteins of *Arabidopsis* possess a novel, unprecedented Ser/Thr protein kinase primary structure. *Plant J* 53:854–863
- Chastain CJ, Failing CJ, Manandhar L, Zimmerman MA, Lakner MM, Nguyen THT (2011) Functional evolution of C<sub>4</sub> pyruvate, orthophosphate dikinase. *J Exp Bot* 62:3083–3091
- Chauton MS, Winge P, Brembu T, Vadstein O, Bones AM (2013) Gene regulation of carbon fixation, storage, and utilization in the diatom *Phaeodactylum tricorutum* acclimated to light/dark cycles. *Plant Physiol* 161:1034–1048
- Clement R, Dimnet L, Maberly SC, Gontero B (2016) The nature of the CO<sub>2</sub>-concentrating mechanisms in a marine diatom, *Thalassiosira pseudonana*. *New Phytol* 209:1417–1427
- Clement R, Jensen E, Prioretti L, Maberly SC, Gontero B (2017) Diversity of CO<sub>2</sub> concentrating mechanisms and responses to CO<sub>2</sub> concentration in marine and freshwater diatoms. *J Exp Bot* 68:3925–3935
- Consortium TU (2014) Activities at the universal protein resource (UniProt). *Nucleic Acids Res* 42:D191–D198
- Davis A, Abbriano R, Smith SR, Hildebrand M (2017) Clarification of photorespiratory processes and the role of malic enzyme in diatoms. *Protist* 168:134–153
- Descolas-Gros C, Oriol L (1992) Variations in carboxylase activity in marine phytoplankton cultures. Beta-carboxylation in carbon flux studies. *Mar Ecol Prog Ser* 85:163–169
- Detarsio E, Maurino VG, Alvarez CE, Müller GL, Andreo CS, Drincovich MF (2008) Maize cytosolic NADP-malic enzyme (ZmCytNADP-ME): a phylogenetically distant isoform specifically expressed in embryo and emerging roots. *Plant Mol Biol* 68:355–367
- Dittrich P, Campbell WH, Black CC (1973) Phosphoenolpyruvate carboxykinase in plants exhibiting crassulacean acid metabolism. *Plant Physiol* 52:357–361
- Edwards GE, Voznesenskaya EV (2011) C<sub>4</sub> photosynthesis: Kranz forms and single-cell C<sub>4</sub> in terrestrial plants. In: Raghavendra AS, Sage RF (eds) C<sub>4</sub> photosynthesis and related CO<sub>2</sub> concentrating mechanisms, vol 32, Springer, Dordrecht, pp 29–61
- Emanuelsson O, Nielsen H, Brunak S, von Heijne G (2000) Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* 300:1005–1016
- Giordano M, Beardall J, Raven JA (2005) CO<sub>2</sub> concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution. *Annu Rev Plant Biol* 56:99–131
- Gould SB, Sommer MS, Kroth PG, Gile GH, Keeling PJ, Maier UG (2006) Nucleus-to-nucleus gene transfer and protein retargeting into a remnant cytoplasm of cryptophytes and diatoms. *Mol Biol Evol* 23:2413–2422
- Granum E, Raven JA, Leegood RC (2005) How do marine diatoms fix 10 billion tonnes of inorganic carbon per year? *Can J Bot (Revue Canadienne De Botanique)* 83:898–908
- Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, Kuo A, Minovitsky S, Nikitin R, Ohm RA, O'tillar R, Poliakov A, Ratnere I, Riley R, Smirnova T, Rokhsar D, Dubchak I (2012) The genome portal of the department of energy joint genome institute. *Nucleic Acids Res* 40:D26–D32
- Gruber A, Kroth PG (2014) Deducing intracellular distributions of metabolic pathways from genomic data. *Methods Mol Biol* 1083:187–211
- Gruber A, Kroth PG (2017) Intracellular metabolic pathway distribution in diatoms and tools for genome-enabled experimental diatom research. *Philos Trans R Soc B* 372:20160402
- Gruber A, Vugrinec S, Hempel F, Gould SB, Maier UG, Kroth PG (2007) Protein targeting into complex diatom plastids: functional characterisation of a specific targeting motif. *Plant Mol Biol* 64:519–530

- Gruber A, Weber T, Río Bártulos C, Vugrinec S, Kroth PG (2009) Intracellular distribution of the reductive and oxidative pentose phosphate pathways in two diatoms. *J Basic Microbiol* 49:58–72
- Gruber A, Rocap G, Kroth PG, Armbrust EV, Mock T (2015) Plastid proteome prediction for diatoms and other algae with secondary plastids of the red lineage. *Plant J* 81:519–528
- Guillard RR, Ryther JH (1962) Studies of marine planktonic diatoms 1. *Cyclotella nana* Hustedt, and *Detonula confervacea* (Cleve) Gran. *Can J Microbiol* 8:229–239
- Haimovich-Dayana M, Garfinkel N, Ewe D, Marcus Y, Gruber A, Wagner H, Kroth PG, Kaplan A (2013) The role of C4 metabolism in the marine diatom *Phaeodactylum tricorutum*. *New Phytol* 197:177–185
- Hao GF, Chen HQ, Wang L, Gu ZN, Song YD, Zhang H, Chen W, Chen YQ (2014) Role of malic enzyme during fatty acid synthesis in the oleaginous fungus *Mortierella alpina*. *Appl Environ Microbiol* 80:2672–2678
- Harrison PJ, Waters RE, Taylor FJR (1980) A broad spectrum artificial seawater medium for coastal and open ocean phytoplankton. *J Phycol* 16:28–35
- Häusler RE, Hirsch HJ, Kreuzaler F, Peterhänsel C (2002) Overexpression of C4-cycle enzymes in transgenic C3 plants: a biotechnological approach to improve C3-photosynthesis. *J Exp Bot* 53:591–607
- Heber U (1974) Metabolite exchange between chloroplasts and cytoplasm. *Annu Rev Plant Physiol Plant Mol Biol* 25:393–421
- Herve V, Derr J, Douady S, Quinet M, Moisan L, Lopez PJ (2012) Multiparametric analyses reveal the pH-dependence of silicon biomineralization in diatoms. *Plos One* 7:e46722
- Holdsworth ES, Colbeck J (1976) The pattern of carbon fixation in the marine unicellular alga *Phaeodactylum tricorutum*. *Mar Biol* 38:189–199
- Hopkinson BM, Meile C, Shen C (2013) Quantification of extracellular carbonic anhydrase activity in two marine diatoms and investigation of its role. *Plant Physiol* 162:1142–1152
- Izui K, Matsumura H, Furumoto T, Kai Y (2004) Phosphoenolpyruvate carboxylase: a new era of structural biology. *Annu Rev Plant Biol* 55:69–84
- Jenkins CLD, Hatch MD (1985) Properties and reaction mechanism of C4 leaf pyruvate, Pi dikinase. *Arch Biochem Biophys* 239:53–62
- Jenkins CLD, Burnell JN, Hatch MD (1987a) Form of inorganic carbon involved as a product and as an inhibitor of C4 acid decarboxylases operating in C4 photosynthesis. *Plant Physiol* 85:952–957
- Jenkins CLD, Harris RLN, McFadden HG (1987b) 3,3-dichloro-2-dihydroxyphosphinoylmethyl-2-propenoate, a new, specific inhibitor of phosphoenolpyruvate carboxylase. *Biochem Int* 14:219–226
- Kaplan A, Reinhold L (1999) CO<sub>2</sub> concentrating mechanisms in photosynthetic microorganisms. *Annu Rev Plant Physiol Plant Mol Biol* 50:539–570
- Kikutani S, Nakajima K, Nagasato C, Tsuji Y, Miyatake A, Matsuda Y (2016) Thylakoid luminal theta-carbonic anhydrase critical for growth and photosynthesis in the marine diatom *Phaeodactylum tricorutum*. *Proc Natl Acad Sci USA* 113:9828–9833
- Kilian O, Kroth PG (2005) Identification and characterization of a new conserved motif within the presequence of proteins targeted into complex diatom plastids. *Plant J* 41:175–183
- Kitao Y, Harada H, Matsuda Y (2008) Localization and targeting mechanisms of two chloroplastic  $\beta$ -carbonic anhydrases in the marine diatom *Phaeodactylum tricorutum*. *Physiol Plant* 133:68–77
- Kroth PG, Chiovitti A, Gruber A, Martin-Jézéquel V, Mock T, Parker MS, Stanley MS, Kaplan A, Caron L, Weber T, Maheswari U, Armbrust EV, Bowler C (2008) A model for carbohydrate metabolism in the diatom *Phaeodactylum tricorutum* deduced from comparative whole genome analysis. *Plos ONE* 3:e1426
- Kustka AB, Milligan AJ, Zheng HY, New AM, Gates C, Bidle KD, Reinfelder JR (2014) Low CO<sub>2</sub> results in a rearrangement of carbon metabolism to support C4 photosynthetic carbon assimilation in *Thalassiosira pseudonana*. *New Phytol* 204:507–520
- Lai LB, Tausta SL, Nelson TM (2002) Differential regulation of transcripts encoding cytosolic NADP-malic enzyme in C3 and C4 *Flaveria* species. *Plant Physiol* 128:140–149
- Leegood RC (2002) C4 photosynthesis: principles of CO<sub>2</sub> concentration and prospects for its introduction into C3 plants. *J Exp Bot* 53:581–590
- Li YY, Xu JJ, Ul-Haq N, Zhang H, Zhu XG (2014) Was low CO<sub>2</sub> a driving force of C4 evolution: *Arabidopsis* responses to long-term low CO<sub>2</sub> stress. *J Exp Bot* 65:3657–3667
- Liaud MF, Lichtle C, Apt K, Martin W, Cerff R (2000) Compartment-specific isoforms of TPI and GAPDH are imported into diatom mitochondria as a fusion protein: evidence in favor of a mitochondrial origin of the eukaryotic glycolytic pathway. *Mol Biol Evol* 17:213–223
- Matsuda Y, Hara T, Colman B (2001) Regulation of the induction of bicarbonate uptake by dissolved CO<sub>2</sub> in the marine diatom, *Phaeodactylum tricorutum*. *Plant Cell Environment* 24:611–620
- Matsuda Y, Nakajima K, Tachibana M (2011) Recent progresses on the genetic basis of the regulation of CO<sub>2</sub> acquisition systems in response to CO<sub>2</sub> concentration. *Photosynth Res* 109:191–203
- McFadden GI (2001) Primary and secondary endosymbiosis and the origin of plastids. *J Phycol* 37:951–959
- Moog D, Stork S, Zauner S, Maier UG (2011) In silico and in vivo investigations of proteins of a minimized eukaryotic cytoplasm. *Genome Biol Evol* 3:375–382
- Nakajima K, Tanaka A, Matsuda Y (2013) SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc Natl Acad Sci USA* 110:1767–1772
- Nelson DM, Tréguer P, Brzezinski MA, Leynaert A, Quéguiner B (1995) Production and dissolution of biogenic silica in the ocean: revised global estimates, comparison with regional data and relationship to biogenic sedimentation. *Global Biogeochem Cycles* 9:359–372
- Nikolau BJ, Ohlrogge JB, Wurtele ES (2003) Plant biotin-containing carboxylases. *Arch Biochem Biophys* 414:211–222
- Offermann S, Okita TW, Edwards GE (2011) Resolving the compartmentation and function of C4 photosynthesis in the single-cell C4 species *Bienertia sinuspersici*. *Plant Physiol* 155:1612–1628
- Parsley K, Hibberd JM (2006) The *Arabidopsis* PDK gene is transcribed from two promoters to produce differentially expressed transcripts responsible for cytosolic and plastidic proteins. *Plant Mol Biol* 62:339–349
- Pedersen AG, Nielsen H (1997) Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. *Ismb-97, Fifth International Conference on Intelligent Systems for Molecular Biology, Proceedings* 5:226–233
- Petersen TN, Brunak S, von Heijne G, Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 8:785–786
- Ratledge C, Wynn JP (2002) The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms. *Adv Appl Microbiol* 51:1–51
- Ray TB, Black CC (1976) Inhibition of oxalacetate decarboxylation during C4 photosynthesis by 3-mercaptopycolinic acid. *J Biol Chem* 251:5824–5826
- Reinfelder JR (2011) Carbon concentrating mechanisms in eukaryotic marine phytoplankton. In: Carlson CA, Giovannoni SJ (eds) *Annual review of marine science, vol 3, Annual Reviews, Palo Alto, pp 291–315*

- Reinfelder JR, Kraepiel AML, Morel FMM (2000) Unicellular C4 photosynthesis in a marine diatom. *Nature* 407:996–999
- Reiskind JB, Bowes G (1991) The role of phosphoenolpyruvate carboxykinase in a marine macroalga with C4-like photosynthetic characteristics. *Proc Natl Acad Sci USA* 88:2883–2887
- Riebesell U, Wolfgladrow DA, Smetacek V (1993) Carbon dioxide limitation of marine phytoplankton growth rates. *Nature* 361:249–251
- Río Bártulos C (2007) Eine mitochondriale Teil-Glykolyse bei *Phaeodactylum tricorutum* und anderen Heterokonten: Nachweis der Kompartimentierung mittels dem “green fluorescent protein” und molekulare Phylogenie. Dissertation, Technische Universität Carolo-Wilhelmina zu Braunschweig
- Roberts K, Granum E, Leegood RC, Raven JA (2007a) C3 and C4 pathways of photosynthetic carbon assimilation in marine diatoms are under genetic, not environmental, control. *Plant Physiol* 145:230–235
- Roberts K, Granum E, Leegood RC, Raven JA (2007b) Carbon acquisition by diatoms. *Photosynth Res* 93:79–88
- Rottberger J, Gruber A, Boenigk J, Kroth PG (2013) Influence of nutrients and light on autotrophic, mixotrophic and heterotrophic freshwater chrysophytes. *Aquat Microb Ecol* 71:179–191
- Sage RF, Christin PA, Edwards EJ (2011) The C4 plant lineages of planet Earth. *J Exp Bot* 62:3155–3169
- Sage RF, Sage TL, Kocacinar F (2012) Photorespiration and the evolution of C4 photosynthesis. In: Merchant SS (ed) *Annual review of plant biology*, vol 63, Annual Reviews, Palo Alto, pp 19–47
- Sambrook J, Fritsch E, Maniatis T (1989) *Molecular cloning: a laboratory manual*. Cold Spring Harbor Laboratory Press, New York
- Samukawa M, Shen C, Hopkinson BM, Matsuda Y (2014) Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynth Res* 121:235–249
- Schaeffer ML, Harper LC, Gardiner JM, Andorf CM, Campbell DA, Cannon EKS, Sen TZ, Lawrence CJ (2011) MaizeGDB: curation and outreach go hand-in-hand. *Database* 2011:bar022
- Scheibe R (1987) NADP<sup>+</sup>-malate dehydrogenase in C3-plants: regulation and role of a light-activated enzyme. *Physiol Plant* 71:393–400
- Shao H, Gontero B, Maberly SC, Jiang HS, Cao Y, Li W, Huang WM (2017) Responses of *Ottelia alismoides*, an aquatic plant with three CCMs, to variable CO<sub>2</sub> and light. *J Exp Bot* 68:3985–3995
- Sheen J (1991) Molecular mechanisms underlying the differential expression of *Maize* pyruvate, orthophosphate dikinase genes. *Plant Cell* 3:225–245
- Sheen JY, Bogorad L (1987) Differential expression of C4 pathway genes in mesophyll and bundle sheath cells of greening *Maize* leaves. *J Biol Chem* 262:11726–11730
- Shen C, Dupont CL, Hopkinson BM (2017) The diversity of carbon dioxide-concentrating mechanisms in marine diatoms as inferred from their genetic content. *J Exp Bot* 68:3937–3948
- Smith CM, Duff SMG, Chollet R (1994) Partial purification and characterization of *Maize*-leaf pyruvate, orthophosphate dikinase regulatory protein: a low-abundance, mesophyll-chloroplast stromal protein. *Arch Biochem Biophys* 308:200–206
- Sommer MS, Gould SB, Lehmann P, Gruber A, Przyborski JM, Maier UG (2007) Der1-mediated preprotein import into the periplastid compartment of chromalveolates? *Mol Biol Evol* 24:918–928
- Sonnhammer EL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. In: *Proceedings of the International Conference on Intelligent Systems for Molecular Biology* 6:175–182
- Stucka R, Dequin S, Salmon JM, Gancedo C (1991) DNA-sequences in chromosome II and chromosome VII code for pyruvate carboxylase isoenzymes in *Saccharomyces cerevisiae*: analysis of pyruvate carboxylase-deficient strains. *Mol General Genet* 229:307–315
- Tachibana M, Allen AE, Kikutani S, Endo Y, Bowler C, Matsuda Y (2011) Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricorutum* and *Thalassiosira pseudonana*. *Photosynth Res* 109:205–221
- Tanaka R, Kikutani S, Mahardika A, Matsuda Y (2014) Localization of enzymes relating to C4 organic acid metabolisms in the marine diatom, *Thalassiosira pseudonana*. *Photosynth Res* 121:251–263
- Taylor L, Nunes-Nesi A, Parsley K, Leiss A, Leach G, Coates S, Wingler A, Fernie AR, Hibberd JM (2010) Cytosolic pyruvate, orthophosphate dikinase functions in nitrogen remobilization during leaf senescence and limits individual seed growth and nitrogen content. *Plant J* 62:641–652
- Tortell PD (2000) Evolutionary and ecological perspectives on carbon acquisition in phytoplankton. *Limnol Oceanogr* 45:744–750
- Tsuji Y, Suzuki I, Shiraiwa Y (2009) Photosynthetic carbon assimilation in the coccolithophorid *Emiliania huxleyi* (Haptophyta): evidence for the predominant operation of the C3 cycle and the contribution of beta-carboxylases to the active anaplerotic reaction. *Plant Cell Physiol* 50:318–329
- Tsuji Y, Suzuki I, Shiraiwa Y (2012) Enzymological evidence for the function of a plastid-located pyruvate carboxylase in the Haptophyte alga *Emiliania huxleyi*: a novel pathway for the production of C4 compounds. *Plant Cell Physiol* 53:1043–1052
- Tsuji Y, Mahardika A, Matsuda Y (2017a) Evolutionarily distinct strategies for the acquisition of inorganic carbon from seawater in marine diatoms. *J Exp Bot* 68:3949–3958
- Tsuji Y, Nakajima K, Matsuda Y (2017b) Molecular aspects of the biophysical CO<sub>2</sub>-concentrating mechanism and its regulation in marine diatoms. *J Exp Bot* 68:3763–3772
- Voznesenskaya EV, Franceschi VR, Kiiirats O, Artyusheva EG, Freitag H, Edwards GE (2002) Proof of C4 photosynthesis without Kranz anatomy in *Bienertia cycloptera* (Chenopodiaceae). *Plant J* 31:649–662
- Wang Y, Bräutigam A, Weber APM, Zhu XG (2014) Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *J Exp Bot* 65:3567–3578
- Weber T, Gruber A, Kroth PG (2009) The presence and localization of thioredoxins in diatoms, unicellular algae of secondary endosymbiotic origin. *Mol Plant* 2:468–477
- Wheeler MCG, Arias CL, Maurino VG, Andreo CS, Drincovich MF (2009) Identification of domains involved in the allosteric regulation of cytosolic *Arabidopsis thaliana* NADP-malic enzymes. *Febs J* 276:5665–5677
- Wynn JP, Ratledge C (1997) Malic enzyme is a major source of NADPH for lipid accumulation by *Aspergillus nidulans*. *Microbiology* 143:253–257
- Wynn JP, Kendrick A, Ratledge C (1997) Sesamol as an inhibitor of growth and lipid metabolism in *Mucor circinelloides* via its action on malic enzyme. *Lipids* 32:605–610
- Yang ZK, Niu YF, Ma YH, Xue J, Zhang MH, Yang WD, Liu JS, Lu SH, Guan YF, Li HY (2013) Molecular and cellular mechanisms of neutral lipid accumulation in diatom following nitrogen deprivation. *Biotechnol Biofuels* 6:67:1–14
- Young JN, Hopkinson BM (2017) The potential for co-evolution of CO<sub>2</sub>-concentrating mechanisms and Rubisco in diatoms. *J Exp Bot* 68:3751–3762
- Young JN, Heureux AM, Sharwood RE, Rickaby RE, Morel FM, Whitney SM (2016) Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *J Exp Bot* 67:3445–3456
- Zaslavskaya LA, Lippmeier JC, Kroth PG, Grossman AR, Apt KE (2000) Transformation of the diatom *Phaeodactylum tricorutum* (Bacillariophyceae) with a variety of selectable marker and reporter genes. *J Phycol* 36:379–386

Zhang Y, Adams IP, Ratledge C (2007) Malic enzyme: the controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. *Microbiology* 153:2013–2025

Zhang YZ, Yin LY, Jiang HS, Li W, Gontero B, Maberly SC (2014) Biochemical and biophysical CO<sub>2</sub> concentrating mechanisms in two species of freshwater macrophyte within the genus *Ottelia* (Hydrocharitaceae). *Photosynth Res* 121:285–297