

# Total spectral count of proteins from R/V Falkor cruise 160115 for the ProteOMZ expedition in the Central Pacific in 2016.

**Website:** <https://www.bco-dmo.org/dataset/737620>

**Data Type:** Cruise Results

**Version:** 3

**Version Date:** 2018-12-10

## Project

» [The ProteOMZ Expedition: Investigating Life Without Oxygen in the Pacific Ocean](#) (ProteOMZ (Proteomics in an Oxygen Minimum Zone))

Contributors	Affiliation	Role
<a href="#">Saito, Mak A.</a>	Woods Hole Oceanographic Institution (WHOI)	Principal Investigator
<a href="#">Ake, Hannah</a>	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager
<a href="#">York, Amber D.</a>	Woods Hole Oceanographic Institution (WHOI BCO-DMO)	BCO-DMO Data Manager

## Abstract

Relative protein abundance data of the upper 1000m water column from the ProteOMZ R/V Falkor expedition. There are 56,577 protein identifications over 103 samples, for ~6 million data points.

---

## Table of Contents

- [Coverage](#)
  - [Dataset Description](#)
    - [Data Processing Description](#)
    - [BCO-DMO Processing Description](#)
  - [Related Datasets](#)
  - [Parameters](#)
  - [Instruments](#)
  - [Deployments](#)
  - [Project Information](#)
  - [Funding](#)
- 

## Coverage

**Spatial Extent:** N:17.4465 E:-139.1089 S:-0.4708 W:-157.3022

**Temporal Extent:** 2016-01-19 - 2016-01-28

---

## Dataset Description

These data are part of the Ocean Protein Portal "ProteOMZ" dataset (<https://proteinportal.whoi.edu/>; Saito et al., 2019).

## Data Processing Description

The raw mass spectra files were searched against SEQUEST within Proteome Discoverer v2.2 software. Processed files were then loaded into Proteome Software and protein and peptide reports as well as fasta files were exported. The files were modified slightly to map to the Protein Portal data model for submission to BCO-DMO.

## BCO-DMO Processing Description

-Date, time, filter min, filter max, lat, lon, and cruise columns were added based on information from the Falkor 160115 Event log and CTD log.

-Column names reformatted to comply with BCO-DMO standards.

Dataset version 3: replaces the earlier data version from date (2018-05-25)

\* Values "#N/A" changed to "Unknown" which has a different meaning than blank values. Unknown = "Protein functional and taxonomic annotations are marked as "Unknown" for protein sequences which did not have any significant hits to known reference sequences or motifs in the metagenome annotation database. "

\* Event log and McLane pump log were updated to fix lat/lon, date/time issues. Since these were sources of information in this dataset, this dataset is also being updated. See respective "processing notes" sections for these two logs for detailed information about changes to those data sources. event log: <https://www.bco-dmo.org/dataset/708384> pump log: <https://www.bco-dmo.org/dataset/708495>

\* ISO\_DateTime\_UTC timestamp added from Date and time columns in the McLane log dataset. Date and time columns were HST time zone so 10 hours were added to make the time in UTC.

\* columns for max and min pump filter size added (min 0.2 max 3.0).

\* some station and target depths for station and target depth combinations did not exist in the McLane pump log so the missing values were added to the mclane pump log. The following columns in this dataset come from the pump log: cruise, cast, date, time, lat, lon, depth (ISO\_DateTime\_UTC is derived from local date and time). Where there were two matched casts in the pump log for a station and target depth, the first cast for the station was used. The only differences between the two possible casts were the date/time columns.

[ [table of contents](#) | [back to top](#) ]

---

## Related Datasets

### IsSupplementedBy

Saito, M. A. (2018) **Total spectral counts of peptides from the R/V Falkor cruise 160115 in the Central Pacific for the ProteOMZ expedition in 2016.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2018-05-25 <http://lod.bco-dmo.org/id/dataset/737596> [[view at BCO-DMO](#)]

*Relationship Description: The peptides related to these proteins*

### IsRelatedTo

Saito, M. A. (2019) **FASTA file of identified protein sequences from the R/V Falkor cruise 160115 for the ProteOMZ expedition in the Central Pacific during 2016.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 2) Version Date 2019-01-02 <http://lod.bco-dmo.org/id/dataset/737611> [[view at BCO-DMO](#)]

*Relationship Description: These datasets are part of the Ocean Protein Portal "ProteOMZ" dataset (<https://proteinportal.whoi.edu/>; Saito et al., 2019).*

Saito, M. A., Saunders, J. (2022) **Relative protein abundance from scaled and corrected exclusive peptide spectral counts from the ProteOMZ R/V Falkor expedition cruise FK160115 in the Pelagic central Pacific Ocean in 2016.** Biological and Chemical Oceanography Data Management Office (BCO-DMO). (Version 1) Version Date 2022-01-13 doi:10.26008/1912/bco-dmo.868030.1 [[view at BCO-DMO](#)]

*Relationship Description: These datasets are part of the Ocean Protein Portal "ProteOMZ" dataset (<https://proteinportal.whoi.edu/>; Saito et al., 2019).*

[ [table of contents](#) | [back to top](#) ]

---

## Parameters

Parameter	Description	Units
row_id	Row number	unitless
Identified_Proteins	JCVI metagenome ID number	unitless
Molecular_Weight	Useful for calculated information	kDa
best_hit_annotation	Product name	unitless
best_hit_species	Taxon	unitless
best_hit_taxon_id	Taxon ID	unitless
KO	Kegg ID	unitless
KO_desc	Kegg name	unitless
KO_pathway	Kegg pathway	unitless
EC	Enzyme Commission ID number	unitless
uniprot	Uniprot database ID number	unitless
PFams	Protein family ID number	unitless
PFams_desc	Protein family description	unitless
cruise	Cruise number	unitless
station	Station number	unitless
date; yyyy/mm/dd	Date at station	unitless
time; hh:mm	Time at station	unitless
lat	Latitude	decimal degrees
lon	Longitude	decimal degrees
cast	Cast number	unitless
depth	Cast depth	meters
spectral_count	Spectral count of each protein	count
min_filter_size	Minimum filter size	microns
max_filter_size	Maximum filter size	microns
ISO_DateTime_UTC	DateTime UTC ISO formatted	unitless

[ [table of contents](#) | [back to top](#) ]

## Instruments

<b>Dataset-specific Instrument Name</b>	Alpkem Autosampler
<b>Generic Instrument Name</b>	Alpkem RFA300
<b>Dataset-specific Description</b>	Used in nutrient analysis
<b>Generic Instrument Description</b>	A rapid flow analyser (RFA) that may be used to measure nutrient concentrations in seawater. It is an air-segmented, continuous flow instrument comprising a sampler, a peristaltic pump which simultaneously pumps samples, reagents and air bubbles through the system, analytical cartridge, heating bath, colorimeter, data station, and printer. The RFA-300 was a precursor to the smaller Alpkem RFA/2 (also RFA II or RFA-2).

<b>Dataset-specific Instrument Name</b>	SeaBird SBE19 CTD
<b>Generic Instrument Name</b>	CTD Sea-Bird
<b>Dataset-specific Description</b>	Used for water sampling
<b>Generic Instrument Description</b>	Conductivity, Temperature, Depth (CTD) sensor package from SeaBird Electronics, no specific unit identified. This instrument designation is used when specific make and model are not known. See also other SeaBird instruments listed under CTD. More information from Sea-Bird Electronics.

<b>Dataset-specific Instrument Name</b>	Technicon AutoAnalyzer II
<b>Generic Instrument Name</b>	Technicon AutoAnalyzer II
<b>Dataset-specific Description</b>	Used to measure phosphate and ammonium
<b>Generic Instrument Description</b>	A rapid flow analyzer that may be used to measure nutrient concentrations in seawater. It is a continuous segmented flow instrument consisting of a sampler, peristaltic pump, analytical cartridge, heating bath, and colorimeter. See more information about this instrument from the manufacturer.

<b>Dataset-specific Instrument Name</b>	Trace Metal Rosette
<b>Generic Instrument Name</b>	Trace Metal Bottle
<b>Dataset-specific Description</b>	Used for nutrient sampling
<b>Generic Instrument Description</b>	Trace metal (TM) clean rosette bottle used for collecting trace metal clean seawater samples.

[ [table of contents](#) | [back to top](#) ]

## Deployments

### FK160115

<b>Website</b>	<a href="https://www.bco-dmo.org/deployment/708387">https://www.bco-dmo.org/deployment/708387</a>
<b>Platform</b>	R/V Falkor
<b>Report</b>	<a href="https://service.rvdata.us/data/cruise/FK160115/doc/FK160115_OfficialCruiseReport_Saito_v3.pdf">https://service.rvdata.us/data/cruise/FK160115/doc/FK160115_OfficialCruiseReport_Saito_v3.pdf</a>
<b>Start Date</b>	2016-01-16
<b>End Date</b>	2016-02-11
<b>Description</b>	Project: Using Proteomics to Understand Oxygen Minimum Zones (ProteOMZ) More information is available from the ship operator at <a href="https://schmidtocean.org/cruise/investigating-life-without-oxygen-in-the...">https://schmidtocean.org/cruise/investigating-life-without-oxygen-in-the...</a> Additional cruise information is available from the Rolling Deck to Repository (R2R): <a href="https://www.rvdata.us/search/cruise/FK160115">https://www.rvdata.us/search/cruise/FK160115</a>

[ [table of contents](#) | [back to top](#) ]

---

## Project Information

### The ProteOMZ Expedition: Investigating Life Without Oxygen in the Pacific Ocean (ProteOMZ (Proteomics in an Oxygen Minimum Zone))

**Website:** <https://schmidtocean.org/cruise/investigating-life-without-oxygen-in-the-tropical-pacific/#team>

**Coverage:** Central Pacific Ocean (Hawaii to Tahiti)

From Schmidt Ocean Institute's ProteOMZ Project page:

Rising temperatures, ocean acidification, and overfishing have now gained widespread notoriety as human-caused phenomena that are changing our seas. In recent years, scientists have increasingly recognized that there is yet another ingredient in that deleterious mix: a process called deoxygenation that results in less oxygen available in our seas.

Large-scale ocean circulation naturally results in low-oxygen areas of the ocean called oxygen deficient zones (ODZs). The cycling of carbon and nutrients – the foundation of marine life, called biogeochemistry – is fundamentally different in ODZs than in oxygen-rich areas. Because researchers think deoxygenation will greatly expand the total area of ODZs over the next 100 years, studying how these areas function now is important in predicting and understanding the oceans of the future. This first expedition of 2016 led by Dr. Mak Saito from the Woods Hole Oceanographic Institution (WHOI) along with scientists from University of Maryland Center for Environmental Science, University of California Santa Cruz, and University of Washington aimed to do just that, investigate ODZs.

During the 28 day voyage named “ProteOMZ,” researchers aboard R/V *Falkor* traveled from Honolulu, Hawaii to Tahiti to describe the biogeochemical processes that occur within this particular swath of the ocean’s ODZs. By doing so, they contributed to our greater understanding of ODZs, gathered a database of baseline measurements to which future measurements can be compared, and established a new methodology that could be used in future research on these expanding ODZs.

[ [table of contents](#) | [back to top](#) ]

---

## Funding

Funding Source	Award
<a href="#">Gordon and Betty Moore Foundation: Marine Microbiology Initiative (MMI)</a>	<a href="#">GBMF3782</a>
Schmidt Ocean Institute (SOI)	<a href="#">R/V Falkor 160115 SOI ProteOMZ Expedition</a>

[ [table of contents](#) | [back to top](#) ]