DATA MANAGEMENT PLAN

I. Project information

1. *Project title*: "Biopolymers produced by diatoms and coccolithophores as carriers for selected natural radionuclides (of Th, Pa, Pb, Po, Be) in the ocean

Responsible data manager and points of contact: Peter Santschi, santschi@tamug.edu; Kathleen Schwehr, schwehrk@tamug.edu; Chen Xu, xuchen66@hotmail.com; Antonietta Quigg, quigga@tamug.edu

II. Types of data: chemical and physicochemical analyses; HPLC, GC-MS, ICP-MS (all at TAMUG), ESI-FTICR-MS and NMR (at ODU).

Data description

1. What is the data and information output expected from this grant?

All images will be stored as *.tiff files. Other processed data in Excel files.

Samples generated by this project will be stored in the PIs laboratories for the duration of this project and made available to other researchers upon request when available, the latest after publication of data derived from those samples. The data will include measurements of organic matter compound classes (hydroxamate and catecholate siderophores, proteins, carbohydrates, amino acids) data. These data will be accompanied by detailed metadata files documenting relevant information about sample collection procedures, data quality (standard curves, measures of analytical error, etc.), transport/handling procedures (when samples are shared across co-PIs), analytical procedures, calibration factors, uncertainties, storage location, etc. These records data will be screened for anomalies and the analytical error will be reported. Where possible, reanalyses will be completed to double check anomalous values. Outliers included in the final, submitted data set will be flagged to alert subsequent data users. The data will primarily consist of laboratory notebooks (from chemical analysis in the laboratory), which will be entered into electronic format and stored on removable memory storage devices.

2. What is the expected size of the collection (number and types of files as well as data volume)?

<150 GB; number of files is unknown at this time.

3. Do standard vocabularies, keywords, or other conventions exist in your discipline for describing data? If so, which ones will you be using?

No special vocabulary will be used; all terms will be defined with units of measure.

Data organization

1. What will be the general data management approach during collection and analysis?

We plan to create a website for this project that will contain the following sections:

a) *Publicly available data* containing information about each of the participants with links to their personal web pages and full contact information as well as a description of their role in the project. Links for downloading published data and publications will also be provided. All project data will be made available to the research community through the publication of several manuscripts in widely available scientific journals.

b) *Resources for the Participants,* which will be password protected, containing monthly updated laboratory data, will be made available to the project's collaborators as it becomes available and to the research community within two-three years after collection through the publication of several manuscripts in widely available scientific journals. As appropriate, we will make all our data available as supplemental material on the journal websites. PI's affirm that data generated, collected, and modeled as part of the proposed research will be preserved for at least 5 yr after the award ends. There are not expected to be any privacy, ethical, intellectual property, or patent issues with the data.

2. How will you keep track of changes in data and different versions, especially season to season or year to year. Will you make use of unique identifiers? How?
All file names include DDMMYYYY and are renumbered as new versions are saved.

Data quality
1. What data quality assurance and control processes will be implemented?
All experiments will be run in at least triplicate (N≥3).

2. How will you ensure the integrity of the data?
All notebooks will be scanned and stored electronically, as well as other information relevant to the collection, processing, and analyses of the samples. Strict lab policy in all participants' labs governs and secures general protocols for data use, access, share, formats, security, backups and long-term archival, taking into consideration research needs, legal obligations, and ethical responsibilities.

3. Please identify a designated person responsible for the quality of the data:
Kathleen A. Schwehr, schwehrk@tamug.edu; Chen Xu, xuchen66@hotmail.com; Antonietta Quigg, quigga@tamug.edu; Peter Santschi, santschi@tamug.edu