



HAL
open science

Inorganic carbon uptake in a freshwater diatom, *Asterionella formosa* (Bacillariophyceae): from ecology to genomics

Stephen C Maberly, Brigitte Gontero, Carine Puppo, Adrien Villain, Ilenia Severi, Mario Giordano

► To cite this version:

Stephen C Maberly, Brigitte Gontero, Carine Puppo, Adrien Villain, Ilenia Severi, et al.. Inorganic carbon uptake in a freshwater diatom, *Asterionella formosa* (Bacillariophyceae): from ecology to genomics. *Phycologia*, 2021, pp.1-12. 10.1080/00318884.2021.1916297 . hal-03251911

HAL Id: hal-03251911

<https://amu.hal.science/hal-03251911>

Submitted on 7 Jun 2021

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

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



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Inorganic carbon uptake in a freshwater diatom, *Asterionella formosa* (Bacillariophyceae): from ecology to genomics

STEPHEN C. MABERLY¹, BRIGITTE GONTERO², CARINE PUPPO², ADRIEN VILLAIN^{3,4}, ILENIA SEVERI^{5,6} AND MARIO GIORDANO⁶

¹*Lake Ecosystems Group, UK Centre for Ecology & Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK*

²*Aix Marseille Univ CNRS, BIP UMR 7281, IMM, FR 3479, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20, France*

³*Aix Marseille Univ CNRS, IGS UMR 7256, IMM, FR 3479, 13009 Marseille, France*

⁴*BIOASTER, 40 Avenue Tony Garnier 69007 Lyon, France*

⁵*Section of Neuroscience and Cellular Biology, Department of Experimental and Clinical Medicine, Marche Polytechnic University, Via Tronto 10/A, 60020 Ancona, Italy*

⁶*Dipartimento di Scienze della Vita e dell'Ambiente, Università Politecnica delle Marche, 60131 Ancona, Italy*

ORCID

S.C. Maberly : ORCID ID 0000-0003-3541-5903

B. Gontero : ORCID ID 0000-0003-1731-712X

A. Villain : ORCID ID 0000-0001-5338-8959

I. Severi : ORCID ID 0000-0002-3985-1131

M. Giordano : ORCID ID 0000-0001-7754-0521

CONTACT

Corresponding author: scm@ceh.ac.uk

RUNNING TITLE

Inorganic carbon uptake in *Asterionella formosa*

ACKNOWLEDGEMENTS

We thank the Laboratoire Information Génomique et Structurale (IGS) for their help with the genomic data. IS was supported by a Campusword Project from the Università Politecnica delle Marche during the part of her MSc at Lancaster University. We dedicate this paper to our co-author and dear friend, the late Mario Giordano, who was a leader in CCM research.

FUNDING

The sequencing of the genome of *A. formosa* was partly supported by the (ANR) A*MIDEX project (n° ANR-11-IDEX-0001-02).

ABSTRACT

Inorganic carbon availability can limit primary productivity and control species composition of freshwater phytoplankton. This is despite the presence of CO₂-concentrating mechanisms (CCMs) in some species that maximise inorganic carbon uptake. Here, we investigated the effects of inorganic carbon on the seasonal distribution, growth rates and photosynthesis of a freshwater diatom, *Asterionella formosa*, and the nature of its CCM using genomics. In a productive lake, the frequency of *A. formosa* declined with CO₂ concentration below air-equilibrium. In contrast, CO₂ concentrations 2.5-times air-equilibrium did not increase growth rate, cell C-quota or the ability to remove inorganic carbon. A pH-drift experiment strongly suggested that HCO₃⁻ as well as CO₂ could be used. Calculations combining hourly inorganic carbon concentrations in a lake with known CO₂ and HCO₃⁻ uptake kinetics suggested that rates of photosynthesis of *A. formosa* would be approximately carbon saturated and largely dependent on CO₂ uptake when CO₂ was at or above air-equilibrium. However, during summer carbon depletion, HCO₃⁻ would be the major form of carbon taken up and carbon saturation will fall to around 30%. Genes encoding proteins involved in CCMs were identified in the nuclear genome of *A. formosa*. We found carbonic anhydrases from subclasses α , β , γ and θ , as well as solute carriers from families 4 and 26 involved in HCO₃⁻ transport, but no periplasmic carbonic anhydrase. A model of the components of the CCM and their location in *A. formosa* showed that they are more similar to *Phaeodactylum tricornutum* than to *Thalassiosira pseudonana*, two marine diatoms.

KEYWORDS

Aquatic photosynthesis; Bicarbonate use; Carbonic anhydrase; CO₂-concentrating mechanism; Solute carrier (SLC)

INTRODUCTION

Photosynthesis in water can rely on CO_2 and HCO_3^- as exogenous sources of inorganic carbon. In lakes, concentrations of CO_2 can be highly variable, temporally and spatially (Maberly & Gontero 2017). Concentrations of HCO_3^- normally exceed the concentration of CO_2 but vary widely among sites as a consequence of catchment geology (Iversen *et al.* 2019). The primary carboxylation enzyme in the Calvin–Benson–Bassham cycle, ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), fixes CO_2 with a relatively low affinity and performs an oxygenation reaction when concentrations of CO_2 at its active site are low. This can lead to photorespiration (Bowes & Ogren 1972) and reduced productivity. Marine diatom RubisCO proteins have a Michaelis–Menten constant between 23 and 68 $\mu\text{mol l}^{-1}$ (Young *et al.* 2016). Assuming that these values also apply to freshwater diatoms, the air-equilibrium concentration of CO_2 in fresh water at an atmospheric CO_2 partial pressures of 400 ppm varies between 25 $\mu\text{mol l}^{-1}$ at 5°C and 14 $\mu\text{mol l}^{-1}$ at 25°C. Consequently, carboxylation by diatom RubisCO could be limited by CO_2 availability at light saturation under these conditions. Furthermore, in productive lakes, the concentration of CO_2 can fall several orders of magnitude below air-equilibrium (Talling 1976; Maberly 1996) potentially limiting photosynthesis even further.

A widespread mechanism to mitigate carbon limitation involves CO_2 -concentrating mechanisms (CCMs) that increase CO_2 around the active site of RubisCO (Giordano *et al.* 2005). These involve a range of processes and structures including active uptake of CO_2 and HCO_3^- , the presence of the metalloenzyme carbonic anhydrase (CA) within different cell compartments to increase the rate of CO_2 and HCO_3^- interconversion, and the concentration of RubisCO in specific areas, such as the pyrenoid, where CO_2 concentration can be elevated (Mackinder *et al.* 2016; Meyer *et al.* 2017; Launay *et al.* 2020). By 2100, atmospheric CO_2 partial pressures are projected to increase to over 900 ppm, depending on the climate change

scenario (Meinshausen *et al.* 2011). Diatom CCMs, like those of many other aquatic photoautotrophs, are often down-regulated at high CO₂ and, therefore, in the future may be down-regulated, assuming air-equilibrium in the water.

Asterionella formosa Hassall, is one of the most abundant, widespread and well-studied freshwater diatoms. It evolved at the end of the Cretaceous period between 85 and 70 Ma (Medlin & Desdevises 2016; Medlin & Desdevises 2020) when atmospheric CO₂ partial pressures were declining but higher than today (Wang *et al.* 2014). In meso-eutrophic lakes in the UK, such as Windermere or Esthwaite Water, *A. formosa* is present in the lake for much of the year, is a major component of the spring bloom (Lund 1949; Maberly *et al.* 1994) and experiences a wide range of CO₂ concentration during its seasonal growth cycle.

Experiments by Talling (1976) showed that cells from the South Basin of Windermere at a background HCO₃⁻ concentration of about 200 μmol l⁻¹ had maximal rates of photosynthesis at free CO₂ concentrations greater than about 10 μmol l⁻¹ and net photosynthesis continued down to less than 0.1 μmol l⁻¹ which is very roughly 1% of air-equilibrium. In Esthwaite Water, Talling (1976) also showed that the reduction in CO₂ concentration during the spring and summer is linked to a seasonal succession of phytoplankton with increasing ability to remove CO₂ from the water: from the diatoms *Aulacoseira italica* (*Melosira italica*) (Ehrenberg) Simonsen to *A. formosa* to *Fragilaria crotonensis* Kitton, to the cyanobacterium *Microcystis aeruginosa* (Kützing) Kützing.

Here we analyse the ecophysiology of inorganic carbon uptake by *A. formosa* from different perspectives: seasonal growth patterns, growth responses to CO₂, photosynthesis as a function of CO₂ and HCO₃⁻ concentration, and the presence of genes encoding CAs and bicarbonate transporters (solute carriers, SLCs) that are involved in CCMs, e.g. Huang *et al.* (2020).

MATERIAL AND METHODS

CO₂ concentrations in Esthwaite Water and the seasonal distribution of *A. formosa*

Esthwaite Water is a meso-eutrophic lake in the English Lake District that has been intensively studied since 1945. Water samples between 1983 and 2014 from the top 5 m were used to estimate the population density of *A. formosa* and, as a comparison, species of *Anabaena* (Cyanophyceae; sometimes now placed in *Dolichospermum*). Concentrations of CO₂ over the same period were calculated following Maberly (1996) using temperature, temperature-corrected pH measured in the laboratory with a combination pH electrode (Radiometer Copenhagen GK 2401C) and alkalinity measured by Gran titration (Mackereth *et al.* 1989) and an ionic strength of 1.35 mmol l⁻¹.

Response of *A. formosa* growth rate to CO₂ and pH

A culture of *A. formosa* CCAP 1005/18, originally isolated from Esthwaite Water (54°22'N, 2°59'W) in 2007, was obtained from the Culture Collection of Algae and Protozoa (CCAP, Oban, Scotland). Cells were grown in filter-sterilised (0.22 µm) modified Diatom Medium (DM; Beakes *et al.* 1988), with the addition of tris-hydroxy methyl amino methane, Tris-HCl (10 mmol l⁻¹) that allowed pH to be adjusted to 7.0, 7.5 and 8.0 with 1 mol l⁻¹ HCl or NaOH. Cells at an initial concentration of 1 x 10⁴ cell ml⁻¹ in 800 mL of culture medium were grown in 1000-mL Schott glass bottles at 20°C, a photon irradiance of 100 µmol photon m⁻² s⁻¹ (photosynthetically available radiation; previously shown to be saturating), and a 16:8 h light:dark cycle. Cultures were bubbled continuously at 150–180 mL min⁻¹ with gas passed through a 0.22-µm filter, either with 390 ppm CO₂ (laboratory air) or with 1000 ppm CO₂ in air, enriched using CO₂ controlled by a mass flow gas mixer (EL FLOW- select, Bronkhorst,

Cambridge, UK). CO₂ partial pressures were checked daily with an infrared gas analyser (ADC 225 Mk3, Analytical Development Company, Amersham, UK). Cells were collected daily and counted in a chamber (Lund 1959) to determine exponential growth rate. Cells were collected for analysis from the exponential phase when cell density was between 2×10^5 and 4×10^5 cells ml⁻¹. Cellular carbon was combusted in an O₂-enriched atmosphere, transported in a He carrier stream to a chromatographic column, and CO₂ was measured using a thermic conductivity detector (EA-1108 CHNS-O, Carlo Erba, Milano, Italy). The software used for data acquisition was EAS-Clarity (DataApex Ltd. Prague, Czech Republic). Triplicate cultures were analysed for each treatment.

pH-drift experiments

Using material grown as described above, drifts were carried out in a medium comprising 1 mM NaHCO₃, 0.1 mM MgSO₄, 0.09 mM KCl and 0.08 mM CaCl₂. A volume of 100 mL, containing from 2 to 4×10^7 cells, was centrifuged at $12000 \times g$, for 15 min at room temperature (Eppendorf Centrifuge 5804 R, Eppendorf AG, Hamburg, Germany). Cells were resuspended in 10 mL of the pH-drift medium, re-centrifuged and then suspended in 75 mL of pH-drift medium in 100-mL bottles. Stoppered bottles were maintained under the same temperature and light conditions as during growth. After careful mixing, pH was measured with a pH-meter (Radiometer PHM 84 Research pH meter, Crawley, UK). Before replacing the stopper, the gas space was flushed for a few seconds with N₂ to remove CO₂ from the air and to reduce oxygen concentration. pH was measured daily until a stable pH was reached. Total inorganic carbon concentration at the end of a drift was measured by injecting 10–25 µL of solution into 2 mL of phosphoric acid (8% v/v) that was carried by a stream of nitrogen to the infrared gas analyzer (ADC 225 Mk3, Analytical Development Company, Amersham, UK) calibrated against 1 mmol l⁻¹ NaHCO₃. Measurements were taken in triplicate.

Calculations of daily CO₂- and HCO₃⁻-dependent rates of *A. formosa* photosynthesis in Esthwaite Water

High frequency (every 15 min) measurements of pH and temperature have been collected in the sub-surface of Esthwaite Water since 1993. Hourly estimates of pH, CO₂ and HCO₃⁻ for 1993 (Maberly 1996) are used here in conjunction with photosynthesis kinetics measured for a strain (CCAP 1005/24) of *A. formosa* from Esthwaite Water (Clement *et al.* 2017a; unpublished) grown at air-equilibrium CO₂ concentration, measured at 16°C. Data were fitted with a model assuming separate kinetics of CO₂ and HCO₃⁻, that contributed to a common maximum rate (Clement *et al.* 2016). Specifically, the maximum rates (μmol O₂ mg⁻¹ chlorophyll *a* h⁻¹) were 71 and 34 for CO₂ and HCO₃⁻, respectively; K_{0.5} values (μmol l⁻¹) were 2.57 and 37.63 for CO₂ and HCO₃⁻, respectively; and the compensation points (μmol l⁻¹) were 0.0001 and 1.3 for CO₂ and HCO₃⁻, respectively. These kinetic data were combined with daily minimum and maximum concentrations of CO₂ and HCO₃⁻ to illustrate how they can affect rates of photosynthesis, assuming light saturation and ignoring temperature effects.

Analysis of the *A. formosa* genome and comparison with CCM components in other diatoms

A culture of *A. formosa* from Esthwaite Water (CCAP 1005/24) was grown at 16.5°C with a 12:12 h light:dark cycle illuminated with a photon irradiance of 50 μmol photon m⁻² s⁻¹ and shaken at 110 rpm. Genomic DNA and RNA were extracted from cultures in exponential phase containing about 2.5·10⁷ and 10⁸ cells ml⁻¹ of *A. formosa*, respectively. Cultures were filtered on 8-μm and then 0.22-μm membranes, frozen in liquid nitrogen and stored at -80°C. DNA was extracted following a hexadecyltrimethylammonium bromide (CTAB)-based protocol adapted from (Bruckner *et al.* 2008). RNA was extracted using TRI-reagent (Sigma-

Aldrich, St. Louis, Missouri, USA) and the SV total RNA isolation system (Promega, Madison, Wisconsin, USA). DNA sequencing was performed using Illumina NextSeq 500 technology (profiXpert platform, Lyon, France). Sequencing was completed by RNAseq to annotate the eukaryotic genes that contain exons and introns, and the data were deposited in NCBI (accession number SRX2949863). Fourteen single-molecule real-time cells (a nanophotonic device) on a Pacific Biosciences RSII sequencer at the GeT-PlaGe platform (INRA, Toulouse, France) were used to improve sequencing quality of the genome. Genome assembly was performed using the hierarchical genome-assembly process software HGAP (Chin *et al.* 2013), and annotated using MAKER (Campbell *et al.* 2014). All sequences included here have been submitted to GenBank (using BankIt ID 2401528) and are listed in Table S1.

The genome of *A. formosa* was compared to existing databases to identify genes encoding components of the CCM: all subclasses of carbonic anhydrase (CA) and SLC families 4 and 26 (SLC4 and SLC26). Using a local database that included 17,133 proteins from *A. formosa* (unpublished data, A. Villain and G. Blanc, personal communication) we performed a BLAST search using BioEdit software and proteins of *Phaeodactylum tricornutum* Bohlin as queries. These queries were CAs, the θ -CA described in Kikutani (2016); the plasma membrane and chloroplast solute carrier transporters SLC4, and SLC26, listed in Matsuda *et al.* (2017) and the thioredoxins f and m (Kikutani *et al.* 2012). The identity of every protein was checked and aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE; <https://www.ebi.ac.uk/Tools/msa/muscle/>). Alignments were imported into Genedoc (<http://www.nrbsc.org/gfx/genedoc>) and similar residues were shaded using the conservation mode.

The topology of the putative SLCs was predicted using TMHMM v2 (<http://www.cbs.dtu.dk/services/TMHMM/>) to confirm that they were transmembrane

proteins. The 2D structure was predicted using Hydrophobic Cluster Analysis (HCA; <https://mobylye.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::HCA>; Callebaut *et al.* 1997). The location of the CA and SLC proteins was determined, based on their signal peptides, using HECTAR (<https://webtools.sb-roscoff.fr/>) that predicted four categories: the chloroplast, the mitochondrion, 'other' and uncategorised. Only four proteins were localised using HECTAR, the other proteins were localised mainly by homology with *P. tricornutum*, except for a few sequences where *Arabidopsis thaliana* (L.) Heynh. was used.

RESULTS

The seasonal distribution of *A. formosa* in relation to CO₂ concentration

Over 32 years, CO₂ concentration in Esthwaite Water varied by 61,000-fold, between $2.7 \cdot 10^{-3}$ and $166 \mu\text{mol l}^{-1}$, with the very low CO₂ minima occurring in summer (Fig. 1). While *A. formosa* was found every week of the year, in some years cell densities in spring were about 200 times higher than in summer (Fig. 2). In contrast, the cyanobacterium *Anabaena* was more abundant in summer and autumn than in spring. *Anabaena* species were present at all concentrations of CO₂ and were most prevalent when CO₂ was low (Fig. 3), while *A. formosa* was absent, or present less than 15% of the dates when CO₂ concentration was less than $0.3 \mu\text{mol l}^{-1}$. This suggests that low CO₂ concentrations are one of the important ecological factors restricting seasonal distribution of *A. formosa*.

Response of *A. formosa* growth rate and carbon cell quota to CO₂ and pH

Although productive lakes such as Esthwaite Water are rarely at equilibrium with the atmosphere, concentrations of CO₂ in water will tend to increase as atmospheric CO₂ partial pressure increases. CO₂ concentration had no significant effect on growth rate (two-way ANOVA, $P < 0.05$; Fig. 4) and growth rates at 390 ppm ($16 \mu\text{mol l}^{-1}$) were not significantly different from those at 1000 ppm CO₂ ($39 \mu\text{mol l}^{-1}$) at any of the pH values (Student's *t*-tests, $P = 0.14$ to 0.22).

However, pH had a significant effect on growth rate: rates were lower at pH 8.0 than at pH 7.0 and 7.5 at both CO₂ concentrations. The cell quota for carbon was, on average, $44.23 \text{ pg C cell}^{-1}$ ($s = 6.05$; Fig. 5) and was not significantly affected by pH or CO₂ (two-way ANOVA, $P > 0.05$). The cell carbon quotas in cells grown at 390 ppm and 1000 ppm were not significantly different (Student's *t*-tests, $P = 0.22$ to 0.92). These results suggest that modest increases in CO₂ concentration above present day air-equilibrium, that may occur in

response to rising atmospheric CO₂, will not directly affect the growth rates or carbon-cell quota in *A. formosa*.

Inorganic carbon uptake during pH-drift experiments

Cells of *A. formosa* grown at 390 and 1000 ppm pCO₂ produced significantly lower final pH values in the drift if previously acclimated at pH 8 than at pH 7 or pH 7.5 ($P < 0.05$; Table 1). A two-way ANOVA showed that although growth pH had a significant effect on the final pH, the pCO₂ did not (two-way ANOVA, $P < 0.05$). Final conditions in cells grown at the two lower pH-values produced very similar final pH values of between 10.24 and 10.33, a final CO₂ concentration of between 0.05 and 0.08 $\mu\text{mol}\cdot\text{l}^{-1}$ and a final HCO₃⁻ concentration of between 0.43 and 0.45 mmol l^{-1} that was probably limited by the direct effect of high pH on the inorganic carbon uptake of *A. formosa*.

Calculation of daily CO₂- and HCO₃⁻-dependent rates of *A. formosa* photosynthesis

Concentrations of CO₂ and HCO₃⁻ change on a diel, episodic, and seasonal basis in Esthwaite Water and are likely to have major effects on rates of photosynthesis by *A. formosa*. High pH and low CO₂ concentrations occurred during the summer when rates of inorganic carbon uptake exceeded rates of resupply (Figs 6, 7). For much of the year, photosynthesis relied on CO₂, but HCO₃⁻ was the dominant carbon source during episodic summer carbon depletion (Fig. 8). Percent carbon saturation of photosynthesis was also highly variable seasonally (Fig. 9). At the relatively high CO₂ concentrations present in early spring and late autumn, rates of inorganic carbon uptake were over 95% carbon-saturated, and at air-equilibrium concentrations of CO₂, rates were about 90% saturated and CO₂-dependent inorganic carbon uptake was about twice that of HCO₃⁻ (Figs 10, 11). However, when CO₂ concentration fell

below about $2 \mu\text{mol l}^{-1}$, HCO_3^- became the dominant source of inorganic carbon. At very low CO_2 concentrations, saturation fell to less than 30%, and this was almost completely dependent on the use of HCO_3^- (Figs 10, 11). In addition, the large diel changes in pH and concentrations of CO_2 and HCO_3^- caused large diel changes in CO_2 -dependent photosynthesis, % carbon saturation, and the relative contribution of CO_2 and HCO_3^- to photosynthesis over the day.

CCM genes in *A. formosa*

Seventeen genes encoding carbonic anhydrases are present within the nuclear genome of *A. formosa*. These CAs are homologous to those of other pennate diatoms, mainly *P. tricornutum*, but also occasionally *Fistulifera solaris* S. Mayama, M. Matsumoto, K. Nemoto & T. Tanaka and *Fragilariopsis cylindrus* (Grunow ex Cleve) Helmcke & Krieger (Table 2). Eight α -CA isoforms were identified. They were localised in an ‘other’ compartment using the HECTAR predictor, but their location was refined by homology with α -CAs isoforms from *P. tricornutum* (Hopkinson *et al.* 2016). This suggested that two subtypes III (four proteins) and VI (three proteins) are in the chloroplast endoplasmic reticulum (CER) and that a single CA I protein is in the periplastidial compartment (PPC).

Two β isoforms were identified and, using HECTAR, both were predicted to be located in the chloroplast. These proteins are homologous to two β -CAs from *P. tricornutum*, PtCA1 (45433) and PtCA2 (51305):

<https://mycocosm.jgi.doe.gov/cgi-bin/dispGeneModel?db=Phatr2&id=51305>, or AAL07493.1

(NCBI) that were isolated, characterised and located in the pyrenoid (Satoh *et al.* 2001; Tanaka *et al.* 2005; Hopkinson *et al.* 2011; Tachibana *et al.* 2011; Kikutani *et al.* 2012). The five residues (M263, L266, I269, L272, L275) involved in the amphiphatic α -helix at the C terminus of the β -CA in *P. tricornutum* that are important to locate these enzymes in the

pyrenoid (Kitao & Matsuda 2009) are replaced at the same position by (I, L, I, L, P) in the β -CAs from *A. formosa* (Fig. 12). These residues are also hydrophobic and were predicted to form an α -helix like in *P. tricornutum*. Thus, the two β -CAs from *A. formosa* might also be pyrenoidal. One of the isoforms, AF09251, possesses cysteine residues in the GCV and CGG motifs and the histidine in the CGH motif that are involved in the binding of Zn^{2+} (Fig. 12). The other isoform (AF16830) has a deletion in the middle of the sequence that contains the Zn^{2+} binding histidine, and therefore this isoform is probably inactive.

The β -CA isoform AF09251 also has two cysteine residues (in the CGH and EQC motifs; Fig. 12) that are targets for thioredoxins (Kikutani *et al.* 2012), suggesting that it might be redox-regulated. However, the isoform AF16830 lacks one of the regulatory cysteine residues, and is therefore unlikely to be redox-regulated (Fig. 12). Thioredoxin f (AF06341) and thioredoxin m (AF02599) were identified by a BLAST search in the genome of *A. formosa* using thioredoxins f (1 thioredoxin f; NCBI accessionnumber EEC47925) and m (XP_002177112.1) from *P. tricornutum* as queries (Fig. S1). Thioredoxin f from *A. formosa* was predicted to be in the chloroplast as is the β -CA (AF09251), potentially allowing redox regulation of this CA. In contrast, thioredoxin m was located in an ‘other’ compartment, and its length of 444 amino acids is rather unusual for a typical thioredoxin. Alignments between these thioredoxins and those from *P. tricornutum* are shown in Fig. S1. They both contain the WCGPC canonical motif responsible for the redox property of thioredoxins (Weber *et al.* 2009) but the cysteine residue at position 60 that is glutathionylated in thioredoxin from angiosperms (Michelet *et al.* 2005) is absent in diatom thioredoxin sequences.

Three γ -CAs were identified (Table 2); the HECTAR predictor did not confirm their location, but three were predicted to be mitochondrial enzymes, indicated by homology to the angiosperm *A. thaliana*. Four putative θ -CAs were identified using a BLAST search with a

sequence query from *P. tricornutum* (Pt43233, NCBI BAV001424.1) that corresponds to a protein that has been identified as a θ -CA (Kikutani *et al.* 2016). These four CAs have the θ -CA signature, containing cysteine, the GPH and CCG residues except for isoform AF09289, where CCG was replaced by CCE (Fig. 13). Using the HECTAR predictor, they were all predicted to be in an ‘other’ compartment; however, by homology to those present in *P. tricornutum*, one can assume a pyrenoid localisation. Based on the genome sequencing data, the full-length proteins were shorter than in *P. tricornutum*. In addition, the peptide at the N-terminus, responsible for location within the thylakoid lumen (Fig. S2, in Kikutani *et al.* 2016), was absent in our sequences.

Four SLCs, transmembrane proteins that are involved in the active or passive transport of HCO_3^- across membranes, were found in the nuclear genome of *A. formosa*. Three of these were putative SLCs from family 4 (Romero *et al.* 2013); AF08596, AF11830 and AF09764). The sequence AF08596 is similar (74.6%) to SLC4-7 from *P. tricornutum* (Prot ID 45656, referring to JGI genome database; <http://genome.jgi.doe.gov/Phatr2/Phatr2.home.html>, or AB733624, NCBI) that was predicted to be chloroplastic. HECTAR software also predicted this SLC4 from *A. formosa* to be in the chloroplast. Another gene encodes the AF09764 protein that is hypothesised to be in the chloroplast as it has a 68% identity to chloroplastic SLC4-6 from *P. tricornutum* (Prot ID 43194, referring to JGI genome database (<http://genome.jgi.doe.gov/Phatr2/Phatr2.home.html>; AB733623, NCBI). However, HECTAR software predicted a signal peptide and a cleavage site at 33 residues for this sequence, but was unable to predict a location. Except for a deletion in the middle of the sequence, AF11830 is identical to AF08596. Using the TMHMM server v2 (SI Fig. 3), all of these SLC4s, as expected, were predicted to possess transmembrane helices.

In *P. tricornutum*, six SLC4 proteins were located in the plasma membrane but no plasma membrane SLC4s could be identified in *A. formosa*. However, in *P. tricornutum*, three SLC26s were predicted potentially to be in the plasma membrane. In *A. formosa* we found a gene encoding a protein (AF11040) that had 50.28% identity to plasma membrane SLC26-2 (Table 2; Pt42556, NCBI XP_002177084.1), and was predicted to possess transmembrane helices (Fig. S2). However, the location of this SLC26-2 remains an open question as it is predicted only *in silico* and should be confirmed experimentally. These results are used to suggest a model of the CCM in *A. formosa* compared to that of *Thalassiosira pseudonana* Hasle & Heimdal and *P. tricornutum* (Figs 14–16).

DISCUSSION

Inorganic carbon is increasingly recognised as an ecological factor that can affect the productivity and species composition of phytoplankton (Low-Decarie *et al.* 2014). Talling (1976) linked the seasonal succession of phytoplankton to decreasing concentrations of CO₂ and increasing ability to access this declining resource. Shapiro (1997) proposed that the dominance of cyanobacteria in summer is enhanced by their greater ability, compared to other species, to access low concentrations of inorganic carbon as a consequence of their highly effective CCMs (see review by Price, 2011). Laboratory and mesocosm experiments have demonstrated that increasing CO₂ concentration alters the competitive ability among different groups of phytoplankton (Low-Decarie *et al.* 2011; Low-Decarie *et al.* 2015) and also increases productivity (Kragh & Sand-Jensen 2018; Hammer *et al.* 2019). Here, we show that *A. formosa*, in contrast to the cyanobacterial genus *Anabaena* with an efficient CCM (Kaplan *et al.* 1980), is absent from a productive lake when CO₂ concentrations fall markedly below air-equilibrium. Multiple ecological factors may affect the seasonality of phytoplankton, including *A. formosa* (Maberly *et al.* 1994). For *A. formosa*, these include silica depletion in the spring that triggers the decline of the spring bloom as this essential resource runs out (Lund 1949), strong stratification that results in large sinking losses of these dense cells and exposure to higher irradiance as the mixed depth becomes shallower (Neale *et al.* 1991; Maberly *et al.* 1994). Historical and seasonal patterns of change in Esthwaite Water of numerous variables are described in Maberly *et al.* (2011). Depletion of inorganic carbon is an additional factor, because *A. formosa* cannot photosynthesise at CO₂ concentrations much below about 0.1 µmol l⁻¹ (Talling 1976; Clement *et al.* 2017b; Table 1), and so populations will be restricted when CO₂ concentrations are low. However, compensation concentrations for growth will be higher than this because of losses such as sinking, flushing,

grazing and parasitism, and this is consistent with its low prevalence at CO₂ concentration slightly higher than the compensation concentration, i.e. at around 0.3 μmol l⁻¹.

The concentration of CO₂ and HCO₃⁻ is highly variable in productive lakes (Maberly 1996), with major implications for the productivity of a species such as *A. formosa*. Growth experiments reported here showed that its growth rate and cellular carbon content were saturated at air-equilibrium concentrations of CO₂ (approximately 15 μmol l⁻¹ in these experiments) and this is consistent with the kinetic parameters of CO₂ and HCO₃⁻ uptake. The illustration of how inorganic carbon depletion by the phytoplankton community can severely limit the productivity of *A. formosa* was calculated for light saturation. When light is at least partly limiting, rates of carbon uptake will be reduced, but when inorganic carbon is strongly limiting, light will have relatively little effect (Talling 1979). The illustration might also underestimate carbon limitation because our pH-drift data suggest that high pH has a direct adverse effect on carbon uptake since final concentrations of HCO₃⁻ are substantially greater than the HCO₃⁻ compensation concentration estimated at pH 7 and 8 (Clement *et al.* 2017a; see Raven *et al.* 2020 for discussion of possible mechanisms).

Physiological data from Clement *et al.* (2017a), and the pH-drift data reported here, suggest that *A. formosa* is able to use HCO₃⁻. It has a low to moderate biophysical CCM that is less effective than some marine diatoms that are able to drive CO₂ concentrations down from 1 to 10 nmol l⁻¹ (Clement *et al.* 2017a), compared to 50 to 80 nmol l⁻¹ for *A. formosa*. Most diatoms that have been studied have a biophysical CCM, while *Thalassiosira weissflogii* (Grunow) G.A. Fryxell & Hasle (currently *Conticribra weissflogii* (Grunow) Stachura-Suchoples & D.M. Williams) is the only known diatom for which a biochemical CCM based on C₄ metabolism is not controversial (Clement *et al.* 2017a).

Multiple proteins are involved in eukaryotic CCMs, but carbonic anhydrases and solute carrier proteins appear to be particularly widespread (Giordano *et al.* 2005). CAs are metalloenzymes that catalyse the interconversion of CO₂ and HCO₃⁻ and are found across all domains of life. Nine CA subclasses have been described to date (α , β , γ , δ , ϵ , ζ , η , θ and ι -CAs), and seven of these are present in diatom genomes. CAs are critical components of CCMs, involved in inorganic carbon uptake. In many microalgae and cyanobacteria they are particularly important in minimising CO₂ leakage from the chloroplast (Matsuda *et al.* 2017). Although CA activity has been measured in *A. formosa* (Clement *et al.* 2017a), the subclasses present, and their putative location, were previously unknown. The nuclear genome of *A. formosa* (full dataset unpublished) contains CAs belonging to the α , β , γ and the θ subclass (Table 2). The reason for the high diversity of CA forms and locations in *A. formosa* and other diatoms, is mysterious. It is possible that different forms of CA are optimal in a specific location. In addition, especially in diatoms with their complex evolutionary history, different CA forms may have arisen from past evolutionary events. The presence of θ -CA in the thylakoid lumen is vital for the diatom CCM (Kikutani *et al.* 2016), but the θ -CA sequences derived from *A. formosa* did not allow us to conclude definitively where the enzymes were located. No sequences were identified for two cambialistic CAs identified in marine diatoms, ζ -CA and δ -CA, that can replace Zn²⁺ with other metal ions (Morel *et al.* 2020). Specifically, ζ -CA is widespread in marine diatoms and other phytoplankton but was not found in *P. tricornutum* (strain CCMP630; Park *et al.* 2007), and appeared to be absent in *A. formosa*. Similarly, δ -CA that is present in *T. weissflogii* (Roberts *et al.* 1997; Lane *et al.* 2005; Del Prete *et al.* 2014; Alterio *et al.* 2015; Angeli *et al.* 2018) was not found in *A. formosa*. Furthermore, the gene encoding ι -CA, a CA that can use Mn²⁺ and was recently discovered in *T. pseudonana* (Jensen *et al.* 2019), could not be detected in *A. formosa*. The ϵ - and the η -

CAs that have a restricted taxonomic distribution and have not previously been reported from diatoms were, unsurprisingly, also not found in *A. formosa*.

Redox-regulation is an important mechanism that fine-tunes photosynthesis under changing environmental conditions, but relatively little is known about this process in diatoms (Michels *et al.* 2005; Maberly *et al.* 2010; Mekhalfi *et al.* 2012). Thioredoxins play a central role in redox-regulation in many organisms, and their targets are well-known in angiosperms (Balmer *et al.* 2003; Balmer *et al.* 2004; Marchand *et al.* 2004), but not in diatoms (Wilhelm *et al.* 2006; Jensen *et al.* 2017; Launay *et al.* 2020). Chloroplastic β -CA from *A. formosa* is a possible target of thioredoxin because it possesses the regulatory cysteine residues, as shown for the β -CA in *P. tricornutum*. The presence of putative thioredoxins m and f, and particularly the predicted location of thioredoxin f in the chloroplast, strengthens the possibility of redox-regulation of β -CA in *A. formosa* by analogy to *P. tricornutum* (Kikutani *et al.* 2012). The presence of all genes encoding thioredoxins in the *A. formosa* genome was not investigated because it was out of the scope of this manuscript.

Multigene families of HCO_3^- transporters such as SLC4 and SLC26 have evolved independently, have been reported in mammals and also occur in the pennate diatom *P. tricornutum* and the centric diatom *T. pseudonana* (Nakajima *et al.* 2013; Tsuji *et al.* 2017). The conservation of these genes in evolutionary distinct species suggests that they are probably widespread and play a critical role. Two genes that potentially encode chloroplast SLC4s, and one gene that could encode a plasma membrane SLC26, were indeed identified in the genome of the pennate diatom *A. formosa*. Since it was shown that *A. formosa* can use bicarbonate (Clement *et al.* 2017a; Table 1), these transporters are likely to be involved in HCO_3^- uptake.

Models are presented in Figs 14–16 that summarise the new information about the CCM of *A. formosa* and compare it to two model marine diatoms. The number and subclasses of CA in the pennate *A. formosa* are more similar to the pennate marine diatom *P. tricornutum* than to the centric marine diatom *T. pseudonana* (Tsuji *et al.* 2017; Jensen *et al.* 2020). *T. pseudonana* lacks β -CA, which is present in the chloroplast of the two other species, but in contrast possesses three δ -CAs and a ζ -CA, both forms of which appear to be absent in the two other species. In *T. pseudonana*, only the δ -CA was found in the chloroplast (Samukawa *et al.* 2014) while, like in *P. tricornutum*, many CAs were predicted within the chloroplast membranes of *A. formosa*. *T. pseudonana* has six subclasses of CA, while *P. tricornutum* and *A. formosa* have five and four, respectively. All three species possess an SLC in the plasmamembrane, SLC4 in the two marine species and SLC26 in *A. formosa*. It is perhaps surprising that no periplasmic CA was identified in *A. formosa*, like in *P. tricornutum* (Tachibana *et al.* 2011) but in contrast to *T. pseudonana*. However, this is not a proof because it is based on *in silico* analysis (absence of sequence homology in existing genome database) rather than on experimental data. Physiological studies are required to establish if this is a true absence. Further work on the properties and location of the proteins involved in diatom CCMs will bring new insights into the mechanisms of inorganic carbon uptake into these cells and their consequences for diatom ecology and biogeochemical cycling in aquatic ecosystems. Furthermore, the sequences and presence of CAs in different compartments in *A. formosa* are more similar to *P. tricornutum* than to *T. pseudonana*. This is possibly related to their evolutionary history since *A. formosa* and *P. tricornutum* are pennate diatoms within the Bacillariophyceae while *T. pseudonana* is a centric diatom within the Mediophyceae (Medlin & Desdevises 2016). There is an indication that Bacillariophyceae may have higher maximal RubisCO catalytic activity k_{cat} (table 1 in Young *et al.* 2016) and higher Rubisco content as a proportion of total cellular protein than Mediophyceae (Losh *et*

al. 2013). However, comparative data are limited, and further study of CCMS from a phylogenetic perspective in this diverse group of microalgae is warranted.

REFERENCES

- Alterio V., Langella E., De Simone G. & Monti S.M. 2015. Cadmium-containing carbonic anhydrase CDCA1 in marine diatom *Thalassiosira weissflogii*. *Marine Drugs* 13: 1688–1697. DOI: 10.3390/md13041688
- Angeli A., Buonanno M., Donald W.A., Monti S.M. & Supuran C.T. 2018. The zinc – but not cadmium – containing ζ -carbonic from the diatom *Thalassiosira weissflogii* is potently activated by amines and amino acids. *Bioorganic Chemistry* 80: 261–265. DOI: 10.1016/j.bioorg.2018.05.027
- Balmer Y., Koller A., del Val G., Manieri W., Schurmann P. & Buchanan B.B. 2003. Proteomics gives insight into the regulatory function of chloroplast thioredoxins. *Proceedings of the National Academy of Sciences of the United States of America* 100: 370–375. DOI: 10.1073/pnas.232703799
- Balmer Y., Koller A., del Val G., Schurmann P. & Buchanan B.B. 2004. Proteomics uncovers proteins interacting electrostatically with thioredoxin in chloroplasts. *Photosynthesis Research* 79: 275–280. DOI: 10.1023/B:PRES.0000017207.88257.d4
- Beakes G.W., Canter H.M. & Jaworski G.H.M. 1988. Zoospore ultrastructure of *Zygorhizidium affluens* and *Z. planktonicum*, two chytrids parasitizing the diatom *Asterionella-formosa*. *Canadian Journal of Botany* 66: 1054–1067. DOI: 10.1139/b88-151
- Bowes G. & Ogren W.L. 1972. Oxygen inhibition and other properties of soybean Ribulose 1,5-diphosphate carboxylase. *Journal of Biological Chemistry* 247: 2171–2176.
- Bruckner C.G., Bahulikar R., Rahalkar M., Schink B. & Kroth P.G. 2008. Bacteria associated with benthic diatoms from Lake Constance: phylogeny and influences on diatom growth and secretion of extracellular polymeric substances. *Applied and Environmental Microbiology* 74: 7740–7749. DOI: 10.1128/aem.01399-08

- Callebaut I., Labesse G., Durand P., Poupon A., Canard L., Chomilier J., Henrissat B. & Mornon J.P. 1997. Deciphering protein sequence information through hydrophobic cluster analysis (HCA): current status and perspectives. *Cellular and Molecular Life Sciences* 53: 621–645. DOI: 10.1007/s000180050082
- Campbell M.S., Holt C., Moore B. & Yandell M. 2014. Genome annotation and curation using MAKER and MAKER-P. *Current Protocols in Bioinformatics* 48: 4.11.1-4.11.39. DOI: 10.1002/0471250953.bi0411s48
- Chin C.-S., Alexander D.H., Marks P., Klammer A.A., Drake J., Heiner C., Clum A., Copeland A., Huddleston J., Eichler E.E., Turner S.W. & Korlach J. 2013. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* 10: 563–569. DOI: 10.1038/nmeth.2474
- Clement R., Dimnet L., Maberly S.C. & Gontero B. 2016. The nature of the CO₂-concentrating mechanisms in a marine diatom, *Thalassiosira pseudonana*. *New Phytologist* 209: 1417–1427. DOI:10.1111/nph.13728
- Clement R., Jensen E., Prioretti L., Maberly S.C. & Gontero B. 2017a. Diversity of CO₂-concentrating mechanisms and responses to CO₂ concentration in marine and freshwater diatoms. *Journal of Experimental Botany* 68: 3925–3935. DOI: 10.1093/jxb/erx035
- Clement R., Lignon S., Mansuelle P., Jensen E., Pophillat M., Lebrun R., Denis Y., Puppo C., Maberly S.C. & Gontero B. 2017b. Responses of the marine diatom *Thalassiosira pseudonana* to changes in CO₂ concentration: a proteomic approach. *Scientific Reports* 7. DOI: 10.1038/srep42333
- Del Prete S., Vullo D., De Luca V., Supuran C.T. & Capasso C. 2014. Biochemical characterization of the delta-carbonic anhydrase from the marine diatom

- Thalassiosira weissflogii*, TveCA. *Journal of Enzyme Inhibition and Medicinal Chemistry* 29: 906–911. DOI: 10.3109/14756366.2013.868599
- Giordano M., Beardall J. & Raven J.A. 2005. CO₂ concentrating mechanisms in algae: Mechanisms, environmental modulation, and evolution. *Annual Review of Plant Biology* 56: 99–131. DOI: 10.1146/annurev.arplant.56.032604.144052
- Hammer K.J., Kragh T. & Sand-Jensen K. 2019. Inorganic carbon promotes photosynthesis, growth, and maximum biomass of phytoplankton in eutrophic water bodies. *Freshwater Biology* 64: 1956–1970. DOI: 10.1111/fwb.13385
- Hopkinson B.M., Dupont C.L., Allen A.E. & Morel F.M.M. 2011. Efficiency of the CO₂-concentrating mechanism of diatoms. *Proceedings of the National Academy of Sciences of the United States of America* 108: 3830–3837. DOI: 10.1073/pnas.1018062108
- Hopkinson B.M., Dupont C.L. & Matsuda Y. 2016. The physiology and genetics of CO₂ concentrating mechanisms in model diatoms. *Current Opinion in Plant Biology* 31: 51–57. DOI: 10.1016/j.pbi.2016.03.013
- Huang W., Han S., Jiang H., Gu S., Li W., Gontero B. & Maberly S.C. 2020. External α -carbonic anhydrase and solute carrier 4 are required for bicarbonate uptake in a freshwater angiosperm. *Journal of Experimental Botany* 71: 6004–6014. DOI: 10.1093/jxb/eraa351
- Iversen L. L., Winkel A., Baastrup-Spohr L., Hinke A. B., Alahuhta J., Baattrup-Pedersen A., Birk S., Brodersen P., Chambers P. A., Ecke F. *et al.* 2019. Catchment properties and the photosynthetic trait composition of freshwater plant communities. *Science* 366: 878. DOI: 10.1126/science.aay5945
- Jensen E., Clement R., Maberly S.C. & Gontero B. 2017. Regulation of the Calvin - Benson - Bassham cycle in the enigmatic diatoms: biochemical and evolutionary variations on

- an original theme. *Philosophical Transactions of the Royal Society B-Biological Sciences* 372. DOI: 10.1098/rstb.2016.0401
- Jensen E.L., Clement R., Kosta A., Maberly S.C. & Gontero B. 2019. A new widespread subclass of carbonic anhydrase in marine phytoplankton. *ISME Journal* 13: 2094–20106. DOI: 10.1038/s41396-019-0426-8
- Jensen E.L., Maberly S.C. & Gontero B. 2020. Insights on the functions and ecophysiological relevance of the diverse carbonic anhydrases in microalgae. *International Journal of Molecular Sciences* 21. DOI: 10.3390/ijms21082922
- Kaplan A., Badger M.R. & Berry J.A. 1980. Photosynthesis and the intracellular inorganic carbon pool in the bluegreen alga *Anabaena variabilis* - response to external CO₂ concentration. *Planta* 149: 219–226. DOI: 10.1007/bf00384557
- 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*. *Proceedings of the National Academy of Sciences of the United States of America* 113: 9828–9833. DOI: 10.1073/pnas.1603112113
- Kikutani S., Tanaka R., Yamazaki Y., Hara S., Hisabori T., Kroth P.G. & Matsuda Y. 2012. Redox regulation of carbonic anhydrases via thioredoxin in chloroplast of the marine diatom *Phaeodactylum tricorutum*. *Journal of Biological Chemistry* 287: 20689–23700. DOI: 10.1074/jbc.M111.322743
- Kitao Y. & Matsuda Y. 2009. Formation of macromolecular complexes of carbonic anhydrases in the chloroplast of a marine diatom by the action of the C-terminal helix. *Biochemical Journal* 419: 681–688. DOI: 10.1042/bj20082315
- Kragh T. & Sand-Jensen K. 2018. Carbon limitation of lake productivity. *Proceedings of the Royal Society B-Biological Sciences* 285: 20181415. DOI: 10.1098/rspb.2018.1415

- Lane T.W., Saito M.A., George G.N., Pickering I.J., Prince R.C. & Morel F.M.M. 2005. A cadmium enzyme from a marine diatom. *Nature* 435: 42–42. DOI: 10.1038/435042a
- Launay H., Huang W. M., Maberly S. C. & Gontero B. 2020. Regulation of carbon metabolism by environmental conditions: a perspective from diatoms and other chromalveolates. *Frontiers in Plant Science* 11. DOI: 10.3389/fpls.2020.01033
- Losh J. L., Young J. N. & Morel F. M. M. 2013. Rubisco is a small fraction of total protein in marine phytoplankton. *New Phytologist* 198: 52–58. DOI: 10.1111/nph.12143
- Low-Decarie E., Bell G. & Fussmann G.F. 2015. CO₂ alters community composition and response to nutrient enrichment of freshwater phytoplankton. *Oecologia* 177: 875–883. DOI: 10.1007/s00442-014-3153-x
- Low-Decarie E., Fussmann G.F. & Bell G. 2011. The effect of elevated CO₂ on growth and competition in experimental phytoplankton communities. *Global Change Biology* 17: 2525–2535. DOI: 10.1111/j.1365-2486.2011.02402.x
- Low-Decarie E., Fussmann G.F. & Bell G. 2014. Aquatic primary production in a high-CO₂ world. *Trends in Ecology & Evolution* 29: 223–232. DOI:10.1016/j.tree.2014.02.006
- Lund J.W.G. 1949. Studies on *Asterionella*: I. The origin and nature of the cells producing seasonal maxima. *Journal of Ecology* 37: 389–419.
- Lund J.W.G. 1959. A simple counting chamber for nanoplankton. *Limnology and Oceanography* 4: 57–65. DOI: 10.4319/lo.1959.4.1.0057
- Maberly S.C. 1996. Diel, episodic and seasonal changes in pH and concentrations of inorganic carbon in a productive lake. *Freshwater Biology* 35: 579–598.
- Maberly S.C., Courcelle C., Groben R. & Gontero B. 2010. Phylogenetically-based variation in the regulation of the Calvin cycle enzymes, phosphoribulokinase and glyceraldehyde-3-phosphate dehydrogenase, in algae. *Journal of Experimental Botany* 61: 735–745. DOI: 10.1093/jxb/erp337

- Maberly S.C., De Ville M.M., Feuchtmayr H., Jones I.D., Mackay E B., May L., Thackeray S.J. & Winfield I.J. 2011. The limnology of Esthwaite Water: historical change and its causes, current state and prospects for the future. UK Centre for Ecology & Hydrology. 155 pp.
- Maberly S.C. & Gontero B. 2017. Ecological imperatives for aquatic CO₂-concentrating mechanisms. *Journal of Experimental Botany* 68: 3797–3814. DOI: 10.1093/jxb/erx201
- Maberly S.C., Hurley M.A., Butterwick C., Corry J.E., Heaney S.I., Irish A.E., Jaworski G.H.M., Lund J.W.G., Reynolds C.S. & Roscoe J.V. 1994. The rise and fall of *Asterionella formosa* in the south basin of windermere - analysis of a 45-year series of data. *Freshwater Biology* 31: 19–34. DOI: 10.1111/j.1365-2427.1994.tb00835.x
- Mackereth F.J.H., Heron J. & Talling J.F. 1989. *Water analysis: some revised methods for limnologists*. Titus Wilson, Kendal, UK. [ADD PAGINATION]
- Mackinder L.C.M., Meyer M.T., Mettler-Altmann T., Chen V.K., Mitchell M.C., Caspari O., Rosenzweig E.S.F., Pallesen L., Reeves G., Itakura A. *et al.* 2016. A repeat protein links Rubisco to form the eukaryotic carbon-concentrating organelle. *Proceedings of the National Academy of Sciences of the United States of America* 113: 5958–5963. DOI: 10.1073/pnas.1522866113
- Marchand C., Le Marechal P., Meyer Y., Miginiac-Maslow M., Issakidis-Bourguet E. & Decottignies P. 2004. New targets of *Arabidopsis* thioredoxins revealed by proteomic analysis. *Proteomics* 4: 2696–2706. DOI: 10.1002/pmic.200400805
- Matsuda Y., Hopkinson B.M., Nakajima K., Dupont C.L. & Tsuji Y. 2017. Mechanisms of carbon dioxide acquisition and CO₂ sensing in marine diatoms: a gateway to carbon metabolism. *Philosophical Transactions of the Royal Society B-Biological Sciences* 372. DOI: 10.1098/rstb.2016.0403

- Medlin L. K. & Desdevises Y. 2016. Phylogeny of 'araphid' diatoms inferred from SSU and LSU rDNA, RBCL and PSBA sequences. *Vie Et Milieu-Life and Environment* 66:129.
- Medlin L. K. & Desdevises Y. 2020. Phylogenetic reconstruction of diatoms using a seven-gene dataset, multiple outgroups, and morphological data for a total evidence approach. *Phycologia* 59: 422–436. DOI: 10.1080/00318884.2020.1795962
- Meinshausen M., Smith S.J., Calvin K., Daniel J.S., Kainuma M.L.T., Lamarque J.F., Matsumoto K., Montzka S.A., Raper S.C.B., Riahi K. *et al.* 2011. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change* 109: 213. DOI: 10.1007/s10584-011-0156-z
- Mekhalfi M., Avilan L., Lebrun R., Botebol H. & Gontero B. 2012. Consequences of the presence of 24-epibrassinolide, on cultures of a diatom, *Asterionella formosa*. *Biochimie* 94: 1213–1220. DOI: 10.1016/j.biochi.2012.02.011
- Meyer M.T., Whittaker C. & Griffiths H. 2017. The algal pyrenoid: key unanswered questions. *Journal of Experimental Botany* 68: 3739–3749. DOI: 10.1093/jxb/erx178:
- Michelet L., Zaffagnini M., Marchand C., Collin V., Decottignies P., Tsan P., Lancelin J.-M., Trost P., Miginiac-Maslow M., Noctor G. & Lemaire S.D. 2005. Glutathionylation of chloroplast thioredoxin f is a redox signaling mechanism in plants. *Proceedings of the National Academy of Sciences of the United States of America* 102: 16478–1683. DOI: 10.1073/pnas.0507498102
- Michels A.K., Wedel N. & Kroth P.G. 2005. Diatom plastids possess a phosphoribulokinase with an altered regulation and no oxidative pentose phosphate pathway. *Plant Physiology* 137: 911–920. DOI: 10.1104/pp.104.055285

- Morel F.M.M., Lam P.J. & Saito M.A. 2020. Trace metal substitution in marine phytoplankton. *Annual Review of Earth and Planetary Sciences* 48: 491–517. DOI: 10.1146/annurev-earth-053018-060108
- Nakajima K., Tanaka A. & Matsuda Y. 2013. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proceedings of the National Academy of Sciences of the United States of America* 110: 1767–1772. DOI: 10.1073/pnas.1216234110
- Neale P.J., Talling J.F., Heaney S.I., Reynolds C.S. & Lund J.W.G. 1991. Long-time series from the English Lake District: Irradiance-dependent phytoplankton dynamics during the spring maximum. *Limnology and Oceanography* 36: 751–760. DOI: 10.4319/lo.1991.36.4.0751
- Park H., Song B. & Morel F.M.M. 2007. Diversity of the cadmium-containing carbonic anhydrase in marine diatoms and natural waters. *Environmental Microbiology* 9: 403–413. DOI: 10.1111/j.1462-2920.2006.01151.x:
- Price G.D. 2011. Inorganic carbon transporters of the cyanobacterial CO₂ concentrating mechanism. *Photosynthesis Research* 109: 47–57. DOI: 10.1007/s11120-010-9608-y:
- Raven J.A., Gobler C.J. & Hansen P.J. 2020. Dynamic CO₂ and pH levels in coastal, estuarine, and inland waters: theoretical and observed effects on harmful algal blooms. *Harmful Algae* 91: 101594. DOI: 10.1016/j.hal.2019.03.012
- Roberts S.B., Lane T.W. & Morel F.M.M. 1997. Carbonic anhydrase in the marine diatom *Thalassiosira weissflogii* (Bacillariophyceae). *Journal of Phycology* 33: 845–850. DOI: 10.1111/j.0022-3646.1997.00845.x
- Romero M.F., Chen A.-P., Parker M.D. & Boron W.F. 2013. The SLC4 family of bicarbonate (HCO₃⁻) transporters. *Molecular Aspects of Medicine* 34: 159–182. DOI: 10.1016/j.mam.2012.10.008

- Samukawa M., Shen C., Hopkinson B.M. & Matsuda Y. 2014. Localization of putative carbonic anhydrases in the marine diatom, *Thalassiosira pseudonana*. *Photosynthesis Research* 121: 235–249. DOI: 10.1007/s11120-014-9967-x
- Satoh D., Hiraoka Y., Colman B. & Matsuda Y. 2001. Physiological and molecular biological characterization of intracellular carbonic anhydrase from the marine diatom *Phaeodactylum tricornutum*. *Plant Physiology* 126: 1459–1470. DOI: 10.1104/pp.126.4.1459
- Shapiro J. 1997. The role of carbon dioxide in the initiation and maintenance of blue-green dominance in lakes. *Freshwater Biology* 37: 307–223. DOI: 10.1046/j.1365-2427.1997.00164.x
- Tachibana M., Allen A.E., Kikutani S., Endo Y., Bowler C. & Matsuda Y. 2011. Localization of putative carbonic anhydrases in two marine diatoms, *Phaeodactylum tricornutum* and *Thalassiosira pseudonana*. *Photosynthesis Research* 109: 205–221. DOI: 10.1007/s11120-011-9634-4
- Talling J. F. 1976. Depletion of carbon dioxide from lake water by phytoplankton. *Journal of Ecology* 64:79-121. DOI: 10.2307/2258685
- Talling J.F. 1979. Factor interactions and implications for the prediction of lake metabolism. *Archive für Hydrobiologie Beihefte Ergebnisse Limnologie* 13: 96–109.
- Tanaka Y., Nakatsuma D., Harada H., Ishida M. & Matsuda Y. 2005. Localization of soluble beta-carbonic anhydrase in the marine diatom *Phaeodactylum tricornutum*. Sorting to the chloroplast and cluster formation on the girdle lamellae. *Plant Physiology* 138: 207–217. DOI: 10.1104/pp.104.058982
- Tsuji Y., Nakajima K. & Matsuda Y. 2017. Molecular aspects of the biophysical CO₂-concentrating mechanism and its regulation in marine diatoms. *Journal of Experimental Botany* 68: 3763–3772. DOI: 10.1093/jxb/erx173

- Wang Y.D., Huang C.M., Sun B.N.A., Quan C., Wu J.Y. & Lin Z.C. 2014. Paleo-CO₂ variation trends and the Cretaceous greenhouse climate. *Earth-Science Reviews* 129: 136–147. DOI: 10.1016/j.earscirev.2013.11.001
- Weber T., Gruber A. & Kroth P.G. 2009. The presence and localization of thioredoxins in diatoms, unicellular algae of secondary endosymbiotic origin. *Molecular Plant* 2: 468–477. DOI: 10.1093/mp/ssp010
- Wilhelm C., Buechel C., Fisahn J., Goss R., Jakob T., LaRoche J., Lavaud J., Lohr M., Riebesell U., Stehfest K. *et al.* 2006. The regulation of carbon and nutrient assimilation in diatoms is significantly different from green algae. *Protist* 157: 91–124. DOI: 10.1016/j.protis.2006.02.003
- Young J.N., Heureux A.M.C., Sharwood R.E., Rickaby R.E.M., Morel F.M.M. & Whitney S.M. 2016. Large variation in the Rubisco kinetics of diatoms reveals diversity among their carbon-concentrating mechanisms. *Journal of Experimental Botany* 67: 3445–356. DOI: 10.1093/jxb/erw163

TABLES

Table 1. Conditions at the end of pH-drift experiments with *Asterionella formosa* grown at two concentrations of CO₂ and three values of pH (mean with standard deviation in parentheses, n = 3).

	Final pH		C _T (mmol·l ⁻¹)		CO ₂ (μmol·l ⁻¹)		HCO ₃ ⁻ (mmol·l ⁻¹)	
	390	1000	390	1000	390	1000	390	1000
pH	ppm	ppm	ppm	ppm	ppm	ppm	ppm	ppm
7.0	10.30	10.25	1.29	1.33	0.05	0.06	0.47	0.53
	(0.15)	(0.15)	(0.01)	(0.04)	(0.00)	(0.02)	(0.02)	(0.06)
7.5	10.24	10.33	1.34	1.28	0.08	0.05	0.54	0.45
	(0.24)	(0.16)	(0.08)	(0.00)	(0.02)	(0.00)	(0.10)	(0.01)
8.0	9.61	9.83	1.79	1.63	0.41	0.28	1.31	1.01
	(0.18)	(0.06)	(0.12)	(0.02)	(0.00)	(0.03)	(0.23)	(0.04)

Table 2. Identification of DNA sequences coding CA or SLC4 from the *Asterionella formosa* nuclear genome.

Protein ¹	Best hit in Heterokonta	Id %	Location ²	Accession number ³
α -CA I	<i>P. tricornutum</i> / <i>Fi. solaris</i>	41.25/ 35.88	PPC (Pt)	AF07424
α -CA III	<i>P. tricornutum</i>	34.40	CER (Pt)	AF06709
α -CA III	<i>P. tricornutum</i>	38.39	CER (Pt)	AF01660
α -CA III	<i>P. tricornutum</i>	28.85	CER (Pt)	AF06411
α -CA III	<i>P. tricornutum</i>	28.85	CER (Pt)	AF14804
α -CA VI	<i>P. tricornutum</i>	46.18	CER (Pt)	AF16403
α -CA VI	<i>P. tricornutum</i>	43.87	CER (Pt)	AF05998
α -CA VI	<i>P. tricornutum</i>	36.08	CER (Pt)	AF08750
β -CA	<i>P. tricornutum</i>	54.61	Chlp (H)/ Pyr (Pt)	AF09251
β -CA	<i>P. tricornutum</i>	42.44	Chlp (H)/ Pyr (Pt)	AF16830
γ -CA	<i>Fr. cylindrus</i>	66.67	Mit (At)	AF16529
γ -CA	<i>Fr. cylindrus</i>	61.51	Mit (At)	AF09968
γ -CA	<i>P. tricornutum</i>	54.73	Mit (At)	AF02107
θ -CA	<i>P. tricornutum</i>	48.35	Oth (H)/ Pyr (Pt)	AF07496
θ -CA	<i>P. tricornutum</i>	50.35	Oth (H)/ Pyr (Pt)	AF06315
θ -CA	<i>P. tricornutum</i>	48.99	Oth (H)/ Pyr (Pt)	AF09289
θ -CA	<i>P. tricornutum</i>	55.49	Oth (H)/ Pyr (Pt)	AF14017
SLC4-6	<i>Fi. solaris</i> / <i>P. tricornutum</i>	68.40/ 68.38	Chlp (Pt)	AF09764
SLC4-6?	<i>P. tricornutum</i>	*	-	AF11380
SLC4-7	<i>P. tricornutum</i>	74.61	Chlp (H)	AF08596
SLC26-2	<i>P. tricornutum</i>	50.28	PM? (Pt)	AF11040

¹SLC nomenclature follows (Nakajima *et al.* 2013).

²CER, chloroplast endoplasmic reticulum; Chlp, Chloroplast; Mit, Mitochondria; Oth, Other; PM, Plasmamembrane; PPC, Periplastidial; Pyr, Pyrenoid. Letters in parenthesis after the location indicate the source of the information: H, HECTAR; or by comparison with sequences in At, *Arabidopsis thaliana*; or Pt, *Phaeodactylum tricornutum*.

³Spliced coding sequences of each protein are given in Table S1.

*100% identical to AF09764 apart from a deletion.

FIGURE LEGENDS

Figs 1–3. CO₂ concentrations and temporal distribution of two phytoplankton taxa in Esthwaite Water between 1983 and 2014.

Fig. 1. Maximum, mean and minimum weekly CO₂ concentration (μmol l⁻¹).

Fig. 2. Mean weekly cell density of *Asterionella formosa* (cell ml⁻¹) and, for comparison, *Anabaena* spp (filament ml⁻¹).

Fig. 3. % frequency of *A. formosa* or *Anabaena* spp. at different CO₂ concentration classes and the number of dates in each class (dashed line).

Figs 4–5. Growth rate and carbon content of *Asterionella formosa* at two CO₂ concentrations and three pH values, 20°C and 100 μmol photon m⁻² s⁻¹.

Fig. 4. Growth rate.

Fig. 5. Carbon cell quota. Error bars represent one standard deviation.

Figs 6–11. Illustration of the effects of changing pH and concentrations of CO₂ and HCO₃⁻ in Esthwaite Water in 1993 on the CO₂-dependent and HCO₃⁻-dependent rates of oxygen evolution of *Asterionella formosa*.

Fig. 6. Daily maximum and minimum pH.

Fig. 7. Daily maximum and minimum concentrations of CO₂ and HCO₃⁻.

Fig. 8. Daily maximum and minimum rates of CO₂-dependent and HCO₃⁻-dependent photosynthesis.

Fig. 9. % carbon saturation.

Fig. 10. Ratio of CO₂ to HCO₃⁻ uptake rates vs CO₂ concentration (log scale).

Fig. 11. % carbon saturation vs CO₂ concentration (log scale).

Grey bars in Figs **10**, **11** show the temperature-dependent range of air-equilibrium concentrations of CO₂. Surface water temperature was within 5°C of the kinetic measurement temperature of 16°C between days 116 and 287.

Fig. 12. Alignment of β -CAs from *Asterionella formosa* and *Phaeodactylum tricorutum*.

β CAs (AF09251 and AF16380 see Tables 2 and S1) from *A. formosa* were aligned with β CA from *P. tricorutum* (Pt; NCBI accession number AAL07493.1 or JGI Prot ID 51305) using MUSCLE. Shown here are the parts of the sequences, numbered from the N-terminus, that bear the regulatory cysteine residues (empty circles) that are targets of thioredoxins and the residues involved in the Zn^{2+} binding site (full circles).

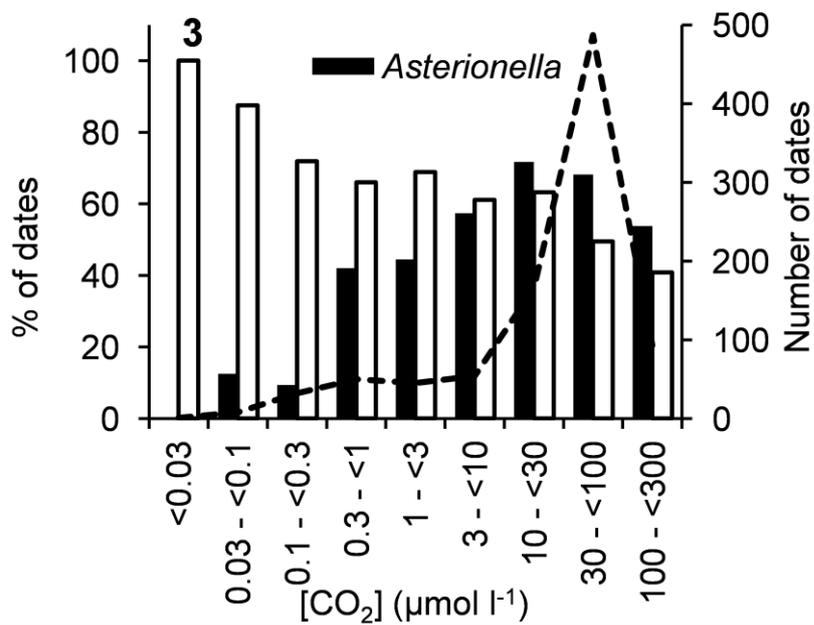
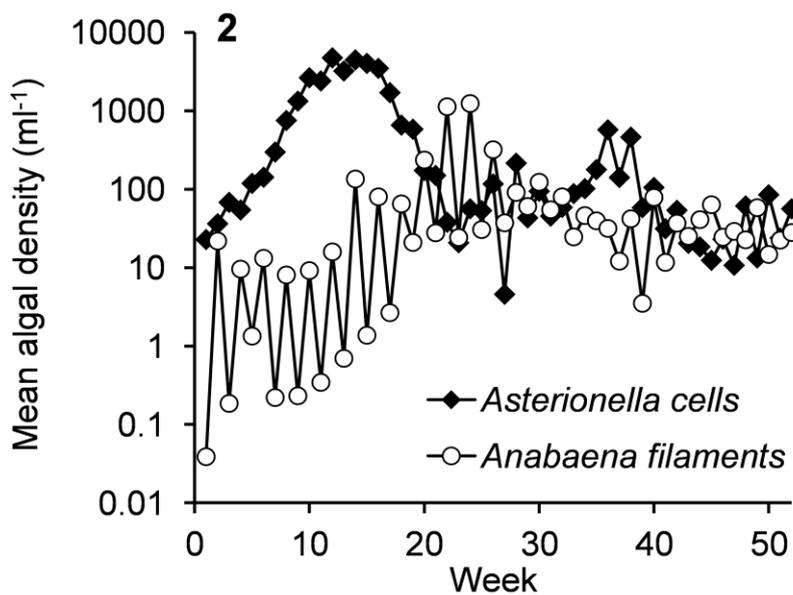
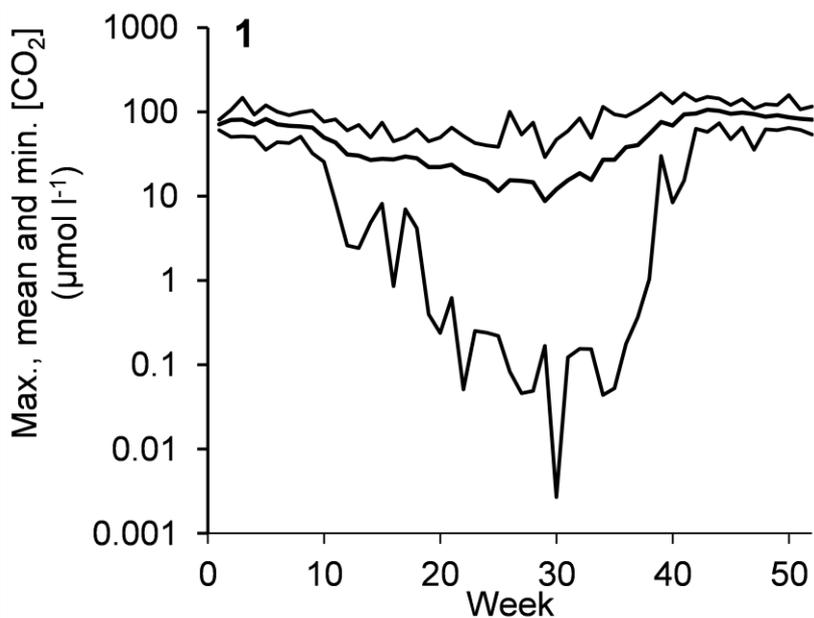
Fig. 13. Alignment of θ -CAs from *Asterionella formosa* and *Phaeodactylum tricorutum*. θ -CAs (AF077496, AF06315, AF09289 and AF14017, Tables 2 and S1) from *A. formosa* were aligned with θ -CA from *P. tricorutum* (Pt; NCBI accession number BAV00142.1) using MUSCLE. The signature of θ -CA is indicated by full circles and the black triangle shows that all sequences except AF09289 contain a G in the canonical motif CCG of θ -CAs.

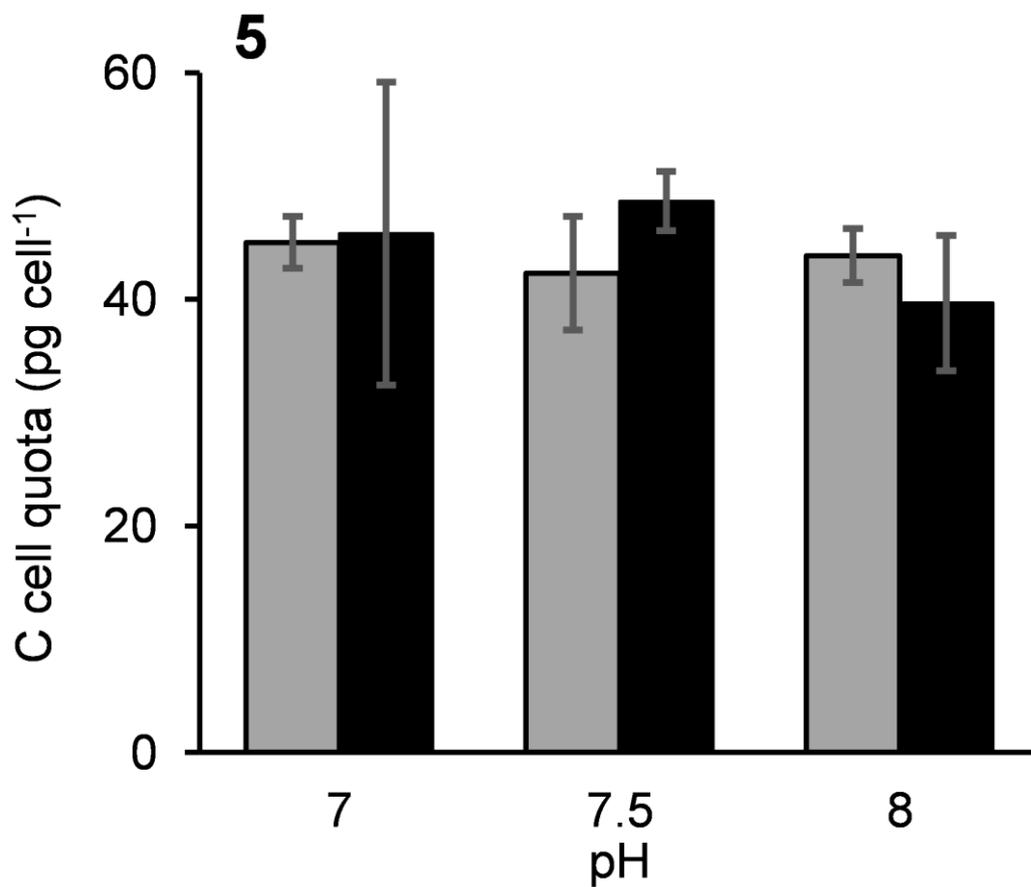
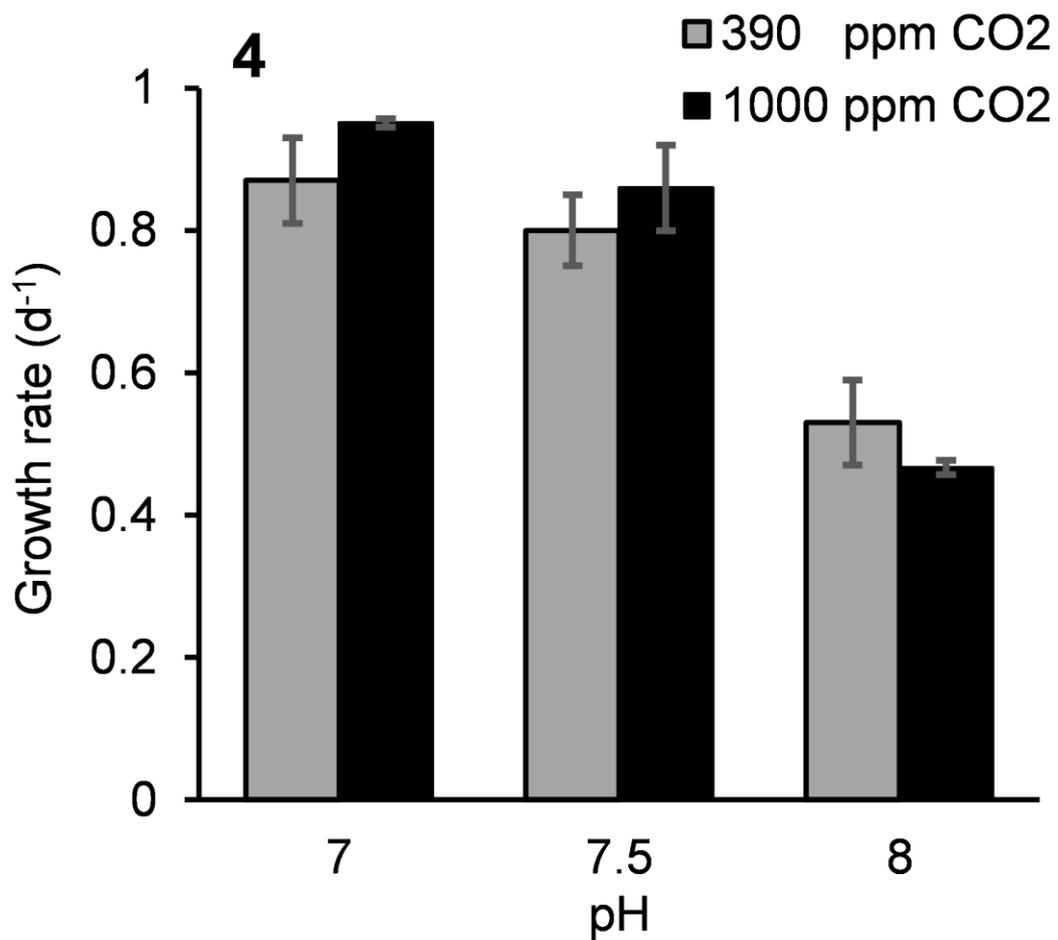
Figs 14–16. Schematic diagrams summarising the presence of proteins linked to the CCM in three diatoms. Putative locations of the proteins, based on sequence homology, are shown for the chloroplast endoplasmic reticulum (CER), chloroplast (Chlp), cytosol (Cyt), mitochondrion (Mit), periplastidial compartment (PPC), periplasmic space (PPS) and pyrenoid (Pyr). Also shown are the cell wall (CW) and plasmamembrane (PM). The scheme for *Phaeodactylum tricorutum* and *Thalassiosira pseudonana* are modified from Tsuji *et al.* (2017 and Jensen *et al.* (2020). Proteins of *Asterionella formosa* are named by homology with those in *P. tricorutum*. α -CA types are shown by Roman numerals.

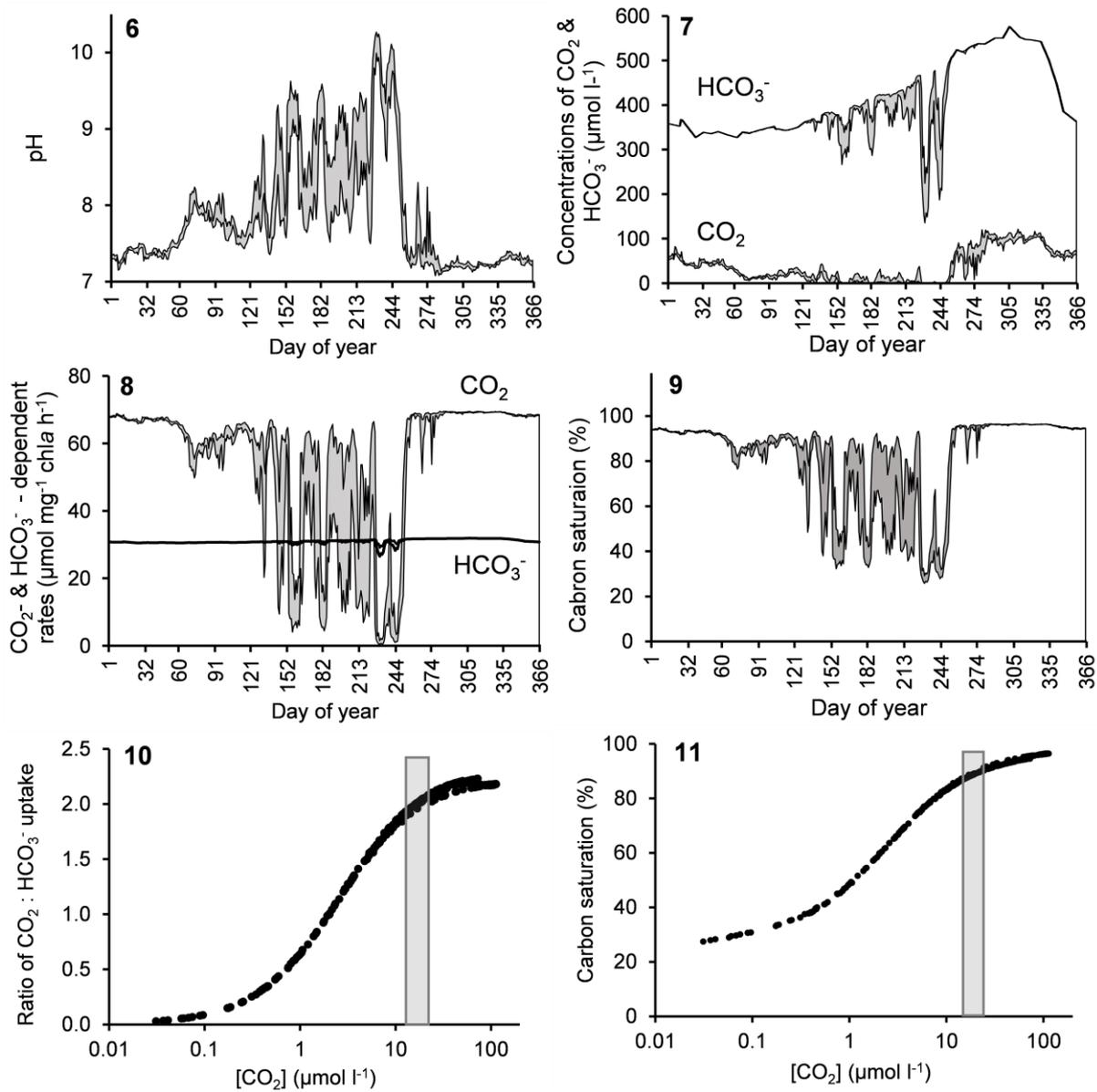
Fig. 14. *Thalassiosira pseudonana*.

Fig. 15. *Phaeodactylum tricorutum*.

Fig. 16. *Asterionella formosa*.







```

Pt      : LYIGCVDAFAEENMIMCTEACTMLITVRNIANMVVNDIAVMSAIQEGINV : 142
AF16830 : MWIGCADARVEANEVVGKLAGTVFVVRNVANMVVNDVNLMAALC----- : 144
AF09251 : MWIGCADARVEANEVVGKLAGTVFVVRNVANMVVNDVNLMAALQAVNV : 149
      ●

Pt      : LKIEHVIVCGHYDCGGVRASYANVDHAPPTSIWLRNIRDVYRIHAREIDA : 192
AF16830 : -----YRIHREKIDS : 154
AF09251 : LKVSIIIVCGHYDCGGIRASMTNKDHAPPTENWLRNIRDVYRIHREKIDS : 199
      ● ● ●

Pt      : IKDEFIRHRRIVELNVIECCVNLKYGVIQAKRIESMCEGFAAAIERVHE : 242
AF16830 : IKDEFIRHRRIVELNVIECCINLFKIGAVQRQRMATYRAGMPYTTTERIHA : 204
AF09251 : IKDEFIRHRRIVELNVIECCINLFKIGAVQRQRMATYRAGMPYTTTERIHA : 249
      ○

Pt      : IVFLEFKTGRIRKLCVDFDKYMSELDIAIYDLYELENKAKIPA : 282
AF16830 : VVFLPITGVLCPLCVNFKKEYISELSAIYDLYSEDDP----- : 240
AF09251 : VVFLPITGVLCPLCVNFKKEYISELSAIYDLYSEDDP----- : 285

```

Fig. 12

```

          360          *          380          *          400
Pt      : IEDGSSCLVVYGPVHGVLDLGNVGVNRRRGREKGGICCGSAFAAAAYISK : 399
AF06315 : IENIGSCLIVFGPHVGIIDSDGIVGRHRRRGRTLIDSCCGSAFAAAAYVAG : 172
AF14017 : IENIGSCLIVFGPHVGIIDSDGIVGRHRRRGRTLIDSCCGSAFAAAAYVAG : 54
AF07496 : IETNGNCIVLFGPHVGVDSDCVVGKVNRRGFLESGCCGSAFAAAAYVHR : 157
AF09289 : IETNGNCIVLFGPHIGVLDACVVGKVNRRGCKHSGICGSAFAAAAYVSS : 149
          ●          ●●●          ●●●

          *          420          *          440          *
Pt      : VFNGEADPPAPAVFESSMDAQCIIYVGNMILPYAERIGNAQDAMVELFYATM : 449
AF06315 : VRSCQHVDAGFETDVLDAEGCFVCGKSILPYGBRLEKFAHFVWELFLCLM : 221
AF14017 : VRSCQHVDAGFETDVLDAEGCFVCGKSILPYGBRLEKFAHSGTGTFF--- : 100
AF07496 : VCRGDCAEHLSVEALDAQCIFVGNMILPYGTRLABEFADEAVELFLALM : 206
AF09289 : VRKEVAKIC-THDIDPELLIQGSFVGNMILQCEHRVPSFTDEAVELSLVLM : 198

          460          *          480          *          500
Pt      : EFLDILMCKIVAKCCGRVGGIGKIALLGGLQINTPAGCFDYFIPLEFEVR : 499
AF06315 : DSQYKLIQRIVARFCNVCVCGNKIALVGGVQINTPSTMSDYFIPLEFEIR : 271
AF14017 : ----LIQRIVARFCNVCVCGNKIALVGGVQINTPSTMSDYFIPLEFEIR : 145
AF07496 : DAQKILMARIVRQVSAVGEAVCIALVGGIQVNTPAGTFDYFVPMTEFLI : 256
AF09289 : EAQRILMRLISSCCGLVEAKSKIALLGGIQVNTPSCIADYFVPTSEKLI : 248

```

Fig. 13

