# Do Cyanobacteria Drive Marine Hydrocarbon Biogeochemistry?

*A Data Management Plan created using DMPTool*

**Creators:** Love, Connor, First Name Surname, First Name Surname, First Name Surname

**Affiliation:** University of California, Santa Barbara (UCSB)

**Funder:** National Science Foundation (NSF)

**Template:** BCO-DMO NSF OCE: Biological and Chemical Oceanography

**ORCID iD:** 0000-0001-5914-9107

**Project abstract:**

While the release of petroleum hydrocarbons into the ocean is recognized as an environmental and human hazard, a recent study has estimated that on an annual basis, the release of natural hydrocarbons by a single phytoplankton group (cyanobacteria) contributes at least ten times more total hydrocarbon to the surface ocean. This project will be the first in-depth study of the latent biogeochemical cycling of this huge pool of biogenic hydrocarbons. Using field studies, laboratory incubations of cyanobacteria, and state-of-the art chemical analysis, the researchers will examine the molecular structures, rates and mechanisms of production and removal, and the environmental conditions that control the cycling of this major pool of oceanic hydrocarbons. The results of this study will reveal significant new knowledge for improved understanding of a major carbon cycle in the ocean. Additionally, data could indicate a role for cyanobacterial hydrocarbons in preparing natural marine bacteria to respond to, and degrade petroleum spills, as well as a possible atmospheric impact (e.g. cloud formation) resulting from air-sea exchange of certain components of the hydrocarbon pool. This project will support undergraduate and graduate students, a postdoctoral investigator, and a new faculty member, and will engage participants from minority-serving institutions in California and North Carolina. Plans are also included to establish links with oil spill and biofuel researchers in order to evaluate additional practical applications for the data resulting from this study. The annual production of 308,000,000 - 771,000,000 tons of hydrocarbons by cyanobacteria has recently been reported and is a factor of 10 larger than marine petroleum hydrocarbon input from spills and natural seeps. Consequently, these biogenic hydrocarbons almost certainly have significant implications for the carbon cycle and the bacterial community composition in the ocean but have never been the subject of rigorous study. This project will investigate the distribution, partitioning, and cycling of biogenic hydrocarbons in the ocean, focusing on the abundance and molecular diversity of biogenic hydrocarbons in relation to cyanobacterial populations; the extent to which volatilization to the atmosphere acts as a sink for biogenic hydrocarbons; and the rate at which hydrocarbons are produced by cyanobacteria and consumed by hydrocarbon-

degrading bacteria. Field studies across natural gradients in phytoplankton community structure and abundance will employ state of the art chemical analysis to evaluate the distribution of biogenic hydrocarbons, and together with incubation experiments will determine quantitative rates for biogenic hydrocarbons cycling in the surface ocean. Laboratory studies will augment field studies by assessing hydrocarbon production and loss mechanisms under carefully controlled laboratory conditions. Together, the project will obtain a quantitative understanding of this important component of the oceanic carbon cycle.

**Last modified:** 07-24-2020

# Do Cyanobacteria Drive Marine Hydrocarbon Biogeochemistry?

## Data Policy Compliance

**Identify any published data policies with which the project will comply, including the NSF OCE Data and Sample Policy as well as other policies that may be relevant if the project is part of a large coordinated research program (e.g. GEOTRACES).**

The project investigators will comply with the data management and dissemination policies described in the NSF Award and Administration Guide (AAG, Chapter VI.D.4) and the NSF Division of Ocean Sciences Sample and Data Policy.

## Pre-Cruise Planning

**If the proposed project involves a research cruise, describe the cruise plans. (Skip this section if it is not relevant to your proposal.) Consider the following questions:**

1. **How will pre-cruise planning be coordinated? (e.g. email, teleconference, workshop)**
2. **What types of sampling instruments will be deployed on the cruise?**
3. **How will the cruise event log be recorded? (e.g. the Rolling Deck to Repository (R2R) event logger application, an Excel spreadsheet, or paper logs)**
4. **Will you prepare a cruise report?**

Pre-cruise planning will be done via teleconferencing and email. Sample instruments deployed will be a CTD-rosette, as well as sediment traps and net tows. Sampling events will be recorded on paper logs and scanned into PDF documents. There will not be a cruise report.

## Description of Data Types

**Provide a description of the types of data to be produced during the project. Identify the types of data, samples, physical collections, software, derived models, curriculum materials, and other materials to be produced in the course of the project. Include a description of the location of collection, collection methods and instruments, expected dates or duration of collection. If you will be using existing datasets, state this and include how you will obtain them.**

The project will produce several observational and experimental datasets described below. Data will be collected on a North Atlantic research cruise that will take place in May.

**Observations and Datasets:**

1. **CTD and Niskin bottle data**: CTD data collected using a SeaBird SBE CTD package; processing to be done using SeaBird's SeaSave software; data will include standard environmental measurements (such as pressure, temperature, salinity, fluorescence) as well as measurements for cyanobacteria hydrocarbon concentrations, cell abundance and nutrients. File types: Raw (.con, .hdr, .hex, .bl) and processed and .cnv, .asc, .btl) ASCII files. Repository: BCO-DMO
2. **Event log:** Cruise scientific sampling event log; will include event numbers, start/end dates, times & locations of instrument deployments. Will be recorded using the R2R event logger (if available) and on paper log sheets. File types: Excel file converted to .csv; scanned PDFs. Repository: BCO-DMO and Rolling Deck to Repository (R2R).
3. **Observational Measurements:** Observational data will consist of membrane-bound hydrocarbon concentrations, phytoplankton cell abundance, and nutrients.

**Experimental Datasets:**

1. **Hydrocarbon Production:** Hydrocarbon production rates will be found using on-board seawater incubators for ~36 hour incubations, using stable isotope 13C bicarbonate as a photosynthetic tracer and compound-specific isotope ratios to calculate production rates. Experiments will be conducted on board the ship and in the ships lab. File types: Excel file(s). Repository: BCO-DMO.
2. **Hydrocarbon Consumption:** Hydrocarbon consumption data will consist of changes in dissolved oxygen of no-headspace incubation bottles with varying hydrocarbon and particle additions. This data will be supplemented by nutrient, cell count and sequncing data.
3. **Genetic sequencing:** mRNA and DNA sequences from animals collected at sea. Sequencing will be performed at the PI's lab in University of California, Santa Barbara following the research cruise. File types: Short-read archive (.sra) and .fasta files. Repository: NCBI; accession numbers to be provided to BCO-DMO.

## Data and Metadata Formats and Standards

**Identify the formats and standards to be used for data and metadata formatting and content. Where existing standards are absent or deemed inadequate, these formats and contents should be documented along with any proposed solutions or remedies. Consider the following questions:**

1. **Which file formats will be used to store your data?**
2. **What type of contextual details (metadata) will you document and how?**
3. **Are there specific data or metadata standards that you will be adhering to?**
4. **Will you be using or creating a data dictionary, code list, or glossary?**
5. **What types of quality control will be used? How will data quality be assessed and flagged?**

Data will be stored in csv files. Field data will include date, time, latitude, longitude, cast number, depth, as appropriate. Quality flags will be assigned according to the ODS IODE Quality Flag scheme (IOC Manuals and Guides, 54, volume 3; http://www.iode.org/mg54_3 ). Metadata will be prepared in accordance with BCO-DMO conventions (i.e. using the BCO-DMO metadata forms) and will include detailed descriptions of collection and analysis procedures.

## Data Storage and Access During the Project

**Describe how project data will be stored, accessed, and shared among project participants during the course of the project. Consider the following:**

1. **How will data be shared among project participants during the data collection and analysis phases? (e.g. web page, shared network drive)**
2. **How/where will data be stored and backed-up?**
3. **If data volumes will be significant, what is the estimated total file size?**

The investigators will store project data on laboratory computers that are backed up by university funded cloud storage (Box). Data will be shared via email and associated cloud storage.

## Mechanisms and Policies for Access, Sharing, Re-Use, and Re-Distribution

**Describe mechanisms for data access and sharing, and describe any related policies and provisions for re-use, re-distribution, and the production of derivatives. Include provisions for appropriate protections of privacy, confidentiality, security, intellectual property, or other rights or requirements. Consider the following:**

1. **When will data be made publicly available and how? Identify the data repositories you plan to use to make data available.**
2. **Are the data sensitive in nature (e.g. endangered species concerns, potential patentability)? If so, is public**

**access inappropriate and how will access be provided? (e.g. formal consent agreements, restricted access)**

3. **Will any permission restrictions (such as an embargo period) need to be placed on the data? If so, what are the reasons and what is the duration of the embargo?**
4. **Who holds intellectual property rights to the data and how might this affect data access?**
5. **Who is likely to be interested in re-using the data? What are the foreseeable re-uses of the data?**

Data will made publicly available at the end of 2020, utilizing the BCO-DMO data system and sequences will be made available on National Center for Biotechnology Information (NCBI) database GenBank. Data are not sensitive in nature. No permission restrictions will be need to be placed on the data.

Data produced by this project may be of interest to chemical and biological oceanographers, and oil spill recovery scientists interested in the role of biogeochemistry in natural oil spill remediation. We will adhere to and promote the standards, policies, and provisions for data and metadata submission, access, re-use, distribution, and ownership as prescribed by the BCO-DMO Terms of Use (http://www.bco-dmo.org/terms-use).

## Plans for Archiving

**Describe the plans for long-term archiving of data, samples, and other research products, and for preservation of access to them. Consider the following:**

1. **What is your long-term strategy for maintaining, curating, and archiving the data?**
2. **What archive(s) have you identified as a place to deposit data and other research products?**

BCO-DMO will also ensure that project data are submitted to the appropriate national data archive.

## Roles and Responsibilities

**Describe the roles and responsibilities of all parties with respect to the management of the data. Consider the following:**

1. **If there are multiple investigators involved, what are the data management responsibilities of each person**
2. **Who will be the lead or primary person responsible for ultimately ensuring compliance with the Data Management Plan?**

Data management responsibilities will mainly fall to the principal investigator David Valentine and his graduate student Connor Love. Connor Love will be the primary person responsible for ensuring compliance with the Data Management Plan.