Semi-quantitative cell proteome of marine Synechococcus WH8102 using DIA-MS, interactive nutrient-temperature responses in stable chemostat bioreactors from laboratory experiments conducted in 2019 at the University of California, Irvine

Website: https://www.bco-dmo.org/dataset/923159
Data Type: experimental
Version: 1
Version Date: 2025-09-11

Project
» Convergence: RAISE: Linking the adaptive dynamics of plankton with emergent global ocean biogeochemistry (Ocean Stoichiometry)
» Collaborative Research: The stoichiometric trait distribution of the marine microbiome (StoichTraitD)

Contributors	Affiliation	Role
Martiny, Adam	University of California-Irvine (UC Irvine)	Principal Investigator
Saito, Mak A.	Woods Hole Oceanographic Institution (WHOI)	Principal Investigator
Garcia, Nathan S.	University of California-Irvine (UC Irvine)	Co-Principal Investigator, Scientist
McIlvin, Matthew R.	Woods Hole Oceanographic Institution (WHOI)	Scientist
Moran, Dawn	Woods Hole Oceanographic Institution (WHOI)	Scientist
York, Amber D.	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

Abstract
This proteomic data set was prepared to determine how the distribution of protein-based traits contribute to cellular elemental stoichiometry of a globally-abundant, surface ocean phytoplankton in response to nutrient stress under variable temperature conditions. Cellular proteins are from laboratory chemostat cultures of marine Synechococcus isolate WH8102 (clade III). Proteins were extracted from cells grown under stable, steady-state conditions in a low-phosphorus medium (N:P=80) and a low-nitrate medium (N:P=1.7) across 3 temperatures of 20, 24, and 28°C (for a total of 6 chemostat cultures) at a white light density of 125 µmol quanta-1 m-2 s-1 on a 12-hour-light:12-hour-dark diel cycle, at a continuous dilution rate of 0.18 d-1, in fall of 2019. Cells and proteins were collected at 5 mid-light-period time points on non-consecutive days to yield a total of 30 samples. Culture work was performed at the University of California, Irvine by Adam Martiny's group. Proteins were analyzed with data-independent acquisition mass spectrometry proteomics methods by Mak Saito's group at the Woods Hole Oceanographic Institution. Data are exclusive peak area intensities of proteins with 2 or more representative peptides, normalized with Scaffold (2.2.1) DIA proteome software. Peptides were analyzed using a Michrom Advance HPLC system coupled to a Q-Exactive mass spectrometer (Thermo Scientific instrument version 2.8) with a Michrom Advance CaptiveSpray source, using the constant injection concentration of 1 µg/µL to allow uniformity across the dataset. Data with more details are published in Garcia et al. (2024) Proteome trait regulation of marine Synechococcus elemental stoichiometry under global change (DOI:10.1093/ismejo/wrae046).

Table of Contents

Coverage

Location: Laboratory in the Department Earth System Science, University of California, Irvine
Temporal Extent: 2019 - 2019

Methods & Sampling

Experimental design

Synechococcus cultures (WH8102) were grown in polycarbonate bottles with a continuous chemostat method used previously (Garcia et al. 2016) in artificial seawater. We used two concentration ratios of macronutrients (NO3-:PO43- = 1.7 and 80) and 3 levels of temperature (20, 24 and 28°C) with a slow dilution rate (0.18 d-1) to ensure treatment-wise culture stability. Target concentrations of NO3- and PO43- in chemostats were 45 µM and 0.56 µM, respectively in the N:P=80 treatment, and 15.9 µM and 9.2 µM, respectively in the N:P=1.7 treatment. White light was supplied at 125 µmol quanta m-2 s-1 on a 12h:12h light:dark cycle. Chemostat equilibria were monitored by measuring culture cell density and forward scatter (*FSCH*) with a Novocyte flow cytometer 1000 (Acea Biosciences, Inc, San Diego, CA). Cells for proteome analysis were collected after an acclimation period on days 38, 43, 47, 50 and 57 with a 47 mm polycarbonate filter (0.2 µm pore size) 7-8 hours into the light period, pelleted by centrifugation (21,130 g for 3 minutes), flash frozen in liquid nitrogen, and stored at -80°C.

Protein extraction and peptide preparation

Proteins were extracted by heating pelleted cells at 95°C for 10 min and gently shaking at room temperature for 30 min in a buffer solution (400 µL – 1760 µL; 50 mM HEPES pH 8.5 (Boston BioProducts #BB-2082), 1% SDS in HPLC grade water) before centrifuging at 14100 g for 20 min at room temperature and removing the supernatant. Sodium dodecyl sulfate (1%) is a strong detergent for diverse matrices including cell membranes (Hughes et al. 2014). Benzonase nuclease (50 units; Novagen #70746-3) was added to 400 µL extracted protein sample and incubated at 37°C for 30 min. Samples were reduced by adding 20 µL of 200 mM DTT (Fisher #BP172-5) in 50 mM HEPES pH 8.5 at 45°C for 30 min and alkylated with 40 µL of 400 mM iodoacetamide (Acros #122270050) in HEPES pH 8.5 for 30 min at 24°C. The reaction was quenched by adding 40 µL of 200 mM DTT in 50 mM HEPES pH 8.5. SpeedBead Magnetic Carboxylate Modified Particles (GE Healthcare #65152105050250 and #45152105050250) were prepared according to Hughes et al. (2014) and added (20 µg/µL) to 400 µL of extracted protein sample. Samples were incubated with formic acid (pH of 2-3) and washed with ethanol and acetonitrile using a magnetic rack. Protein was measured with the BCA method (Thermo Scientific Micro BCA Protein Assay Kit #23235) and digested overnight at 37°C with 1 part trypsin (Promega #V5280; dissolved in HEPES pH 8.0, 0.5 µg/µL) and 25 parts protein. Peptides were washed with acetonitrile and ethanol using a magnetic rack and diluted to a target concentration of 0.1% trifluoroacetic acid or 1% formic acid and a final concentration of 1 µg/µL.

Mass spectrometry of peptides

Similar to other analyses (Searle et al. 2018), peptides were analyzed using a Michrom Advance HPLC system coupled to a Q-Exactive mass spectrometer (Thermo Scientific instrument version 2.8) with a Michrom Advance CaptiveSpray source, using the constant injection concentration of 1 µg/µL (one microgram per microliter) to allow uniformity across the dataset. Samples were concentrated onto a C18 column (Reprosil-Gold, Dr. Maisch GmbH) and eluted in a non-linear, 200-min gradient of formic acid and acetonitrile buffers. Full MS1 scans were performed (35,000 resolution, 3e6 AGC target, 60 ms maximum IT, 385 to 1015 m/z) with overlapping DIA scans (17,500 resolution, 1e6 AGC target, 60 ms maximum IT, 24.0 m/z isolation windows, normalized collision energy of 27, loop count 25).

Proteomic data analysis

Data-independent acquisition mass spectrometry sample data were analyzed using Scaffold DIA (2.2.1), converted to mzML format (ProteoWizard 3.0.11748, Chambers et al (2012)), and individually searched against

Syn8102_uniprot-proteome_UP000001422.fasta with a peptide and fragment mass tolerance of 10.0 ppm. Percolator (3.01) filtered peptides for a maximum false discovery rate of 0.01. Charged peptides (2-3) with length (6-30) were considered. EncyclopeDIA (0.9.6) selected the 5 highest quality fragment ions for quantitation (Searle et al. 2018).

Data Processing Description

Data-independent acquisition mass spectrometry sample data were analyzed using Scaffold DIA (2.2.1), converted to mzML format (ProteoWizard 3.0.11748, Chambers et al (2012)), and individually searched against Syn8102_uniprot-proteome_UP000001422.fasta with a peptide and fragment mass tolerance of 10.0 ppm. Percolator (3.01) filtered peptides for a maximum false discovery rate of 0.01. Charged peptides (2-3) with length (6-30) were considered. EncyclopeDIA (0.9.6) selected the 5 highest quality fragment ions for quantitation (Searle et al. 2018).

.raw proteomics files are accessible at the PRIDE Archive (https://www.ebi.ac.uk/pride/archive/projects/PXD043180).

BCO-DMO Processing Description

Dataset version 1:

Wide format table: "923159_v1_norm-peak-area_exclusive-intensity-wide-format.xlsx" history:
* This contains the wide format of the dataset (separate normalized peak area columns per Protein_ID_(long_form),Temp_chemostat,Nutrient_chemostat,Treatment_replicate,Chemostat_day). It is the original Excel file provided to BCO-DMO on Sept 11th, 2025.
* This format contains 5 header lines containing sample and treatment metadata values. The data rows start as row 6.
* The last row contains no values except ("#VALUE!" as gene name 1 and protein name.)

Long format table: "923159_v1_norm-peak-area_exclusive-intensity.csv" history:
* This table is in "long" format with one column for Normalized peak area, and treatment/sample metadata contained in additional columns:
"Sample_Number,Temp_chemostat,Nutrient_chemostat,Treatment_replicate,Chemostat_day."
* Sheet 1 of submitted file "Normalized peak areas (exclusive intensity for BCO-DMO).xlsx" (submitted Apr 2nd, 2024) was modified to add a row of sample ID numbers (1-30) before import into the bco-dmo data system, discussed with the file submitter. The last row was not imported, it contained no values except ("#VALUE!" as gene name 1 and protein name.)
* The wide to long transformation was performed with an unpivot operation.
* Values "NaN" were imported as missing data values.
** Missing data values are displayed differently based on the file format you download. They are blank in csv files, "NaN" in MatLab files, etc.
* Column names adjusted to conform to BCO-DMO naming conventions designed to support broad re-use by a variety of research tools and scripting languages. [Only numbers, letters, and underscores. Can not start with a number]

[table of contents | back to top]

Data Files

File	
923159_v1_norm-peak-area_exclusive-intensity.csv	(Comma Separated Values (.csv), 7.23 MB) MD5:0e72a38f21caee0338e52145fe16af8c
Primary data file for dataset ID 923159, version 1 which includes a row per sample measured for each protein. This form of the data is referred to as a long-form data table. See supplemental files for an alternate form of this table (wide-form, rows per protein with additional columns per sample).	

[table of contents | back to top]

Supplemental Files

File
923159_v1_norm-peak-area_exclusive-intensity-wide-format.xlsx (Microsoft Excel, 437.10 KB) MD5:1b7d758f4126fa56411e03a515046680
Normalized peak areas, exclusive intensity (wide format table). This is the same data as in "923159_v1_norm-peak-area_exclusive-intensity.csv," however, this is the "wide" form of the dataset with normalized peak areas (exclusive intensities) in separate columns per (Temp_(deg_C)_chemostat,Nutrient_(N:P)_chemostat,Treatment_replicate,Chemostat_day).

[table of contents | back to top]

Related Publications

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., … Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. Nature Biotechnology, 30(10), 918–920. doi:10.1038/nbt.2377
Software

Garcia, N. S., Bonachela, J. A., & Martiny, A. C. (2016). Interactions between growth-dependent changes in cell size, nutrient supply and cellular elemental stoichiometry of marine Synechococcus. The ISME Journal, 10(11), 2715–2724. https://doi.org/10.1038/ismej.2016.50
Methods

Garcia, N. S., Du, M., Guindani, M., McIlvin, M. R., Moran, D. M., Saito, M. A., & Martiny, A. C. (2024). Proteome trait regulation of marine Synechococcus elemental stoichiometry under global change. The ISME Journal, 18(1). https://doi.org/10.1093/ismejo/wrae046
Results

Hughes, C. S., Foehr, S., Garfield, D. A., Furlong, E. E., Steinmetz, L. M., & Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. Molecular Systems Biology, 10(10), 757. doi:10.15252/msb.20145625
Methods

Searle, B. C., Pino, L. K., Egertson, J. D., Ting, Y. S., Lawrence, R. T., MacLean, B. X., Villén, J., & MacCoss, M. J. (2018). Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. Nature Communications, 9(1). https://doi.org/10.1038/s41467-018-07454-w
Methods

[table of contents | back to top]

Related Datasets

IsRelatedTo

Matthew McIlvin (2024). Temperature (20-28ºC) and nutrient stress (N- and P-stress) on Synechococcus WH8102, pride, V1. https://www.ebi.ac.uk/pride/archive/projects/PXD043180.

ProteomeXchange dataset (2024). PXD043180: Temperature (20-28ºC) and nutrient stress (N- and P-stress) on Synechococcus WH8102. Available from https://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD043180 http://proteomecentral.proteomexchange.org/cgi/GetDataset?ID=PXD043180

[table of contents | back to top]

Parameters

Parameter	Description	Units
Protein_ID_long_form	This is the protein ID (long form) (this is the same as the Raw protein ID)	unitless
Protein_ID_short_form	This is the protein ID (short form)	unitless
Gene_name1	Gene name 1	unitless
Gene_name2	Gene name 2	unitless
Protein_name	Protein name (e.g. "C-phycoerythrin class 2 subunit alpha")	unitless
Molecular_Weight	Molecular Weight of protein	kilodalton (kDa)
Peptide_Count	This is the number of peptides that are identified by the mass-spec method and subsequently used to identify the protein.	unitless
Sample_Number	Sample number (1_through_30)	unitless
Temp_chemostat	Temperature of chemostat	Degrees Celsius
Nutrient_chemostat	Nitrate:phosphate (NO3_to_PO4) input ratio into chemostats	unitless
Treatment_replicate	Treatment replicate (1_through_5)	unitless
Chemostat_day	Sampling day from chemostat start day	days
Norm_peak_area	Normalized peak areas (exclusive intensity). [relative values, unitless]	unitless

[table of contents | back to top]

Instruments

Dataset-specific Instrument Name	Michrom Advance HPLC system
Generic Instrument Name	High-Performance Liquid Chromatograph
Dataset-specific Description	HPLC system and mass spectrometer: A Michrom Advance HPLC system was coupled to a Q-Exactive mass spectrometer (Thermo Scientific instrument version 2.8) with a Michrom Advance CaptiveSpray source.
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	Q-Exactive mass spectrometer
Generic Instrument Name	Mass Spectrometer
Dataset-specific Description	HPLC system and mass spectrometer: A Michrom Advance HPLC system was coupled to a Q-Exactive mass spectrometer (Thermo Scientific instrument version 2.8) with a Michrom Advance CaptiveSpray source.
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.

Dataset-specific Instrument Name	8 L polycarbonate bottles
Generic Instrument Name	no_bcodmo_term
Dataset-specific Description	Synechococcus cultures (WH8102) were grown in polycarbonate bottles.
Generic Instrument Description	No relevant match in BCO-DMO instrument vocabulary.

Dataset-specific Instrument Name	
Generic Instrument Name	no_bcodmo_term
Dataset-specific Description	47 mm polycarbonate filter (0.2 µm pore size, Millipore): Cells for proteome analysis were collected with a 47 mm polycarbonate filter (0.2 µm pore size, Millipore) 7-8 hours into the light period, pelleted by centrifugation in 2 mL centrifuge tubes (21,130 g for 3 minutes), flash frozen in liquid nitrogen, and stored at -80°C.
Generic Instrument Description	No relevant match in BCO-DMO instrument vocabulary.

[table of contents | back to top]

Project Information

Convergence: RAISE: Linking the adaptive dynamics of plankton with emergent global ocean biogeochemistry (Ocean Stoichiometry)

NSF Award Abstract:
Due to their sheer abundance and high activity, microorganisms have the potential to greatly influence how ecosystems are affected by changes in their environment. However, descriptions of microbial physiology and diversity are local and highly complex and thus rarely considered in Earth System Models. Thus, the researchers focus on a convergence research framework that can qualitatively and quantitatively integrate eco-evolutionary changes in microorganisms with global biogeochemistry. Here, the investigators will develop an approach that integrates the knowledge and tools of biologists, mathematicians, engineers, and geoscientists to understand the link between the ocean nutrient and carbon cycles. The integration of data and knowledge from diverse fields will provide a robust, biologically rich, and computationally efficient prediction for the variation in plankton resource requirements and the biogeochemical implications, addressing a fundamental challenge in ocean science. In addition, the project can serve as a road map for many other research groups facing a similar lack of convergence between biology and geoscience.

Traditionally, the cellular elemental ratios of Carbon, Nitrogen, and Phosphorus (C:N:P) of marine communities have been considered static at Redfield proportions but recent studies have demonstrated strong latitudinal variation. Such regional variation may have large - but poorly constrained - implications for marine biodiversity, biogeochemical functioning, and atmospheric carbon dioxide levels. As such, variations in ocean community C:N:P may represent an important biological feedback. Here, the investigators propose a convergence research framework integrating cellular and ecological processes controlling microbial resource allocations with an Earth System model. The approach combines culture experiments and omics measurements to provide a molecular understanding of cellular resource allocations. Using a mathematical framework of increasing complexity describing communicating, moving demes, the team will quantify the extent to which local mixing, environmental heterogeneity and evolution lead to systematic deviations in plankton resource allocations and C:N:P. Optimization tools from engineering science will be used to facilitate the quantitative integration of models and observations across a range of scales and complexity levels. Finally, global ocean modeling will enable understanding of how plankton resource use impacts Earth System processes. By integrating data and knowledge across fields, scales and complexity, the investigators will develop a robust link between variation in plankton C:N:P and global biogeochemical cycles.

Collaborative Research: The stoichiometric trait distribution of the marine microbiome (StoichTraitD)

Coverage: Indian Ocean; Laboratory Ecophysiology

NSF abstract:
The elemental ratios of carbon, nitrogen, and phosphorus (C:N:P) have been considered fixed proportions in marine environments but recent work has demonstrated changes across latitudes. Such variation can have large implications for marine biogeochemistry and atmospheric CO_2 levels. As such, future variations in ocean community C:N:P could be a key feedback to global change. The elemental composition of particulate organic matter (POM) represents the aggregate value of diverse microorganisms as well as non-living particles. However, we currently do not understand how individual cells and particles contribute to observed variation in the C:N:P of POM. This project is determining the biomass C:N:P of individual microbial cells grown under a range of conditions and sampled from diverse ocean biomes. Because different individuals are likely to have different fates (e.g., loss to sinking, lysis, predation), understanding how the trait distribution of microbial biomass C:N:P relates to cell size and trophic mode and how environmental conditions affect each trait's distribution offers new perspectives and insight into marine C, N, and P cycles. This project supports two PhD students and multiple undergraduate students. The PhD students are integrated into existing networks on microbiome science at each institution with opportunities to collaborate across diverse disciplines. Undergraduate students at both institutions are being recruited through existing training programs that target

underrepresented groups. In addition, PI Hall is part of a collaborative of Ecologists and Poets at CSU, that look at ways to translate ecological relationships to non-traditional media to make it more accessible and impactful to the general public. This group is exploring the nature of individuality within the marine microbiome by creating trait distributions of written text and removing different modes of individuals (e.g., words) to better understand and communicate how individuals from the smallest organisms on the planet (marine microorganisms) can have large effects on the surrounding ecosystem (i.e., the planet). Results from this project are being incorporated into future projects of the working group at Colorado State University including public presentations, art installations, and published materials.

The aim of this project is to quantify the relationship between environmental conditions and marine microbial C:N:P by assessing the individuality in microbial elemental stoichiometry. To achieve this the project uses energy dispersive spectroscopy (EDS) to assess the stoichiometric trait distribution of populations and communities under different resource and temperature treatments and oceanographic field work across a broad latitudinal gradient. The researchers hypothesize that the relationship between the stoichiometry of particulate organic matter (POM) and environmental conditions are masked by distinct responses of individual constituents of the marine microbiome. They hypothesize that these distinct responses result in a multi-modal distribution of particulate carbon (C), nitrogen (N), and phosphorus (P) of the marine microbiome. The investigators are characterizing the distribution of three stoichiometric traits (biomass C:N, C:P, and N:P) for 50 marine isolates (both autotrophs and heterotrophs) under different resource and temperature environments. They are characterizing the same trait distribution for marine communities sampled from different regions of the ocean for the same resource and temperature environments as the population experiments. They are participating on an ocean going cruise to characterize stoichiometric trait distribution of the marine microbiome across natural gradients of resources and temperature. Understanding how constituent members of microbial communities alter their biomass in response to environmental change is providing a missing link between the variation in the ocean’s environment and particulate C:N:P ratios for diverse marine environments.

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-1848576
NSF Division of Ocean Sciences (NSF OCE)	OCE-2135035
NSF Division of Integrative Organismal Systems (NSF IOS)	IOS-2137339

[table of contents | back to top]