Data Management Plan

1. Roles and responsibilities

The PIs at University of Florida (UF – Yoon, Casanova) and Texas Tech University (TTU - Li) will be responsible for data management. Project staff (host lab personnel) will be instructed on data archival policy. The PI's will make decisions on these data, which include destroying them when no longer needed.

2. Products of the research

The proposed research will be carried out at UF and TTU. Assessment will be carried out by UF and TTU personnel. A record of all experiments, procedures, notes, and observations performed over the course of the proposed program will be preserved in notebooks or databases that are maintained by each researcher and supervised by senior investigators. Other data that will be generated include analytical data, images from various microscopies, videos, and physical samples of materials.

3. Types of data, data formats, and metadata

Data acquired from instrumentation will be stored in both electronic and hard copy formats. Depending on the output of the instrument, the electronic outputs may be a graph displayed as an image or as a table in ASCII or other suitable format developed by the manufacturer or representative images. Numerical data obtained from instrumentation may also be imported in spreadsheets to aide in statistical analysis. Original unaltered microscopy images will be preserved on the original instruments, and original images will be copied into the notebooks of individual researchers. The format of this information varies considerably depending on the program used. Computational data will be acquired with standard package programs in standard formats. Data in locally-developed software is also stored in well-known formats, typically in plain text, including off, mol2, xyz, and tables with column headers. Lossless compression of the data files may be performed on large data sets. There is a dedicated backup server available for each user and located within a secure facility.

Files will be labeled with descriptive titles and dates, and additional text files will be stored containing keys to these data. Data from individual participants will be associated with a code and made anonymous in the stored form, wherever possible. A spreadsheet will relate the codes to demographic information, and another will provide the key that unlocks the name to code link. This measure enables the investigators to mine data collected from the same participant at a later time.

4. Research data sharing policies and intellectual property

Specific data generated from the proposed research accumulated over the course of research will be

shared via publication in peer-reviewed journals. It is not planned to release data prior to publication. An exception is the sharing of data within the context of a meeting with collaborators, reports to the NSF, or a scientific conference. For data sharing, whether or not access is granted to any data generated depends on the nature of the third party. In general, groups within academia will be allowed access to unpublished data. Additionally, data resulting from the proposed research will be shared between the partnering Institutions in accordance to the collaboration outlined in the proposal. Any sensitive data will be regarded as such and will be treated with appropriate diligence. Data that are shared will include standards and notations needed to interpret the data, following commonly accepted practices in the field. Data will be available for access and sharing as soon as is reasonably possible, normally no longer than two years after its acquisition. In the event that discoveries or inventions are made in direct connection with this data, access will be granted upon request once appropriate invention disclosures and/or provisional patent filings are made. Key data relevant to the discovery will be preserved until all issues of intellectual property are resolved.

5. Policies and provisions for re-use, re-distribution, and the production of derivatives

Any data generated from the proposed research is not planned to be released publically prior to publication. There will be no need for disclaimers or conditions regarding use of the data.

6. Data retention

Data will be retained for 10 years after the end of the funding period. This is to provide access for further data mining.

7. Assessment data dissemination

Assessment data will be available outside the team three years after completion of the project. Requests may be made by e-mail to one of the team members. Any data sharing outside the team will only be provided in an anonymized form.

8. Data achieving, storage and access

Hand-written notebooks will be stored away from the laboratory when not in use. Printout copy of instrumentation data will also be stored in an office. Electronic data, whether generated from instrumentation or computationally, will be stored on a computer and will be backed up on the server at the partner institution, in accordance with institutional policies, to prevent accidental data loss. All generated data will be archived according to date of collection and sample name, for easy identification of the electronic data associated with the described experiments.

Assessment data and metadata will be stored on the “Mimmetic Dropbox” and backed up on “Research Data Storage Service,” a secure and stable collaborative research data storage platform. Research Data Storage Service is supported by UF IT and offers the following features and benefits:

- 1 TB storage space per account at no cost
- Compliant for storage for most legally or contractually restricted data.
- Daily snapshots and data replication for disaster recovery saved for 28 days.
- Auditing capabilities for reporting on access, creation, replication, updates, and deletion of files upon request.
- Ability to share data with internal and external collaborators.
- Accessible via VPN when off campus.*
- Accepts common network drive mapping for Windows, Mac OS and Linux - SMB/CIFS access protocols.
- Ability to manage user read-only and read/write access through established security groups using Outlook Web Access (OWA).