ORIGINAL PAPER

Clarification of Photorespiratory Processes and the Role of Malic Enzyme in Diatoms

Aubrey Davis^{a,1}, Raffaella Abbriano^{a,1}, Sarah R. Smith^{b,c}, and Mark Hildebrand^{a,2}

^aMarine Biology Research Division, Scripps Institution of Oceanography, UC San Diego, La Jolla, CA, U.S.A.

^bIntegrative Oceanography Division, Scripps Institution of Oceanography, UC San Diego, La Jolla, CA, U.S.A.

^cJ. Craig Venter Institute, La Jolla, CA, U.S.A.

Submitted May 3, 2016; Accepted September 28, 2016

Monitoring Editor: Saul Purton

Running title: Photorespiratory Processes in Diatoms

¹These authors contributed equally to this work

²Corresponding author; fax +1 858 534-7313

e-mail mhildebrand@ucsd.edu (M. Hildebrand).

Evidence suggests that diatom photorespiratory metabolism is distinct from other photosynthetic eukaryotes in that there may be at least two routes for the metabolism of the photorespiratory metabolite glycolate. One occurs primarily in the mitochondria and is similar to the C2 photorespiratory pathway, and the other processes glycolate through the peroxisomal glyoxylate cycle. Genomic analysis has identified the presence of key genes required for glycolate oxidation, the glyoxylate cycle, and malate metabolism, however, predictions of intracellular localization can be ambiguous and require verification. This knowledge gap leads to uncertainties surrounding how these individual pathways operate, either together or independently, to process photorespiratory intermediates under different environmental conditions. Here, we combine in silico sequence analysis, in vivo protein localization techniques and gene expression patterns to investigate key enzymes potentially involved in photorespiratory metabolism in the model diatom *Thalassiosira pseudonana*. We demonstrate the peroxisomal localization of isocitrate lyase and the mitochondrial localization of malic enzyme and a glycolate oxidase. Based on these analyses, we propose an updated model for photorespiratory metabolism in *T. pseudonana*, as well as a mechanism by which C2 photorespiratory metabolism and its associated pathways may operate during silicon starvation and growth arrest.

Key words: Diatom; photorespiration; glyoxylate cycle; malic enzyme; *Thalassiosira pseudonana*.

Introduction

Diatoms are a ubiquitous and highly productive group of eukaryotic phytoplankton that contribute significantly to the global carbon cycle. Diatoms are estimated to fix 10 billion metric tons of inorganic carbon annually, which is at least a quarter of the carbon fixed each year in the ocean (Granum et al. 2005). Photosynthetic carbon fixation by diatoms sustains diverse marine ecosystems and contributes significantly to total export production (Sarhou et al. 2005). The high productivity that drives these large-scale processes is ultimately connected to the efficiency of carbon fixation and the metabolism of fixed carbon within diatom cells. Photorespiration can represent the second largest route for carbon flux in photosynthetic organisms (Bauwe et al. 2010; Pick et al. 2013). Photorespiration occurs when RubisCO catalyzes a reaction with O₂ rather than CO₂, resulting in reduced photosynthetic carbon fixation (Eisenhut et al. 2008). The metabolic by-product of photorespiration, 2-phosphoglycolate, cannot be used directly in biosynthetic pathways and additional energy-requiring steps are necessary to recover this carbon in a usable form. Although active carbon concentrating mechanisms can maintain a higher concentration of CO₂ within the proximity of RubisCO (Reinfelder 2011; Reinfelder et al. 2004), experimental evidence suggests that diatoms must contend with photorespiratory by-products under certain environmental conditions (Schnitzler Parker et al. 2004). The existence of photorespiration in diatoms is supported by extracellular glycolate release and transcript-level induction of key enzymes involved in glycolate metabolism in response to high light exposure (Parker and Armbrust 2005; Schnitzler Parker et al. 2004). Adaptation to high light led to less glycolate production, indicating that either photorespiration was reduced or that efficient recycling pathways were in operation (Schnitzler Parker et al. 2004). In plants, 2-phosphoglycolate recycling is accomplished through a series of reactions known as the photorespiratory C₂ cycle, which is distributed among the chloroplasts, leaf peroxisomes and mitochondria (Bauwe et al. 2012). However, the specific pathways utilized by diatoms to metabolize photorespiratory by-products are not yet fully understood, and represent a significant knowledge gap in terms of building accurate metabolic models for carbon metabolism in these important primary producers.

While many of the enzymes involved in the classical C₂ photorespiratory pathway have been identified in diatom genomes (Kroth et al. 2008), some discrepancies exist between past

experimental data and genomic information. The first step in processing photorespiratory glycolate is to convert it to glyoxylate, which can be accomplished by either glycolate dehydrogenases (GDH) or by glycolate oxidases (GOX). Biochemical studies suggest that diatoms utilize glycolate dehydrogenases (GDH) to oxidize glycolate (Suzuki et al. 1991; Winkler and Stabenau 1992), however only putative glycolate oxidases (GOX) have been identified in diatom genomes (Kroth et al. 2008). It is also unclear how similar diatom C2 pathways are to those most commonly described in other organisms. For example, glycerate kinase, the enzyme that catalyzes final step of the C2 cycle in plants, was not originally identified in diatoms (Kroth et al. 2008), although a candidate gene has been proposed (Fabris et al. 2012). Beyond this uncertainty, additional experimental evidence suggests that diatom photorespiratory metabolism may not be limited to the steps of the classical C2 cycle. For example, glycolate oxidation has been detected in both diatom peroxisomes and mitochondria, providing support for at least two possible fates for photorespiratory glycolate (Winkler and Stabenau 1995). Furthermore, reducing CO₂ concentration and increasing available light did not increase the activity of particular canonical C2 pathway enzymes in the diatom *Nitzschia laevis* (Winkler and Stabenau 1995), again supporting the possibility of a distinct photorespiratory mechanism from that commonly described.

Metabolism via the peroxisomal glyoxylate cycle has been proposed in the literature to be an alternate route to recycle photorespiratory intermediates (Fig. 1), and enzymes in this pathway are responsive to conditions that induce photorespiration in diatoms (Winkler and Stabenau 1995). The glyoxylate cycle enzymes isocitrate lyase (ICL) and malate synthase (MS) serve as a metabolic bypass to the TCA cycle, but MS can also combine photorespiratory glyoxylate with acetyl-CoA to produce malate, which then re-enters the TCA cycle in the mitochondria (Igamberdiev and Lea 2002; Winkler and Stabenau 1992) or is converted to pyruvate by malic enzyme (ME) (Sweetlove et al. 2010). This alternate fate for photorespiratory metabolites implicates malate as a potentially important compound linking the photorespiratory processes in the peroxisome with mitochondrial metabolism (Igamberdiev and Lea 2002). Another proposed pathway for photorespiratory metabolism based on research in the diatom *Cylindrotheca fusiformis* suggested that glycolate may be metabolized via enzymes derived from a bacterial glycerate pathway using a glyoxylate carboligase and tartronate semialdehyde reductase to

produce glycerate (Paul and Volcani 1976). However, genome annotation efforts have not been able to identify key enzymes involved in this pathway with confidence (Kroth et al. 2008).

Photorespiration serves an additional purpose beyond carbon recovery, as the high energy requirement of phosphoglycolate recycling also makes this pathway an important means to dissipate excess energy and prevent reaction center damage (Wingler et al. 2000; Igamberdiev et al. 2001). At high light intensities, reducing power generated by the light reactions of photosynthesis exceeds the capacity of carbon assimilating enzymes to regenerate NADP⁺ (Wilhelm and Jakob 2011). Stress conditions that further inhibit carbon fixation can exacerbate the imbalance, resulting in overreduction of the photosynthetic electron transport chain, photoinhibition, and the production of reactive oxygen species (Wingler et al. 2000). Photorespiratory metabolism is one of several alternative electron pathways (Wilhelm and Jakob 2011) that serve to prevent overreduction of photosynthetic electron transport components, since photorespiratory intermediates transported out of the chloroplast carry energy to be distributed to other organelles. Strong evidence from inhibitor studies in *Phaeodactylum tricornutum* support a role for mitochondrial electron transport processes in dissipating excess reducing power from diatom chloroplasts (Bailleul et al. 2015), but the molecules that facilitate this interaction remain to be identified. Additionally, photorespiratory intermediates are substantially integrated with other aspects of central carbon and nitrogen metabolism (Bauwe et al. 2012), including the production of metabolites that can further boost antioxidant production (Wingler et al. 2000). The role of photorespiration in alleviating photooxidative stress may be of particular importance in diatoms, as it has been shown that an extreme light shift is sufficient to induce a photorespiratory response (Allen 2005; Schnitzler Parker et al. 2004).

We previously reported on physiological and transcript level changes in response to silicon starvation in *T. pseudonana* (Smith et al. 2016). Silicon is a requirement for cell wall synthesis in most diatom species, and therefore necessary for cell division. During silicon starvation, cell division halts and fatty acid and triglyceride (TAG) synthesis are induced (Smith et al. 2016; Yu et al. 2009), signaling considerable changes in intracellular carbon partitioning. We found that the first 8 hours following transfer into silicon-free medium is characterized by initial photodamage to the PSII reaction centers, followed by a sustained decrease in the maximal rate of carbon fixation (Smith et al. 2016). In addition, we observed a large increase in the de-epoxidation state of the xanthophyll pool, indicating increased xanthophyll-dependent non-

photochemical quenching (NPQ) to reduce electron flow and protect PSII from photooxidative damage (Smith et al. 2016). The increase in NPQ suggests the need for alternative electron cycling pathways such as photorespiration under these experimental conditions. Silicon starvation potentially offers advantages over some other experimental conditions used to induce photorespiratory responses because most cellular metabolic processes are minimally affected by the lack of silicon. In addition, silicon-starvation induced photorespiration does not involve an artificial shift in CO₂ levels. While we recognize the limitations of using transcript data as a proxy to interpret shifts in metabolic processes, ample evidence exists in the literature to demonstrate that when coordinated and consistent changes in transcript levels occur at several steps in a pathway, it is most likely indicative of a metabolic shift (Fernie et al. 2002; Lancien et al. 1999; Smith et al. 2016; Sweetlove et al. 2010). Given the lack of clarity regarding photorespiratory carbon cycling in diatoms, investigation of the transcriptional response under silicon limitation may provide new insights into this process.

Using a combination of in silico analysis, transcript expression analysis, and in vivo localization to assess metabolic potential, our goal in this study was to clarify and update a model for the organization and operation of photorespiratory metabolism in diatoms by investigating the response during silicon starvation in *T. pseudonana*. Based on the existing model, we chose to focus on the following questions: 1) What are the potential routes for glycolate oxidation and transport in *T. pseudonana*? 2) Does evidence support a complete C2 cycle with carbon recycled back to the plastid? 3) What are the respective roles of the mitochondria and the peroxisome in terms of the metabolism of photorespiratory intermediates? 4) Could malic enzyme play an auxiliary role in photorespiratory metabolism? By combining our data and analysis with data available in the literature, we generate an updated model (Fig. 1) that incorporates previous experimental and analytical approaches into a self-consistent picture of photorespiratory C2 metabolism in diatoms, and identifies aspects in need of further clarification.

Results and Discussion

Photophysiology and Transcript Changes

Transfer of *T. pseudonana* to silicon-deplete medium at half of its original culture density induced light stress, as evidenced by photophysiology data (Smith et al. 2016). We analyzed the whole genome microarray dataset from Smith et al. (2016) and determined that transcripts for a subset of genes potentially involved in photorespiratory C2 metabolism were upregulated in accordance with the timing of photostress (Fig. 2). These genes included previously identified markers for high light induction of photorespiration in diatoms such as GDC-T and PGP (Parker and Armbrust 2005; Schnitzler Parker et al. 2004). The photophysiology data combined with the transcript data are consistent with induction of photorespiration under these experimental conditions.

We also observed highly coordinated changes in photorespiratory transcripts (Fig. 2), linking genes under what are likely common regulatory processes. Co-expression of many genes related to photorespiratory metabolism due to transcript regulatory sequences called CO₂-cAMP-Responsive Elements (CCREs) has been previously discussed (Hennon et al. 2015; Ohno et al. 2012). However, not all putative photorespiratory genes share the same expression pattern (Fig. 2), suggesting that multiple regulatory factors may come into play in the transcript-level response. Three distinct expression clusters containing multiple genes were identified (Fig. 2A-C), and the genes in each cluster fell into functional categories that will be described in detail below. Transcripts for three other genes did not cluster with the rest (Fig. 2D). All genes included in subsequent analysis are listed in Table 1.

Multiple Routes for Glycolate Oxidation in Diatoms

The pathway that determines the metabolic fate of photorespiratory glycolate remains unresolved in diatoms. Our analysis, coupled with previous biochemical analyses, suggests the presence of a single glycolate oxidizing activity in the peroxisome (GOX1), and potentially two genes with glycolate oxidizing activity (GOX2 and a previously unidentified GDH), with the possibility of a third (2-HADH – see section on “Processing of photorespiratory glyoxylate in the mitochondria” below) in the mitochondria (Fig. 1).

Glycolate oxidation can be catalyzed by either glycolate oxidase (GOX) or glycolate dehydrogenase (GDH). Characterized GOX enzymes use O₂ as an electron acceptor, generate H₂O₂, are insensitive to cyanide and carry out a secondary reaction with L-lactate (Frederick et al. 1973). In contrast, GDHs utilize NAD⁺/NADP⁺ as an electron acceptor, are inhibited by

cyanide and carry out a secondary reaction with D-lactate (Frederick et al. 1973). Biochemical studies in whole cell and organelle fractions of diatoms have shown a lack of H₂O₂ production, consistent with diatoms using GDH to generate glyoxylate in both the mitochondria and the peroxisome (Paul et al. 1975; Suzuki et al. 1991; Winkler and Stabenau 1992). However, lack of H₂O₂ production could be explained by efficient catalase activity, ascorbate peroxidase activity, or use of alternate electron acceptors (Mittler 2002; Paul et al. 1975; Paul and Volcani 1975; Suzuki et al. 1991), and insensitivity to cyanide and a preference by purified peroxisomes for oxidation of L-lactate is consistent with GOX activity (Suzuki et al. 1991; Winkler and Stabenau 1992). The conflicting biochemical evidence prompted a bioinformatic analysis of putative GDH and GOX genes from diatom genomes to attempt to clarify the enzymatic potential in each cellular compartment.

Two gene candidates for glycolate oxidation (GOX1 and GOX2) have previously been identified from genomic sequence (Kroth et al. 2008). To investigate the relationship of these sequences to other GOXs, we generated a maximum likelihood tree to place GOX1 and GOX2 among other characterized lactate oxidase (LOX) and GOX proteins (Fig. 3). The peroxisomal GOX1 protein, encoded by Thaps3_406, has a PTS1 C-terminal tripeptide targeting sequence (Kroth et al. 2008) (Table 1; Fig. 1A) and groups with other LOX/GOX proteins found in higher plants, green algae, and cyanobacteria (Fig. 3). This finding is similar to previous reports (Kern et al. 2011), and is consistent with the GOX-like biochemical characteristics observed in diatom peroxisomes (i.e. cyanide insensitivity and L-lactate preference). The GOX1 transcript is not coordinately expressed with other genes involved in metabolizing photorespiratory glycolate in transcriptome datasets (Fig. 2; Hennon et al. 2015), however, a substantial induction of peroxisomal glycolate oxidation activity was demonstrated under more severe photorespiratory stimulating conditions (Winkler and Stabenau 1995). We were unable to identify any other candidate genes encoding glycolate oxidation activity in the peroxisome, suggesting that GOX1 is the gene responsible for this activity.

Diatom GOX2 sequences were highly divergent from diatom GOX1 (Fig. 3). GOX2 sequences, along with other GOX-like sequences from various eukaryotic algae, are more similar to GOX-like proteins in proteobacteria. None of the proteins within this group have been biochemically characterized, so it is possible that they may be distinct from other GOXs studied to date. The *in silico* targeting of GOX2 (Thaps3_3353) is ambiguous in *T. pseudonana* (Table

1), despite a clear mitochondrial targeting signal in the homolog of this gene in *P. tricornutum* (Kroth et al. 2008) (Fig. 1A). We clarified the localization of this protein in *T. pseudonana* by generating and expressing a C-terminal GFP fusion. GFP fluorescence confirmed that the GOX2 protein was localized to the mitochondria (SFig. 1) as verified through co-localization with the mitochondrial citrate synthase (Fig. 4). Supporting a role in metabolizing phosphoglycolate, GOX2 transcript levels were co-induced with other known photorespiratory genes (Fig. 2; Hennon et al. 2015).

Glycolate oxidizing enzymatic activity in isolated diatom mitochondria suggested a preference for a secondary reaction with D-lactate, consistent with the presence of a GDH (Winkler and Stabenau 1992), which prompted us to search the *T. pseudonana* genome for candidates responsible for this activity. A putative GDH gene (Thaps3_42965) containing a GlcD domain with sequence similarity to FAD-linked oxidases and D-lactate dehydrogenases (BLAST e-value $1e^{-101}$) that has not previously been suggested as a candidate for glycolate oxidation was found. Homologs for this gene were also identified in *P. tricornutum* and *Phytophthora sojae*. The gene product of Thaps3_42965 had a weak Mitoprot score (0.312), but was predicted to be targeted to mitochondria by TargetP (Emanuelsson et al. 2007), and the *P. tricornutum* homolog had high predicted mitochondrial targeting by both programs. Thus, the gene product is a candidate for the mitochondrial GDH activity previously characterized biochemically (Winkler and Stabenau 1992). Although induced in our conditions, the GDH-like protein was not tightly coexpressed with other photorespiratory genes as was GOX2 (Fig. 2). It also did not cluster with other photorespiratory genes in other transcriptomics data (Hennon et al. 2015). Ultimately, we are unable to definitively assign a lone glycolate-oxidizing candidate to the mitochondria of *T. pseudonana* based on the sequence content, gene expression, or localization information. However, evidence supports at least two candidates for this reaction (Thaps3_3353 and Thaps3_42965), and it is possible that either may operate in diatom mitochondria depending on the specific set of conditions.

Recycling of Photorespiratory Intermediates into the Chloroplast

Whether or not diatoms recycle photorespiratory intermediates completely back to the chloroplasts has been debated (Fabris et al. 2012; Igamberdiev and Lea 2002; Kern et al. 2011; Kroth et al. 2008; Winkler and Stabenau 1995) due to an apparent lack of a plastidial glycerate

kinase (GLYK) (Kern et al. 2011; Kroth et al. 2008) (Fig. 1B). However, a more recent study using an orthology-based approach identified a GLYK candidate in *P. tricornutum* (Fabris et al. 2012). We found that although the homolog of this putative GLYK in *T. pseudonana* (Thaps3_261750) is similar to other GLYK sequences, it has substantially higher sequence similarity to glycerate-2-kinases (GKs) that produce glycerate 2-phosphate (2PGA) rather than 3PGA (Bartsch et al. 2008), including a biochemically characterized GK from the archaea *Thermoproteus tenax* (Kehrer et al. 2007) (Fig. 5). The Thaps3_261750 protein has predicted periplastid targeting (Gould et al. 2006) (Table 1), and our analysis suggests the same for the *P. tricornutum* (Phatr2_54867) homolog. In addition, the transcript for the putative GK is substantially downregulated, which is inconsistent with the other genes in the C2 pathway (Fig. 2), implying a reduction of glycerate processing by this enzyme under our experimental conditions. Taken together, the data does not support the idea that a GLYK returns photorespiratory carbon to diatom plastids.

Another route for reintroducing 3PGA back into the Calvin-Benson cycle is possible, which is consistent with the predicted location of enzymes. If a GK converts glycerate to 2PGA in the periplastid compartment, 2PGA could be transported into the chloroplast by an unidentified transporter (2PGA transporters have been characterized in bacteria) (Saier et al. 1975), then converted to 3PGA by a chloroplast-localized phosphoglycerate mutase (PGAM; (Thaps3_263397). Alternatively, a periplastid-localized PGAM (Thaps3_260919) could generate 3PGA which could be transported into the plastid. Arguing against these processes occurring under our experimental conditions is the fact that the GK transcript is highly downregulated.

Transport is another consideration in the completion of a C2 cycle. In higher plants, glycerate is transported into the plastid before it is converted into 3PGA by a GLYK. This import is mediated by a glycolate/glycerate transporter (PLGG1) in plastid membranes (Pick et al. 2013; Reumann and Weber 2006). We identified two genes with homology to plant PLGG1 (Thaps3_9352, Thaps3_14106) in *T. pseudonana* with predicted chloroplast and periplastid membrane targeting, respectively (Table 1). Although the PLGG1 transporter is commonly depicted as a glycolate/glycerate exchanger, it actually functions via a unidirectional proton symport mechanism and can recognize either substrate (as well as lactate and glyoxylate) and transport them in either direction (Young and McCarty 1993). Thus there is no a priori reason to correlate the presence of PLGG1 with a requirement for glycerate uptake into the chloroplast

since the transporter may simply function as a glycolate exporter. The lack of predicted chloroplast targeted enzymes to process glycerate in the chloroplast is consistent with this possibility.

The lack of an identifiable GLYK candidate and downregulation of GK are consistent with diatoms using a distinct mechanism for recovering photorespiratory C2 rather than recycling it directly back into the Calvin-Benson cycle. This led us to investigate other possible routes for reprocessing carbon in photorespiratory metabolism in *T. pseudonana*.

Processing of Photorespiratory Glyoxylate in the Mitochondria

The metabolism of glyoxylate by C2 cycle enzymes is predicted to occur primarily in the mitochondria of diatoms (Kroth et al. 2008) (Table 1, Fig. 1). It has been proposed (Kroth et al. 2008) that mitochondrial glyoxylate is converted to glycine by a serine-pyruvate/alanine-glyoxylate aminotransferase (SPT/AGAT), followed by conversion to serine by stepwise reactions of glycine decarboxylase (GDC) and serine hydroxymethyltransferase (SHMT). Exhibiting substrate flexibility, the same SPT/AGAT is proposed to use serine to generate hydroxypyruvate (Fig. 1, gray box), which is further converted to glycerate by a hydroxypyruvate reductase (HPR). In this scheme, SPT/AGAT would compete with itself by both generating and utilizing serine, which may be inefficient. In addition, careful biochemical analysis failed to identify HPR activity in diatom mitochondria, even under conditions when other photorespiratory enzymes were detected (Winkler and Stabenau 1995). These issues, plus the possible lack of importance of glycerate in recycling photorespiratory carbon (discussion above), prompted us to consider alternative means of processing of glyoxylate in diatom mitochondria.

We identified an additional transaminase similar to alanine aminotransferase/glutamate glyoxylate aminotransferase (ALAT/GGAT, Thaps3_412) that exhibited similar transcript changes as other mitochondrial C2 pathway genes (Fig. 2B). The Thaps v.3 gene model is prematurely truncated, but N-terminal extension using RNAseq data revealed a mitochondrial signal peptide (Table 1). A similar ALAT with GGAT activity catalyzes the transamination of glyoxylate in the photorespiratory pathway of *Arabidopsis* (Liepman and Olsen 2003), and the presence of this enzyme is consistent with previous detection of GGAT activity in diatom mitochondria (Winkler and Stabenau 1995).

SPT/AGAT and ALAT/GGAT are both capable of carrying out multiple reactions with different substrates. The lack of HPR activity (Winkler and Stabenau 1995) suggests no need for the reaction by SPT/AGAT that converts glyoxylate and serine to hydroxypyruvate and glycine, therefore, we propose that this enzyme (Thaps3_22208) carries out its alternative reaction by converting glyoxylate and alanine to pyruvate and glycine (Fig. 1C). Induction of transcripts (Fig. 2B) for GDC-T (Thaps3_36208) and SHMT (Thaps3_269942) suggest that glycine is further converted to serine (Fig. 1). We propose that the ALAT/GGAT (Thaps3_412) catalyzes the conversion of pyruvate and glutamate to α -ketoglutarate and alanine. The transcript pattern for the ALAT/GGAT, GDC-T, SHMT, and SPT/AGAT are similar, with SPT/AGAT having a more prolonged upregulation than the others (Fig. 2B). The end results of this proposed scheme are that glycine and serine would be the primary amino acids synthesized from photorespiratory C₂, and that α -ketoglutarate levels would also be increased under photorespiratory conditions. Diatom labeling studies documented specific labeling of serine and glycine as a result of photorespiratory processes, and a substantial increase in labeling of α -ketoglutarate relative to non-photorespiratory conditions (Huang et al. 2015). Thus, the model is entirely consistent with labeling studies (Huang et al. 2015; Zheng et al. 2013).

This model and biochemical evidence (Winkler and Stabenau 1995) predict that hydroxypyruvate is not formed in the mitochondria via photorespiratory processes. A mitochondrially-targeted HPR was previously proposed (Kroth et al. 2008), but our analysis indicated that it was the previously discussed GK (Fig. 5). We identified another *T. pseudonana* gene, Thaps3_2846, with weak similarity to plant and human HPRs, but high similarity to bacterial 2-hydroxyacid dehydrogenases (2-HADHs). 2-HADHs can perform diverse functions in many aspects of metabolism, and previous phylogenetic analyses have demonstrated that members of the HADH family clade well according to substrate specificity (Fauvart et al. 2007). In an effort to better annotate the 2-HADH proteins in the *T. pseudonana* genome, we generated a sequence alignment that includes 2-HADH proteins that have experimentally verified substrate specificity. We observed that the diatom 2-HADH sequences, including Thaps_2846, fall within a large and diverse group containing glyoxylate/hydroxypyruvate reductases, 2-ketogluconate reductases, and phenylpyruvate reductases (Supplementary Material Fig. S2), thus its substrate specificity still remains unclear. However, comparison of conserved amino acids known to be functionally important based on the human HPR 3D structure indicate that two critical residues

(W-141 and S-296) are missing in the diatom 2-HADHs (data not shown), arguing against HPR activity. The transcript pattern for Thaps3_2846 clusters tightly with PGP and GOX2 (Fig. 2), suggesting that the enzyme may instead play a role in glyoxylate reduction. In addition, these 2-HADHs have a preference for D-stereoisomers, which is consistent with the D-lactate preference in the mitochondria (Winkler and Stabenau 1992). Confirmation of the exact activities and role of Thaps3_2846 will require experimental validation.

Peroxisomal Processing of Glycolate Through the Glyoxylate Cycle

The glyoxylate cycle offers an alternate route to convert glycolate into useful biosynthetic precursors (Kunze et al. 2006). Glyoxylate is produced in the peroxisome by isocitrate lyase (ICL, Thaps3_35523) during β -oxidation, and can be produced from photorespiratory glycolate by GOX1, (Thaps3_406). Glyoxylate combined with acetyl-CoA forms malate by the activity of malate synthase (MS, Thaps3_26293) (Fig. 1D). Biochemical studies showed that MS and ICL enzyme activities are induced under conditions of high irradiance and low CO₂ availability, suggesting that the glyoxylate cycle is involved in processing glycolate in the peroxisome (Winkler and Stabenau 1995). Our transcript data show coordinate upregulation of MS, ICL and malate dehydrogenase (MDH, Thaps3_41425), but not GOX1, suggesting that the glyoxylate cycle is responsive to our conditions, but that the formation of glyoxylate via photorespiratory glycolate may not be a prominent contributor to glyoxylate formation. Alternatively, GOX1 may be constitutively expressed and not regulated at the transcript level. The GOX1 transcript response also does not cluster with the photorespiratory response in other datasets (Hennon et al. 2015). To substantiate the possible role of the glyoxylate cycle, we analyzed and experimentally verified the location of enzymes involved in it. MS sequences from both *T. pseudonana* and *P. tricornutum* were analyzed and found to contain canonical C-terminal PTS1 signals recognized by a Pex5 receptor protein that facilitates transport into the peroxisome (Rucktäschel et al. 2011). In isolated peroxisomes of the diatom *Fragilaria* ¹⁴C-labeled malate is produced from ¹⁴C-labeled glycolate, supporting the localization of this reaction (Winkler and Stabenau 1992). For these reasons, we used MS as a peroxisomal marker. *T. pseudonana* cells transformed with GFP-tagged MS produced labeling in two closely-associated spots per cell (Fig. 6) that were similar in size ($\leq 1\mu\text{m}$) as peroxisomes described in other organisms (Van den Bosch and Schutgens 1992).

While significant efforts have advanced our understanding of the signals that direct proteins into secondarily-derived plastids of diatoms (Apt et al. 2002; Gruber et al. 2007, 2015), signals that dictate targeting to other cellular compartments, such as the peroxisome, are just beginning to emerge (Gonzalez et al. 2011). In plants, ICL is targeted to the peroxisome through PTS1, but ICL is known to be cytosolic in some fungi (Kunze et al. 2006). *T. pseudonana* ICL (Thaps3_35523) lacks the PTS1 sequence (Table 1), yet expression of GFP-fused ICL resulted in similar labeling as that observed for MS, supporting peroxisomal localization that was expected based on previous biochemical studies (Fig. 6) (Winkler and Stabenau 1995). Diatoms appear to lack alternate peroxisomal import systems that have been described in other organisms (i.e. PTS2 targeting) (Gonzalez et al. 2011). Therefore, the mechanism targeting ICL to diatom peroxisomes remains unclear and points to the possibility of the presence of an uncharacterized peroxisomal import signal within the ICL sequence. In support of this possibility, so-called non-PTS proteins have been described, including MS and ICL from *Candida albicans* (Piekarska et al. 2008) and ICL from castor bean (Parkes et al. 2003), that are imported to peroxisomes in a PTS1-independent but Pex5-dependent manner (Rucktäschel et al. 2011; van der Klei and Veenhuis 2006). The in vivo localization of ICL demonstrates the presence of key glyoxylate cycle enzymes in peroxisomes of *T. pseudonana*.

In higher plants, the glyoxylate cycle utilizes acetyl-CoA generated by the β -oxidation of lipids, for example, during mobilization of storage lipids during seed germination (Doubnerová 2010) or during senescence due to periods of prolonged dark exposure when endomembrane systems are degraded (Kunze et al. 2006). Therefore, the glyoxylate cycle may be an important mechanism for the repackaging of acetyl-CoA from oxidized lipids. The breakdown and repackaging of membrane lipids is a documented response in diatoms during nutrient-limited growth arrest and neutral lipid accumulation (Yang et al. 2013).

The Anapleurotic Role of Malic Enzyme

Malate or OAA generated by the glyoxylate cycle may be transported into the mitochondria and enter the TCA cycle directly (Igamberdiev and Lea 2002), or be converted into pyruvate via the action of malic enzyme (Fig. 1E) (Plaxton and Podestá 2007; Sweetlove et al. 2010). Malic enzyme is a C4 decarboxylase with diverse functional roles in photosynthetic eukaryotes. Malic enzyme isozymes use either NAD^+ or NADP^+ as a cofactor and can be targeted to the cytosol,

chloroplast, and/or the mitochondria. While a chloroplast-localized NADP-dependent ME would be more consistent with roles associated with C₄ photosynthesis and/or fatty acid biosynthesis, a cytosolic or mitochondrial NAD-dependent ME is less likely to be involved in these pathways. Previous analyses of ME amino acid sequence suggests that NAD may be the preferred cofactor for diatom MEs (Kroth et al. 2008). To confirm this analysis, we compared NAD(P) binding sites from diatom MEs (Supplementary Material Table S1) and from several manually curated and biochemically validated MEs from the Uniprot database (Supplementary Material Fig. S3). Our analysis demonstrates that the *T. pseudonana* ME (Thaps3_34030) nucleotide binding sequence groups with other known NAD-specific binding sites, as do most of the other diatom sequences. A notable exception is a second *P. tricornutum* isoform (Pt_51970), which groups with NADP-dependent ME sequences found in plants, highlighting the possibility that this ME may serve a different functional role in *P. tricornutum* (Smith et al. 2012).

Conflicting information concerning the intracellular localization of diatom ME exists in the literature. Although the Thaps3_34030 homolog found in *P. tricornutum* also has a mitochondrial targeting sequence (Kroth et al. 2008), experimental localization in *T. pseudonana* suggested this protein was cytosolic (Tanaka et al. 2014). With additional RNAseq support, we extended the N-terminus of the protein model based on RNAseq coverage to uncover a mitochondrial signal (Supplementary Material Fig. S4). We also generated a GFP fusion construct using native promoter/terminator regions for ME and observed colocalization to the mitochondria with the TCA cycle enzyme citrate synthase (CS; Thaps3_11411) (Fig. 7; Supplementary Material Fig. S1).

Both the similarity to NAD-dependent enzymes and the mitochondrial localization of ME suggests an anapleurotic role through the provision of pyruvate for the synthesis of TCA cycle intermediates. This is consistent with previous transcriptomic, proteomic, and enzyme activity data that suggest ME is not directly involved in a biochemical carbon concentration mechanism (Kustka et al. 2014; Clement et al. 2015). In addition, coordinate expression of ME with photorespiratory (Hennon et al. 2015) and glyoxylate cycle genes (Fig. 2C) suggest this enzyme may have an important supporting role in mitochondrial metabolism.

In plants, the activity of mitochondrial ME maintains TCA cycle function during conditions of low glycolytic flux and reduced pyruvate availability (Kromer, 1995). Under such scenarios, ME generates an alternative source of pyruvate for acetyl-CoA formation and utilization in the

TCA cycle (Sweetlove et al., 2010). This arrangement conveys some metabolic flexibility, as it allows for malate to serve as the sole substrate for the TCA cycle if necessary (Sweetlove et al. 2010). While the contribution of ME-derived pyruvate is typically small, flux through this pathway is significant under certain conditions, including during fatty acid accumulation in oilseed embryos (Plaxton and Podestá 2007; Schwender et al. 2004, 2006; Sweetlove et al. 2010). In our conditions, malate produced by the glyoxylate cycle may be routed to mitochondria where it can serve a similar role in supplementing the TCA cycle. Furthermore, the observed increase in glyoxylate cycle and ME transcripts during silicon starvation supports the idea that ME serves as a connection between peroxisomal metabolism and the repurposing of carbon for biosynthetic precursors. Increases in TAG and chlorophyll content are hallmarks of silicon starvation in *T. pseudonana* (Smith et al. 2016), and these carbon-intensive processes may reduce pyruvate availability for mitochondrial respiration and increase the flux of carbon skeletons (in the form of TCA cycle intermediates) out of this organelle. Malate, coupled with the activity of ME, can serve as an alternative source of organic carbon with the ability to support biosynthetic and respiratory processes of mitochondrial metabolism during silicon starvation in diatoms.

Implications of the Model

Physiological data (Smith et al. 2016), coupled with an induction of photorespiratory transcripts, suggests that photorespiration may be induced as a mechanism to manage an imbalance in cellular reducing potential in our experimental conditions. Recent evidence suggests that diatoms maintain favorable ATP/NADPH ratios in the chloroplast through the transfer of reducing equivalents to mitochondria, where reductive potential can be used to drive ATP production or be dissipated in the electron transport chain via transfer to O₂ (Allen et al. 2006, 2008; Bailleul et al. 2015; Hockin et al. 2011; Kim et al. 2015; Prihoda et al. 2012). The latter process involves the activity of alternative oxidase (AOX), which transfers electrons directly to molecular oxygen without contributing to ATP synthesis (Bailleul et al. 2015; Vanlerberghe 2013). In plants, AOX is subject to many layers of regulation including redox regulation and activation by pyruvate and photorespiratory intermediates (Vanlerberghe 2013), suggesting the possibility that the buildup of these metabolites in diatom mitochondria may also signal AOX that the cellular reducing status needs to be adjusted. The details of a mechanism for transferring reducing potential between the chloroplast and the mitochondria remain unclear (Bailleul et al. 2015), although

dihydroxyacetone phosphate (DHAP) was recently put forth as a possible candidate carrier molecule (Kim et al. 2015). Given our identification of three possible candidates for glycolate oxidation in the mitochondria of *T. pseudonana* (Thaps3_3353, Thaps3_42965, Thaps3_2846), including confirmation of mitochondrial localization for GOX2 (Fig. 4, SFig1A), we suggest glycolate may also participate in the transport of reducing potential and carbon skeletons from the chloroplast to the mitochondria under certain conditions (Fig 1A), and may also support active energy dissipating mechanisms (such as AOX) within the mitochondria. Consistent with this was the transcript-level induction (1.5-fold) of a mitochondrial AOX (Thaps3_42992) in our experimental conditions.

Our analysis does not support the presence of a GLYK or a convincing HPR in diatom genomes (Fig. 5), and is in agreement with previous assertions that the C2 cycle is incomplete (Igamberdiev and Lea 2002; Kroth et al. 2008). The implication of this finding is that photorespiratory carbon is repurposed initially into other metabolic processes, rather than being returned directly to the Calvin-Benson cycle. Instead, we propose that the transaminases SPT/AGAT (Thaps3_22208) and ALAT/GGAT (Thaps3_412) work in conjunction to produce precursors for amino acid biosynthesis (Fig. 1C). This alternative approach simultaneously allows for the dissipation of energy and the repurposing of carbon skeletons within the same organelle and may be especially advantageous under extreme conditions when the rate of carbon fixation (and therefore the requirement for 3PGA in the plastid) is low.

Additionally, we suggest that peroxisomal processes may be interconnected with photorespiratory metabolism in diatoms, and this multi-compartmentation may afford some flexibility in responding to different environmental conditions. For example, mitochondrial photorespiratory processes permitted cells to satisfy carbon backbone demand for reassimilation of nitrogen released from intracellular stores (Huang et al. 2015). Consistent with this was down-regulation of enzymes involved in the glyoxylate cycle during nitrogen limitation (Yang et al. 2013). Alternatively, when diatoms were cultivated in low CO₂ conditions, biochemical studies found that glycolate oxidation activity was substantially increased in the peroxisome as were glyoxylate cycle enzymes (Winkler and Stabenau 1995). This suggests carbon-limited cells salvage photorespiratory glycolate and convert it into malate, which can be metabolized through other pathways, minimizing carbon and energy losses associated with the C2 cycle (Winkler and Stabenau 1992). The differential biochemical response in the mitochondria and peroxisome to

moderate versus severe photorespiration (Winkler and Stabenau 1995) coupled with transcriptome analysis (Hennon et al. 2015) (Fig. 2) suggest that primary processing of photorespiratory carbon occurs in the mitochondrion, and that peroxisomal processing can contribute as an “overflow valve” when conditions are more severe. Collectively these findings provide a framework through which we can interpret the response of diatoms during silicon starvation-induced growth arrest.

In our conditions, we observe induction of transcripts likely involved in mitochondrial glycolate oxidation (Fig. 2A), including a mitochondrial GOX2 (Fig. 4), concomitant with a high light shift and the onset of cell cycle arrest. The transport and subsequent oxidation of glycolate in the mitochondria may be one strategy to mitigate damage from overreduction of the chloroplast. Additionally, we observe an increase in the transcripts of genes that likely participate in the conversion of glyoxylate to glycine, serine, and TCA cycle intermediates (Fig. 2B). Due to a lack of cell division, cellular energy demands are reduced, facilitating a non-cyclic mode for the TCA cycle that feeds isocitrate into the glyoxylate cycle. The induction of peroxisomal glyoxylate cycle enzymes (Figs 2C, 6; Supplementary Material Fig. S1B) may participate in salvaging carbon produced during β -oxidation and membrane remodeling that occurs during lipid accumulation in diatoms (Yang et al. 2013). This situation also permits decarboxylating and NADH producing steps of the TCA cycle to be bypassed, thereby limiting carbon loss and excess energy production. The entry of carbon into the mitochondria as malate permits cells to utilize carbon from non-glycolytic sources to provision the TCA cycle. This is accomplished through the action of a mitochondrial ME (Fig. 7; Supplementary Material Fig. S1A), which supplies the pyruvate needed to replenish carbon skeletons via the partial TCA cycle (Hanning and Heldt 1993; Sweetlove et al. 2010; Turpin et al. 1988). In this way, our model for silicon starvation suggests that both mitochondrial photorespiratory processes and the peroxisomal glyoxylate cycle contribute carbon for biosynthetic purposes, unlike the recycling of photorespiratory carbon back into the Calvin-Benson cycle that has been described in plants.

The model presented in Figure 1 places our results in the context of earlier research and clarifies areas of ambiguity from previous models. Taken together, our data supports the idea that photorespiratory intermediates processed in the mitochondria are routed into amino acid metabolism rather than back into the chloroplast. In addition, we suggest that malate derived from the peroxisomal glyoxylate cycle may support the operation of the TCA cycle in certain

conditions through the action of malic enzyme. This organizational scheme is a step towards understanding these unique carbon processing pathways in diatoms and serves to provide a framework for future investigation.

Methods

Cultivation: Cultivation for expression analysis was performed as previously described (Smith et al. 2016). Briefly, axenic 8L cultures of *Thalassiosira pseudonana* (CCMP1335) were grown in artificial seawater medium (ESAW, <http://www3.botany.ubc.ca/cccm/NEPCC/esaw.html>) at 18 °C under continuous light ($150 \mu\text{mol} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$) to a concentration of approximately 1×10^6 cells $\cdot \text{ml}^{-1}$ and then harvested by centrifugation for 12 min at 3100 x g and placed in 8L of silicic acid free medium in a polycarbonate bottle at a concentration of approximately 5×10^5 cells $\cdot \text{ml}^{-1}$. Cultures were stirred and bubbled with air under constant continuous light. Cultures were sampled at 0, 4, 8, 12, 18, and 24-hour time points following inoculation into silicic acid free medium.

Transcriptomics: RNA for transcriptomic analysis was extracted from *T. pseudonana* cells during a 24-hour time course of silicon limitation. Details of the experimental design, RNA sequencing, and processing are described in Smith et al. 2016. Briefly, cells were treated with cycloheximide (final concentration, $20 \mu\text{g ml}^{-1}$), harvested by filtration, pelleted and stored at -80 °C prior to total RNA isolation (Hildebrand and Dahlin 2000). RNA was processed either for hybridization to an Affymetrix GeneChip whole genome tiling array or processed for Illumina-based RNA-Seq.

Bioinformatic analyses: The genome of *Thalassiosira pseudonana* (<http://genome.jgi-psf.org/Thaps3/Thaps3.home.html>) was searched for genes involved in photorespiratory metabolism and carbon concentration based on BLAST similarity with previously annotated sequences available from the Uniprot database. RNAseq reads generated during duplicate silicon starvation experiments (described in Smith et al. 2016) were mapped against the *T. pseudonana* genome version 3 using TopHat2 (Kim et al. 2013). Alignment files of uniquely mapped reads in BAM format were used to create index files using the Bam to Bai function available through the Galaxy platform (Blankenberg et al. 2010; Giardine et al. 2005; Goecks et al. 2010). The *T.*

pseudonana genome, existing version 3 gene models, and the mapped reads were visualized together on the Integrated Genomics Viewer (Robinson et al. 2011; Thorvaldsdóttir et al. 2012) to assess the reliability of each gene model. A FASTA file of evaluated protein sequences used in this study is available in the supplemental material (Supplementary Material Fig. S5). When necessary, the N-terminus of the gene was extended to the first in-frame methionine to recover any excluded targeting information. Several bioinformatic tools were used to predict subcellular localization, including SignalP (Bendtsen et al. 2004; Nielsen and Krogh 1998), ChloroP (Emanuelsson et al. 1999), ASAFind (Gruber et al. 2015), MitoProt (Claros and Vincens 1996), and PTS1 Predictor (Neuberger et al. 2003a, b). In diatoms, nuclear encoded proteins targeted to the chloroplast have a bipartite targeting signal necessary for import across several membrane systems (Apt et al. 2002). Proteins predicted to have a signal peptide (SignalP) and a chloroplast transit-peptide (ChloroP), but not meeting the sequence requirements for stromal import (ASAFind) were predicted to have periplastid localization. An assessment of the conservation of targeting predictions from available diatom sequences was attempted, and in many cases supported our observations from *P. tricornutum* and *T. pseudonana*. However, the variability in accurate gene model predictions, especially at the gene model boundaries, precludes confidence in drawing rigorous conclusions.

Phylogenetic analyses: Homologous sequences were selected by BLAST search of the NCBI nr protein database (<http://www.ncbi.nlm.nih.gov>). Sequences from model organisms that had been manually annotated and reviewed (submitted to the UniprotKB/Swiss-Prot database) were included as points of reference. Diatom protein sequences were obtained from the JGI genome browser and (<http://genome.jgi-psf.org/cgi-bin/searchGM?db=Thaps3>). Sequences were aligned using MAFFT with default parameters (Katoh et al. 2002) and manually trimmed to conserved functional domains in Geneious version 8.1.2 (<http://www.geneious.com>) (Kearse et al. 2012). Maximum likelihood trees were constructed from the alignment in the MEGA 6 software (Tamura et al. 2013) using the best fit model (LG+G) with complete gap deletion and 1,000 bootstrap replicates. For the ME tree, sequences were trimmed to the previously characterized nicotinamide-binding region (Scrutton et al. 1990) and a maximum likelihood tree was constructed using PHYML (Guindon et al. 2010) with 1000 bootstrap replicates. We were not able to produce a meaningful alignment between the GK and GLYK proteins, and therefore decided it was not appropriate to include them in the same phylogenetic tree. In order to

represent the idea that GK and GLYK sequences are nonhomologous proteins, we generated a hierarchically clustered heatmap from the percent identity matrix using the gplots package (Warnes et al. 2015) in R (R Core Team 2015).

Vector construction for protein localization: Destination and expression vectors were constructed as described in Shrestha and Hildebrand (2015) using MultiSite Gateway Technology (Life Technologies). Briefly, the constitutive expression vector pTpfcgGFP (Poulsen et al. 2006) was modified by cloning a Gateway frame B cassette upstream of *egfp* creating a destination vector. In cases where higher expression levels were required for *in vivo* localization, another destination vector was created utilizing the regulatory elements that drive constitutive expression of the ribosomal protein L41 (Shrestha et al. 2013) instead of FCP. Native expression constructs were created by incorporating 1000 bp of the 5' flanking region and 500 bp of the 3' flanking region of a gene of interest into the pMHL_71 vector to provide the native regulatory elements as described in Shrestha and Hildebrand (2015). Primers used to amplify genes and regulatory elements for construct synthesis are listed in Supplementary Material Table S2. Each expression vector was co-transformed with pMHL_9 expressing the *nat1* gene under the control of the acetyl coenzyme A carboxylase promoter, which confers resistance to the antibiotic nourseothricin.

Diatom transformation: The protocol for diatom transformation was adapted from Poulsen et al., 2006. Axenic exponential-phase wild type *T. pseudonana* cells (1×10^8 cells total) were pelleted ($3,000 \times g$ for 10 minutes) and plated on artificial seawater medium (ASW) 1.2% agar plates. M-17 tungsten particles were coated with 5 μ g of plasmid DNA for each construct with the CaCl_2 -spermidine method as per manufacturer's instructions (BIORAD, 165-2267). Each plate was bombarded twice with 2-5 mg of coated tungsten beads using the Biolistic DS-1000/He particle delivery system with a 1350 psi rupture disk. Following bombardment, cells were immediately resuspended in 10 ml of ASW medium and incubated for 24 hours under constant illumination ($150 \text{ uE} \cdot \text{m}^{-2} \cdot \text{s}^{-1}$). The following day, the cell density was determined and a range of concentrations (1×10^6 - 5×10^6 cells) were plated on ASW agar plates containing the antibiotic nourseothricin at a final concentration of $100 \text{ ug} \cdot \text{ml}^{-1}$. The plates were incubated at 18°C in continuous light for approximately 2 weeks. Individual colonies from the plates were isolated and screened by PCR and imaging flow cytometry to confirm nuclear insertion of the plasmid and for GFP expression, respectively.

Fluorescence microscopy: Diatom transformants were imaged with a Zeiss Axio Observer Z1 inverted microscope equipped with an ApoTome and a Zeiss AxioCam MRm camera (Carl Zeiss Microimaging, Inc., Thornwood, NY, USA). Non-fluorescent images were taken using differential interference contrast (DIC). The filter sets used for fluorescent imaging were as follows: chlorophyll (Zeiss #16: Excitation BP 485/20 nm, Dichromatic mirror FT 510 nm, Emission LP 515 nm), GFP (Zeiss #38HE Excitation BP 470/40 nm, Dichromatic mirror FT 495 nm, Emission BP 525/50 nm), and RFP (Zeiss #43HE Excitation BP 550/25 nm, Dichromatic mirror FT 570 nm, Emission BP 605/70 nm). Images were acquired with 63x/1.4 objective oil immersion Plan-Apochromat objective and processed using Axiovision 4.7.2 software.

Acknowledgements

This research was funded by Department of Energy Grants DE-SC0012556 and DE-EE0001222 and National Science Foundation grant CBET-0903712. Additional support was provided through a Department of Energy Office of Science Graduate Fellowship Program (DOE SCGF), made possible in part by the American Recovery and Reinvestment Act of 2009, administered by ORISE-ORAU under contract no. DE-AC05-06OR23100.

References

- Allen A** (2005) Defining the molecular basis for energy balance in marine diatom under fluctuating environmental conditions. *J Phycol* 41:1073–1076
- Allen AE, Vardi A, Bowler C** (2006) An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Curr Opin Plant Biol* 9:264–273
- Allen AE, LaRoche J, Maheswari U, Loomer M, Schauer N, Lopez PJ, Finazzi G, Fernie AR, Bowler C** (2008) Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *Proc Natl Acad Sci USA* 105:10438–10443
- Apt KE, Zaslavkaia L, Lippmeier JC, Lang M, Kilian O, Wetherbee R, Grossman AR,**

682 **Kroth PG** (2002) In vivo characterization of diatom multipartite plastid targeting signals. J
683 Cell Sci **115**:4061–4069

684 **Bailleul B, Berne N, Murik O, Petroutsos D, Prihoda J, Tanaka A, Villanova V, Bligny R,**
685 **Flori S, Falconet D, Krieger-Liskay A, Santabarbara S, Rappaport F, Joliot P,**
686 **Tirichine L, Falkowski P, Cardol P, Bowler C, Finazzi C** (2015) Energetic coupling
687 between plastids and mitochondria drives CO₂ assimilation in diatoms. Nature **524**:366–369

688 **Bartsch O, Hagemann M, Bauwe H** (2008) Only plant-type (GLYK) glycerate kinases produce
689 d-glycerate 3-phosphate. FEBS Lett **582**:3025–3028

690 **Bauwe H, Hagemann M, Fernie AR** (2010) Photorespiration: players, partners and origin.
691 Trends Plant Sci **15**:330–336

692 **Bauwe H, Hagemann M, Kern R, Timm S** (2012) Photorespiration has a dual origin and
693 manifold links to central metabolism. Curr Opin Plant Biol **15**:269–275

694 **Bendtsen JD, Nielsen H, von Heijne G, Brunak S** (2004) Improved prediction of signal
695 peptides: SignalP 3.0. J Mol Biol **340**:783–795

696 **Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrutenko**
697 **A, Taylor J** (2010) Galaxy: A web-based genome analysis tool for experimentalists. Curr
698 Prot Mol Biol **89**:19.10:19.10.1–19.10.21

699 **Claquin P, Martin-Jezequel V, Kromkamp JC, Veldhus MJ, Kraay GW** (2002) Uncoupling
700 of silicon compared with carbon and nitrogen metabolisms and the role of the cell cycle in
701 continuous cultures of *Thalassiosira pseudonana* (Bacillariophyceae) under light, nitrogen,
702 and phosphorus control. J Phycol **38**:922–930

703 **Claros MG, Vincens P** (1996) Computational method to predict mitochondrially imported
704 proteins and their targeting sequences. Eur J Biochem **241**:779–786

705 **Clement R, Dimnet L, Maberly SC, Gontero B** (2015) The nature of the CO₂-concentrating
706 mechanisms in a marine diatom, *Thalassiosira pseudonana*. New Phytol **209**:1417–1427

707 **Eisenhut M, Ruth W, Haimovich M, Bauwe H, Kaplan A, Hagemann M** (2008) The
708 photorespiratory glycolate metabolism is essential for cyanobacteria and might have been
709 conveyed endosymbiontically to plants. Proc Natl Acad Sci USA **105**:17199–17204

710 **Emanuelsson O, Nielsen H, von Heijne G** (1999) ChloroP, a neural network-based method for
711 predicting chloroplast transit peptides and their cleavage sites. Prot Sci **8**:978–984

712 **Fabris M, Matthijs M, Rombauts S, Vyverman W, Goossens A, Baart GJE** (2012) The
713 metabolic blueprint of *Phaeodactylum tricornutum* reveals a eukaryotic Entner-Doudoroff
714 glycolytic pathway. Plant J **70**:1004–1014

715 **Fauvart M, Braeken K, Daniels R, Vos K, Ndayizeye M, Noben J, Robben J, Vanderleyden**
716 **J, Michiels J** (2007) Identification of a novel glyoxylate reductase supports phylogeny-

717 based enzymatic substrate specificity prediction. *Biochim Biophys Acta Prot Proteom*
718 **1774**:1092–1098

719 **Fernie AR, Tiessen A, Stitt M, Willmitzer L, Geigenberger P** (2002) Altered metabolic fluxes
720 result from shifts in metabolite levels in sucrose phosphorylase-expressing potato tubers.
721 *Plant Cell Environ* **25**:1219-1232

722 **Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y,**
723 **Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A** (2005) Galaxy: A
724 platform for interactive large-scale genome analysis. *Genome Res* **15**:1451–1455

725 **Goecks J, Nekrutenko A, Taylor J** (2010) Galaxy: a comprehensive approach for supporting
726 accessible, reproducible, and transparent computational research in the life sciences.
727 *Genome Biol* **11**:1-13

728 **Gould SB, Sommer MS, Hadfi K, Zauner S, Kroth PG, Maier U** (2006) Protein targeting into
729 the complex plastid of cryptophytes. *J Mol Evol* **62**:674–681

730 **Granum E, Raven JA, Leegood RC** (2005) How do marine diatoms fix 10 billion tonnes of
731 inorganic carbon per year? *Can J Bot* **83**:898–908

732 **Gruber A, Vugrinec S, Hempel F, Gould SB, Maier U, Kroth PG** (2007) Protein targeting
733 into complex diatom plastids: functional characterization of a specific targeting motif. *Plant*
734 *Mol Biol* **64**:519-530

735 **Gruber, A, Rocap G, Kroth PG, Armbrust EV, Mock T** (2015) Plastid proteome prediction
736 for diatoms and other algae with secondary plastids of the red lineage. *Plant J* **81**:519–528

737 **Guindon S, Dufayard J, Lefort V, Anisimova M, Hordijk W, Gascuel O** (2010) New
738 algorithms and methods to estimate maximum-likelihood phylogenies: assessing the
739 performance of PhyML 3.0. *Syst Biol* **59**:307–321

740 **Hanning I, Heldt HW** (1993) On the function of mitochondrial metabolism during
741 photosynthesis in spinach (*Spinacia oleracea* L.) leaves: Partitioning between respiration
742 and export of redox equivalents and precursors for nitrate assimilation products. *Plant Phys*
743 **103**:1147–1154

744 **Hennon GMM, Ashworth J, Groussman RD, Berthiaume C, Morales RL, Baliga NS,**
745 **Orellana MV, Armbrust EV** (2015) Diatom acclimation to elevated CO₂ via cAMP
746 signalling and coordinated gene expression. *Nat Clim Change* **5**:761–765

747 **Hildebrand M, Dahlin K** (2000) Nitrate transporter genes from the diatom *Cylindrotheca*
748 *fusiformis* (Bacillariophyceae): mRNA levels controlled by nitrogen source and by the cell
749 cycle. *J Phycol* **36**:702-713

750 **Hockin N, Mock T, Mulholland F** (2011) The response of diatom central carbon metabolism to
751 nitrogen starvation is different to that of green algae and higher plants. *Plant Phys* **158**:299-
752 312

753 **Huang A, Liu L, Yang C, Wang G** (2015) *Phaeodactylum tricornutum* photorespiration takes
754 part in glycerol metabolism and is important for nitrogen-limited response. *Biotechnol*
755 *Biofuels* **8**:1-16

756 **Igamberdiev AU, Lea PJ** (2002) The role of peroxisomes in the integration of metabolism and
757 evolutionary diversity of photosynthetic organisms. *Phytochemistry* **60**:651–674

758 **Igamberdiev AU, Bykova NV, Lea PJ, Gardestrom P** (2001) The role of photorespiration in
759 redox and energy balance of photosynthetic plant cells: A study with a barley mutant
760 deficient in glycine decarboxylase. *Physiologia Plantarum* **111**:427–438

761 **Katoh K, Misawa K, Kuma K, Miyata T** (2002) MAFFT: a novel method for rapid multiple
762 sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**:3059–3066

763 **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A,**
764 **Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A** (2012)
765 Geneious Basic: An integrated and extendable desktop software platform for the
766 organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649

767 **Kehrer D, Ahmed H, Brinkmann H, Siebers B** (2007) Glycerate kinase of the
768 hyperthermophilic archaeon *Thermoproteus tenax*: new insights into the phylogenetic
769 distribution and physiological role of members of the three different glycerate kinase classes.
770 *BMC Genomics* **8**:1-18

771 **Kern R, Bauwe H, Hagemann M** (2011) Evolution of enzymes involved in the
772 photorespiratory 2-phosphoglycolate cycle from cyanobacteria via algae toward plants.
773 *Photosyn Res* **109**:103–114

774 **Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL** (2013) TopHat2: accurate
775 alignment of transcriptomes in the presence of insertions, deletions and gene fusions.
776 *Genome Biol* **14**:1-13

777 **Kim J, Fabris M, Baart G, Kim MK, Goossens A, Vyverman W, Falkowski PG, Lun DS**
778 (2015) Flux balance analysis of primary metabolism in the diatom *Phaeodactylum*
779 *tricornutum*. *Plant J* **85**:161–176

780 **Kromer S** (1995) Respiration during photosynthesis. *Annu Rev Plant Phys Plant Mol Biol*
781 **46**:45-70

782 **Kroth PG, Chiovitti A, Gruber A, Martin-Jezequel V, Mock T, Schnitzler Parker M,**
783 **Stanley MS, Kaplan A, Caron L, Weber T, Maheswari U, Armbrust EV, Bowler C**
784 (2008) A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum*
785 deduced from comparative whole genome analysis. *PloS ONE* **3**:e1426

786 **Kunze M, Pracharoenwattana I, Smith SM, Hartigm A** (2006) A central role for the
787 peroxisomal membrane in glyoxylate cycle function. *Biochim Biophys Acta Mol Cell Res*
788 **1763**:1441–1452

789 **Kustka AB, Milligan AJ, Zheng H, New AM, Gates C, Bidle KD, Reinfelder JR** (2014) Low
790 CO₂ results in a rearrangement of carbon metabolism to support C₄ photosynthetic carbon
791 assimilation in *Thalassiosira pseudonana*. *New Phytologist* **204**:507-520

792 **Lancien M, Ferrario-Mery S, Roux V, Bismuth E, Masclaux C, Hirel B, Gadal P, Hodges**
793 **M** (1999) Simultaneous expression of NAD-dependent isocitrate dehydrogenase and other
794 Krebs cycle genes after nitrate resupply to short-term nitrogen starved tobacco. *Plant Physiol*
795 **120**:717-725

796 **Mittler R** (2002) Oxidative stress, antioxidants and stress tolerance. *Trends Plant Sci* **7**:405–410

797 **Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F** (2003a) Motif
798 refinement of the peroxisomal targeting signal 1 and evaluation of taxon-specific differences.
799 *J Mol Biol* **328**:567–579

800 **Neuberger G, Maurer-Stroh S, Eisenhaber B, Hartig A, Eisenhaber F** (2003b) Prediction of
801 peroxisomal targeting signal 1 containing proteins from amino acid sequence. *J Mol Biol*
802 **328**:581–592

803 **Nielsen H, Krogh A** (1998) Prediction of signal peptides and signal anchors by a hidden Markov
804 model. *Proc Int Conf Intell Sys Mol Biol* **6**:122-130

805 **Ohno N, Inoue T, Yamashiki R, Nakajima K, Kitahara Y, Ishibashi M, Matsuda Y** (2012)
806 CO₂-cAMP-responsive cis-elements targeted by a transcription factor with CREB/ATF-like
807 basic zipper domain in the marine diatom *Phaeodactylum tricornutum*. *Plant Physiol*
808 **158**:499–513

809 **Parker M, Armbrust EV** (2005) Synergistic effects of light, temperature, and nitrogen source
810 on transcription of genes for carbon and nitrogen metabolism in the centric diatom
811 *Thalassiosira pseudonana* (Bacillariophyceae). *J Phycol* **41**:1142–1153

812 **Parkes JA, Langer S, Hartig A, Baker A** (2003) PTS1-independent targeting of isocitrate lyase
813 to peroxisomes requires the PTS1 receptor Pex5p. *Mol Membrane Biol* **20**:61-69

814 **Paul JS, Volcani BE** (1975) Photorespiration in diatoms III. Glycolate:cytochrome *c* reductase
815 in the diatom *Cylindrotheca fusiformis*. *Plant Sci Lett* **5**:281–285

816 **Paul JS, Sullivan CW, Volcani BE** (1975) Photorespiration in diatoms III. Mitochondrial
817 glycolate dehydrogenase in *Cylindrotheca fusiformis* and *Nitzschia alba*. *Arch Biochem*
818 *Biophys* **169**:152–159

819 **Pick TR, Bräutigam A, Schulz MA, Obata T, Fernie AR, Weber APM** (2013) PLGG1, a
820 plastidic glycolate glycerate transporter, is required for photorespiration and defines a unique
821 class of metabolite transporters. *Proc Natl Acad Sci USA* **110**:3185–3190

822 **Piekarska K, Hardy G, Mol E, van den Burg J, Strijbis K, van Roermund C, van den Berg**
823 **M, Distel B** (2008) The activity of the glyoxylate cycle in peroxisomes of *Candida albicans*
824 depends on a functional β -oxidation pathway: evidence for reduced metabolite transport

across the peroxisomal membrane. *Microbiology* **154**:3061–3072

Plaxton WC, Podestá FE (2007) The functional organization and control of plant respiration. *Critical Rev Plant Sci* **25**:159-198

Poulsen N, Chesley PM, Kröger N (2006) Molecular genetic manipulation of the diatom *Thalassiosira pseudonana* (Bacillariophyceae) *J Phycol* **42**:1059–1065

Prihoda J, Tanaka A, de Paula WBM, Allen JF, Tirichine L, Bowler C (2012) Chloroplast-mitochondria cross-talk in diatoms. *J Exp Bot* **63**:1543-1557

R Core Team 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>

Reinfelder JR (2011) Carbon concentrating mechanisms in eukaryotic marine phytoplankton. *Annu Rev Mar Sci* **3**:291–315

Reinfelder JR, Milligan AJ, Morel FMM (2004) The role of the C4 pathway in carbon accumulation and fixation in a marine diatom. *Plant Phys* **135**:2106-2111

Reumann S, Weber APM (2006) Plant peroxisomes respire in the light: Some gaps of the photorespiratory C2 cycle have become filled - others remain. *Biochim Biophys Acta Mol Cell Res* **1763**:1496–1510

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP (2011) Integrative genomics viewer. *Nature Biotech* **29**:24–26

Rucktäschel R, Girzalsky W, Erdmann R (2011) Protein import machineries of peroxisomes. *Biochim Biophys Acta Biomembranes* **1808**:892–900

Sarthou G., Timmermans KR, Blain S, Tréguer P (2005) Growth physiology and fate of diatoms in the ocean: a review. *J Sea Res* **53**:25–42

Schnitzler Parker M, Armbrust EV, Piovia-Scott J, Keil RG (2004) Induction of photorespiration by light in the centric diatom *Thalassiosira weissflogii* (Bacillariophyceae): Molecular characterization and physiological consequences. *J Phycol* **40**:557–567

Schwender J, Ohlrogge J, Shachar-Hill Y (2004) Understanding flux in plant metabolic networks. *Curr Opin Plant Biol* **7**:309–317

Schwender J, Shachar-Hill Y, Ohlrogge JB (2006) Mitochondrial metabolism in developing embryos of *Brassica napus*. *J Biol Chem* **281**:34040–34047

Scrutton NS, Berry A, Perham RN (1990) Redesign of the coenzyme specificity of a dehydrogenase by protein engineering. *Nature* **343**:38–43

Shrestha RP, Haerizadeh F, Hildebrand M (2013) Molecular Genetic Manipulation of Microalgae: Principles and Applications. In Richmond AM, Hu X (eds) *Handbook of*

858 Microalgal Culture: Applied Phycology and Biotechnology, 2nd edn, Second Edition, John
859 Wiley & Sons, Ltd., Oxford UK, pp 146-167

860 **Smith SR, Glé C, Abbriano RM, Traller JC, Davis A, Trentacoste E, Vernet M, Allen AE,**
861 **Hildebrand M** (2016) Transcript level coordination of carbon pathways during silicon
862 starvation-induced lipid accumulation in the diatom *Thalassiosira pseudonana*. New
863 Phytologist **210**:890–904

864 **Smith SR, Abbriano RM, Hildebrand M** (2012) Comparative analysis of diatom genomes
865 reveals substantial differences in the organization of carbon partitioning pathways. Algal
866 Research **1**:2-16

867 **Sweetlove LJ, Beard KFM, Nunes-Nesi A, Fernie AR, Ratcliffe RG** (2010) Not just a circle:
868 flux modes in the plant TCA cycle. Trends Plant Sci **15**:462–470

869 **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S** (2013) MEGA6: molecular
870 evolutionary genetics analysis version 6.0. Mol Biol Evol **30**:2725–2729

871 **Tanaka R, Kikutani S, Mahardika A, Matsuda Y** (2014) Localization of enzymes relating to
872 C4 organic acid metabolisms in the marine diatom, *Thalassiosira pseudonana*. Photosyn Res
873 **121**:251-63

874 **Thorvaldsdóttir H, Robinson JT, Mesirov JP** (2012) Integrative Genomics Viewer (IGV):
875 high-performance genomics data visualization and exploration. Briefings Bioinformatics
876 **14**:178-92

877 **Turpin DH, Elrifí IR, Birch DG, Weger HG, Holmes JJ** (1988) Interactions between
878 photosynthesis, respiration, and nitrogen assimilation in microalgae. Can J Bot **66**: 2083 -
879 2097

880 **van der Klei IJ, Veenhuis M** (2006) PTS1-independent sorting of peroxisomal matrix proteins
881 by Pex5p. Biochim Biophys Acta Mol Cell Res **1763**:1794–1800

882 **Vanlerberghe G** (2013) Alternative oxidase: a mitochondrial respiratory pathway to maintain
883 metabolic and signaling homeostasis during abiotic and biotic stress in plants. Int J Mol Sci
884 **14**:6805–6847

885 **Warnes GR** (2015) gplots: various R programming tools for plotting data. R package version
886 2.17.0. <http://CRAN.R-project.org/package=gplots>

887 **Wingler A, Lea PJ, Quick WP, Leegood RC** (2000) Photorespiration: metabolic pathways and
888 their role in stress protection. Phil Trans R Soc B Biol Sci **355**:1517–1529

889 **Winkler U, Stabenau H** (1992) Compartmentation of Peroxisomal Enzymes in the Diatom
890 *Fragilaria*. In Stabenau H (ed) Phylogenetic Changes in Peroxisomes of Algae —
891 Phylogeny of Plant Peroxisomes. University Press, Oldenburg, pp 130-139

892 **Winkler U, Stabenau H** (1995) Isolation and characterization of peroxisomes from diatoms.

Planta **195**:403–407

Yang ZK, Yang WD, Liu JS, Lu SH, Guan Y, Li HY (2013) Molecular and cellular mechanisms of neutral lipid accumulation in diatom following nitrogen deprivation. *Biotechnol Biofuels* **6**:1–67

Yu ET, Zendejas FJ, Lane PD, Gaucher S, Simmons BA, Lane TW (2009) Triacylglycerol accumulation and profiling in the model diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum* (Bacillariophyceae) during starvation. *J Appl Phycol* **21**:669–681

Zheng Y, Quinn AH, Sriram G (2013) Experimental evidence and isotopomer analysis of mixotrophic glucose metabolism in the marine diatom *Phaeodactylum tricornutum*. *Microb Cell Factories* **12**:109

FIGURE LEGENDS

Figure 1. Proposed metabolic map based on in silico analysis and select *in vivo* validation. Map features glyoxylate cycle in blue and the C2 photorespiratory pathway in pink. Dashed lines indicate uncharacterized transport processes. Circled letters indicate metabolic processes featured in the discussion: **A**) Glycolate oxidation in the peroxisome and the mitochondria; **B**) Completion of the C2 pathway by a glycerate kinase; **C**) transamination of glyoxylate; **D**) Peroxisomal isocitrate lyase and malate synthase; **E**) Decarboxylation of malate in the mitochondria by malic enzyme; **?**) Grey box contains previously proposed reactions unsupported by our analysis. **Enzyme abbreviations:** 2HADH, 2-hydroxyacid dehydrogenase; ALAT/GGAT, alanine aminotransferase/glutamate glyoxylate aminotransferase; GDH, glycolate dehydrogenase; G2K, glycerate-2-kinase; GOX, glycolate oxidase, HPR, hydroxypyruvate reductase; ICL, isocitrate lyase; MDH, malate dehydrogenase; ME, malic enzyme; MS, malate synthase; PDC, pyruvate dehydrogenase complex; PGP, phosphoglycolate phosphatase; PLGG1, plastidic glycolate glycerate transporter; SHMT; SPT/AGAT, serine-pyruvate aminotransferase/alanine-glyoxylate aminotransferase.

Figure 2. Expression of genes associated with photorespiratory processes and the glyoxylate cycle in *T. pseudonana* during the silicon starvation timecourse. **A**) Expression of genes

putatively involved in the production of and oxidation of photorespiratory glycolate; **B**) Expression of genes proposed to participate in the metabolism of photorespiratory intermediates and amino acid synthesis in diatoms; **C**) Expression of genes proposed to participate in the glyoxylate cycle in diatoms; **D**) Genes with expression patterns that are distinct from those involved in photorespiratory processes and the glyoxylate cycle. Clustering was done manually by normalizing transcript levels across each time point, then calculating the standard deviation of the mean for each point. The mean variation of the standard deviation in terms of percent for each point ranged from 3.7-13%.

Table 1. In silico targeting predictions for proteins associated with C2 photorespiratory metabolism and the glyoxylate cycle. Proteins are organized based on predicted or demonstrated localization to A) mitochondria, B) peroxisome, and C) chloroplast. Protein models were extended when appropriate using RNAseq evidence to obtain the most accurate targeting prediction. Predictions were made with programs available through the ExPASy Bioinformatics Resource Portal (<http://www.expasy.org/tools/>), with the exception of ASAFind, which is hosted by the Rocap Lab (<http://rocaplab.ocean.washington.edu/tools/asafind>) and is specific for secondary endosymbionts of the red algal lineage.

Figure 3. Phylogenetic analysis of glycolate oxidizing proteins found in proteobacteria, photosynthetic eukaryotes and cyanobacteria. 31 sequences were aligned using the MAFFT algorithm, and the MEGA6 software was used to estimate the maximum likelihood tree using 1000 bootstrap replicates. The distance scale (substitutions per site) is shown in the bottom left-hand corner. Enzymes that are included in the SWISS-Prot database with biochemical evidence to support their annotation are indicated with an asterisk. Diatom glycolate oxidase-like proteins (GOXs) form two distinct clades. In *T. pseudonana*, GOX1 and GOX2 are localized in peroxisomes and mitochondria, respectively.

Figure 4. Colocalization of glycolate oxidase 2 (GOX2) with mitochondrial citrate synthase (CS). Representative *T. pseudonana* cells. False color images represent GFP in green, TagRFP in pink and chlorophyll autofluorescence in red. White demonstrates co-localization of the fluorescent proteins. Expression of GOX2 from the plasmid pTp FCP/Tp3353_gfp and CS from

from pTp rpL41/Tp11411-TagRFP demonstrates that GOX2 is co-localized with mitochondrial CS.

Figure 5. Heatmap representing calculated percent identity among 27 glycerate kinase (GLYK, 3-PGA forming, blue bar) and glycerate-2-kinase (GK, 2-PGA forming, gray bar) protein sequences. Biochemically characterized enzymes are noted with an asterisk. Diatom GK sequences are distinct from known GLYKs found in cyanobacteria, algae, and higher plants. Diatom GK sequences are most similar to archaeal GKs, such as the characterized sequence from *Thermoproteus tenax*, suggesting a unique origin for these enzymes. In addition, the lack of a GLYK sequence in diatom genomes supports the idea that the canonical C2 photorespiratory pathway is incomplete in diatoms. Hierarchical clustering of percent identity scores was done in R (<http://www.R-project.org/>) using the heatmap.2 function within the gplots package (<http://CRAN.R-project.org/package=gplots>).

Figure 6. Localization of malate synthase and isocitrate lyase in *T. pseudonana*. Representative cells of *T. pseudonana* in different orientations. False color images represent GFP in green and chlorophyll autofluorescence in red. **A)** Expression of MS from plasmid pTp_native/Tp26293-bfloa1_gfp and ICL from plasmid pTp_FCP/Tp35523-gfp demonstrate peroxisomal localization that typically appeared as two distinct, but closely associated spots per cell that are less than or equal to 1µm.

Figure 7. Colocalization of malic enzyme (ME) with mitochondrial citrate synthase (CS). Representative *T. pseudonana* cells. False color images represent GFP in green, TagRFP in pink and chlorophyll autofluorescence in red. White demonstrates co-localization of the fluorescent proteins. Expression of ME from the plasmid pTp Native/Tp5100-gfp and CS from pTp rpL41/Tp11411-tag_rfp demonstrate that ME is co-localized with mitochondrial CS.

A

Glycolate

B

Amino Acids

C

Glyoxylate Cycle

D

Other Responses

Si Starvation (h)

0.3

Malate Synthase

Isocitrate Lyase

	Chlorophyll + Citrate Synthase	Chlorophyll + Malic Enzyme	Citrate Synthase + Malic Enzyme	Chlorophyll + Citrate Synthase + Malic Enzyme	DIC
--	--------------------------------------	----------------------------------	---------------------------------------	---	-----

Cell 1

5μm

5μm

5μm

5μm

5μm

Cell 2

5μm

5μm

5μm

5μm

5μm

Cell 3

5μm

5μm

5μm

5μm

5μm

A) Mitochondria

Protein	Thaps3_PID	SignalP 3.0	ChloroP	ASAFind	Mitoprot	PTS1
GOX2	3353	-	N, 0.451	-	0.287	-
GDH	42965	-	Y, 0.565	-	0.312	-
GDC-H	28521	-	Y, 0.535	-	0.993	-
GDC-T	36208	-	Y, 0.507	-	0.999	-
GDC-P	39799	SNA-LL	N, 0.479	Yes, low	0.780	-
GDC-L	36716	-	Y, 0.555	-	0.486	-
SHMT	269942	-	Y, 0.520	-	0.955	-
ALAT/GGAT	412	-	Y, 0.517	-	0.949	-
AGAT/SPT	22208	-	Y, 0.583	-	0.997	-
2HADH	2846	-	N, 0.469	-	0.969	-
MDH	20726	-	Y, 0.575	-	0.9616	-
ME	34030	-	Y, 0.506	-	0.996	-

B) Peroxisome

Protein	Thaps3_PID	SignalP 3.0	ChloroP	ASAFind	Mitoprot	PTS1
GOX1	406	-	N, 0.453	-	0.430	SKL
CAT	11615	-	N, 0.447	-	0.026	SRL
MS	26293	-	N, 0.434	-	0.036	SKL
ICL	35523	-	N, 0.449	-	0.471	-
MDH	41425	-	N, 0.442	-	0.762	SRI

C) Chloroplast

Protein	Thaps3_PID	SignalP 3.0	ChloroP	ASAFind	Mitoprot	PTS1
PGP	25544	TTA-FS	Y, 0.588	Yes, high	0.914	-
PLGG1	9352	SVA-LV	Y, 0.563	Yes, low	0.828	-
	14106	IVA-TT	Y, 0.512	No	0.904	-
GK	261750	SSA-SS	Y, 0.544	No	0.968	-