Protein Abundance data from Zn and CO2 culture experiments on the diatom Phaeodactylum tricornutum

Website: https://www.bco-dmo.org/dataset/1011111
Data Type: experimental
Version: 1
Version Date: 2026-11-03

Project
» Parsing the biogeochemistry of marine carbonic anhydrases (Carbonic Anhydrase)

Contributors	Affiliation	Role
Saito, Mak A.	Woods Hole Oceanographic Institution (WHOI)	Co-Principal Investigator
Subhas, Adam V.	Woods Hole Oceanographic Institution (WHOI)	Co-Principal Investigator
Burdige, Emily	Woods Hole Oceanographic Institution (WHOI)	Student, Contact
Hayden, Matthew G.	Woods Hole Oceanographic Institution (WHOI)	Technician
McIlvin, Matthew R.	Woods Hole Oceanographic Institution (WHOI)	Technician
Soenen, Karen	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
Phaeodactylum tricornutum was grown under 6 different experimental treatments all carried out in biological triplicate. In a CO2 experiment, cultures were grown with replete Zn under varying pCO2 (200, 400, and 750 ppm CO2). In a Zn experiment, cultures were grown under ambient pCO2 (no modification to bubbling) with variable Zinc concentrations (3 nM, 30 nM, and 100 nM [Zn]T) Here, proteomic data from experimental treatments are reported. We specifically look at the protein abundances of 3 relevant groups of proteins: carbonic anhydrase enzymes, bicarbonate transporters, metal stress biomarkers (Fe and Zn) The data collected for this project are used to assess the plasticity of the model diatom Phaeodactylum tricornutum's CCMs, specifically carbonic anhydrase (CA) enzymes and bicarbonate transporters.

Table of Contents

- Dataset Description
  - Methods & Sampling
  - Data Processing Description
  - Problem Description
- Parameters
- Instruments
- Project Information
- Funding

Methods & Sampling

Proteomic Analysis: "Harvested cell pellets were removed from the freezer, resuspended in lysis buffer (2% SDS, 50mM TEAB pH 8.5), and incubated at 95°C and mixed at 350 rmp for 10 (eppendorf; Thermomixer R). The samples were cooled on ice following the heated lysis buffer treatment, and 2 µL of benzonase nuclease (MilliporeSigma 707463) was added to each. This solution was left to incubate at 37°C for 30 min, shaking at 350 rpm. The extract was centrifuged at 14,100 × g for 10 minutes to remove cellular debris, and the extracted proteins in the supernatant were transferred to an ethanol washed tube. After the sample protein extraction, the total protein was quantified using a colorimetric assay at 562 nm (Pierce BCA total protein assay, Thermo Scientific 23225). Samples containing 25 µg of total protein were each subsampled into 200 µL of lysis buffer for the reduction, alkylation, and digestion of proteins.

To alkylate the samples, 4 µL of 500 mM dithiothreitol (DTT; Thermo Fisher Scientific 12-3-3483) in 50 mM ammonium bicarbonate (Ambic; Thermo Fisher Scientific 1066-33-7) was added and incubated at 45°C for 30 minutes. The alkylation was completed with an addition of 12 µL 500mM iodoacetamide (MilliporeSigma 144-48-

9) in 50 mM Ambic, which incubated in the dark for 30 min at room temperature. Finally, the reaction was quenched with a second addition of 4 µL of DTT with a 10 min incubation and acidification using 12% phosphoric acid. Following extraction, reduction, alkylation, and acidification, the samples were transferred to an S-trap column (Protifi, S-Trap Mini Column) to clean, isolate, digest, and elute the proteins for analysis. Briefly, samples were thoroughly washed with binding buffer and a final wash with 90% methanol, digested with trypsin at a ratio of 25 µg protein:1 µg trypsin, left to incubate overnight at 37°C, and eluted with 50 mM Ambic, 50% formic acid, and 50% acetonitrile (ACN, Thermo Fisher Scientific A955-4) in 0.2% formic acid. This volume was quantified, and the purified protein was quantified using the BCA colorimetric assay.

Tryptic peptides were analyzed with liquid chromatography coupled with tandem mass spectrometry (LS/MS/MS) where a Thermo Vanquish Neo HPLC reverse phase chromatography system was paired with a Thermo Scientific Fusion mass spectrometer using a Thermo Flex nanospray source. The peptides were eluted on a reverse phase C18 column (0.1 x 500 mm ID, 2 µm particle size, 120 Å pore size, C18 Reprosil-Gold, Dr. Maisch GmbH) using a 500 nl/min flow rate. All solvents were prepared with Fisher Optima grade reagents. Liquid chromatography was a nonlinear 60-minute gradient from 5%-95% buffer B, where A was 0.1% formic acid in water and B was 0.1% formic acid in ACN. Data independent acquisition (DIA) was used for analysis of the biological triplicates. The DIA scans comprised a precursor range of 380 m/z to 980 m/z on the Orbitrap with 24 m/z MS2 non-overlapping windows. The MS2 scans ranged from 150 to 2000 m/z, and the Orbitrap resolution was 60K for MS1 and 30K for MS2.

The generated mass spectra from DIA analysis were searched against the translated genomes obtained from NCBI using MSFragger within FragPipe (v22.0: https://github.com/Nesvilab/FragPipe). An initial precursor mass tolerance of -10 to 10 ppm, a fragment mass tolerance of 10 ppm, a parent tolerance of 10.0 ppm, and a +57 carbamidomethyl fixed modification of cysteine were used, and MSBooster rescoring was used. Variable modifications were a +16 oxidation modification on methionine and a +42 acetyl modification on the N-terminus. Only peptides of length 6 to 50 and precursor charge states 2 to 4 were included, and a maximum of 1 missed cleavage was allowed. False discovery rates were calculated using decoys generated from a reverse database. A spectral library was generated from MSFragger, and DIA quantification was performed using DIA-NN to obtain a maximum protein false discovery rate of 0.01 (1%)."

Data Processing Description

Replicate measurements as well as calculated average and standard deviations are reported.

Problem Description

One technical replicate in the Mid CO2 treatment identified significantly fewer unique proteins relative to all other samples, likely due to an insufficient protein concentration analyzed. Using a Grubbs' test, we identify that the number of uniquely identified proteins in this replicate is an outlier, and therefore the protein abundances were justifiably neglected when calculating mean and standard deviations in the Mid CO2 treatment.

[table of contents | back to top]

Parameters

Parameters for this dataset have not yet been identified

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Thermo Vanquish Neo HPLC reverse phase chromatography
Generic Instrument Name	High-Performance Liquid Chromatograph
Dataset-specific Description	Thermo Vanquish Neo HPLC reverse phase chromatography: Tryptic peptides were analyzed with liquid chromatography coupled with tandem mass spectrometry (LS/MS/MS)
Generic Instrument Description	A High-performance liquid chromatograph (HPLC) is a type of liquid chromatography used to separate compounds that are dissolved in solution. HPLC instruments consist of a reservoir of the mobile phase, a pump, an injector, a separation column, and a detector. Compounds are separated by high pressure pumping of the sample mixture onto a column packed with microspheres coated with the stationary phase. The different components in the mixture pass through the column at different rates due to differences in their partitioning behavior between the mobile liquid phase and the stationary phase.

Dataset-specific Instrument Name	Thermo Scientific Fusion mass spectrometer
Generic Instrument Name	Mass Spectrometer
Dataset-specific Description	Thermo Scientific Fusion mass spectrometer: Tryptic peptides were analyzed with liquid chromatography coupled with tandem mass spectrometry (LS/MS/MS)
Generic Instrument Description	General term for instruments used to measure the mass-to-charge ratio of ions; generally used to find the composition of a sample by generating a mass spectrum representing the masses of sample components.

[table of contents | back to top]

Project Information

Parsing the biogeochemistry of marine carbonic anhydrases (Carbonic Anhydrase)

Coverage: Woods Hole Oceanographic Institution, Bermuda Atlantic Timeseries Station

Although carbon dioxide (CO2) is found both in the atmosphere and in sea water, it does not react quickly with water. This presents a problem for marine organisms that need to wait for CO2 to react with water before using it for photosynthesis. Therefore, organisms use an enzyme to greatly speed up this reaction. The enzyme, known as carbonic anhydrase (CA), is so useful that it is found across the entire tree of life. However, the role of carbonic anhydrase in the CO2 cycle has been hard to measure. In this project, a team of investigators from the Woods Hole Oceanographic Institution will use a combination of instrument techniques and laboratory experiments to investigate the role of different types of CA in marine organisms. The team will identify the different types of CA and their inherent catalytic rates. The team will then identify the metal cofactors present in these enzymes, which are key for enzyme function. This work will deepen our understanding of the biochemical basis for carbon dioxide cycling in marine phytoplankton, and will provide key insights into one of the most ubiquitous and fastest enzymes on the planet. The team will collaborate with Professor Indu Sharma and her students from Hampton University which has just launched a new marine science major program. Team members will visit Hampton University for a workshop on careers in oceanography. In Year 2, Professor Sharma and her students will accompany the team on a research expedition to the Bermuda-Atlantic Time Series and will help design a sampling campaign focused on the shared research interests of Professor Sharma and the investigators.

Carbonic anhydrases (CAs) are ubiquitous enzymes involved in almost every carbon transformation in the marine environment. Every CA uses a metal cofactor in its active site, and the expression of CAs is tied to the uptake and utilization of trace metals such as zinc and cobalt in the surface ocean. We are starting to understand that organisms express many different forms of CA, all of which serve distinct physiological roles; exhibit unique catalytic rates; and can use different metal cofactors at the enzyme active site. In this proposal, the team will exploit this enzymatic diversity to investigate carbonic anhydrase kinetics in marine diatoms. By combining an isotope-based enzyme activity assay with quantitative proteomics, they will document the inherent kinetics of multiple CA isoforms. In addition, they will identify the metal cofactors associated with the most abundant CAs, therefore providing a direct linkage between enzyme activity and trace metal utilization.

This award reflects NSF's statutory mission and has been deemed worthy of support through evaluation using the Foundation's intellectual merit and broader impacts review criteria.

[table of contents | back to top]

Funding

Funding Source	Award
NSF Division of Ocean Sciences (NSF OCE)	OCE-2123055

[table of contents | back to top]