**EXAMPLE DATA MANAGEMENT PLAN: CHEMISTRY**

1. **Research Products.**

Data collected through this proposed study will include synthesis of chemical compounds, data pertaining to the ability of the iodonium salts to activate thioglycosides, the compatibility of this technology with additives to control reaction stereochemistry, and the use of these promoters in various platforms which can be transitioned into automated synthesis. Products will be purified, and fully characterized to establish chemical structure/purity using spectroscopic methods. Typical spectroscopic data will include 1H NMR, 13C NMR, IR, specific rotation and mass spectroscopy. For selectivity studies, 1H NMR data on diastereomeric mixtures of reaction products prior to separation will be used to assess the selectivity of all reactions described in this proposal.

1. **Data Format**

Data will be collected according to methods described in the proposal (see section 6 project description) using common software appropriate to the instrument. This data will be stored in electronic format (both raw FID files and PDF format) and in the laboratory as hardcopies of notebooks and spectra. All spectra are indexed according to the page in the laboratory notebook in which they are recorded for rapid identification. Raw FID files collected on Bruker Topspin software can be processed by any number of NMR software programs, and can be made available upon request. PDF files will contain both images of the original spectra, as well as numerical tables detailing chemical shift, coupling constants, mass values, and other numerical interpretations of data.

**3. Metadata**

Metadata will be provided.  The project will document information about the context, content, quality, provenance, and/or accessibility of the data used. This will also include information embedded in the raw FID files. Additionally, the project will seek to document information about authors, dates and brief descriptions for scanned PDFs, notebooks and lab work.

**4. Access to Data and Data Sharing Policies**

This data will be of interest to the synthetic organic and carbohydrate chemistry communities, and will be made available upon request. When requested, the data will be made available by the PI so long as the request does not interfere with publication, compromise intellectual property interests, or precede data analysis. Shared data will include both spectroscopic data, and reagents if not commercially available. Data will be made available for access as soon as reasonably possible.

Data will be disseminated by publication in scholarly journals. Data in both journal articles and supporting information will be presented in standard format. This information will include descriptions of experimental procedures, numerical descriptions of spectra and electronic copies of the actual spectra.

Tufts University maintains a distributed information technology environment, with central as well as local aspects of overall planning and control. Tufts' information security program is structured in a similar manner. Operationally, Tufts central IT organization, Tufts Technology Services and each local IT group maintain standards of quality and professionalism regarding operational processes and procedures that enable effective operational security. For TTS managed systems, the emphasis is on centralized resources such as administration and finance, telecommunications, research computing and networking, systems and operations as well as directory, email, LDAP, calendaring, storage and Windows domain services. TTS also provides data center services and backups for all of these systems. Additionally, a large number of management systems (for patching), anti-virus and firewall services are centrally provided and/or managed by TTS. Within TTS, processes and procedures exist for managed infrastructure changes, as change control is required for all critical central systems. Tufts University provides anti-virus software for computers owned by the University, and makes anti-virus software available at no charge for users who employ personally owned computers in the course of their duties at the University.

All devices and users are subject to the Tufts Information Stewardship Policy located here: http://it.tufts.edu/ispol

1. **Re-Use, Re-Distribution and Production of Derivatives & Responsible Conduct of Research**

All laboratory members with access to data will receive instruction in the Responsible Conduct of Research (RCR), provided by the Office of the Vice Provost at Tufts. Instruction is provided 2-3 times a year in six 2 hour sessions.

Original notebooks and hardcopies of all NMR and mass spectra are stored in the PI’s laboratory Additional electronic data will be stored on the PI’s computer, which is backed up daily. Additionally, the laboratory will make use of the PI’s lab server space at Tufts for a second repository of data storage. The PI’s lab has access to up to 1 terabyte of information storage at Tufts, which can be expanded if needed.

The PI will retain principal legal rights to the intellectual property developed under this grant, as is compliant with the Tufts University policies, which can be found at: http://techtransfer.tufts.edu/resources/tufts-policies/intellectual-property-policies/

All the project data will be stored using Tufts University Information Technology resources. Research data for this study will be stored on Tufts University's centrally provided “shared research storage facility”, which is provisioned on Tufts’ NetApp enterprise networked storage. Provisioning occurs through self-service at <http://researchstorage.uit.tufts.edu>. Data security and confidentiality are protected by using Microsoft Active Directory authentication, and the storage is backed up to LTO-4 tape on a daily and weekly basis and stored offsite at Iron Mountain facilities. As storage needs for the research study increase over time, additional storage can be easily provisioned on the same appliance. TTS issues storage space in 50GB increments, up to several TB.

1. **Archiving Data**

Data will be stored for a minimum of three years beyond award period, per NSF guidelines. If inventions or new technologies are made in connection data, access to data will be restricted until invention disclosures and/or provisional patent filings are made with the Tufts Office of Technology Licensing and Industry Collaboration.

The data acquired and preserved, as part of the proposed research will be governed by Tufts University’s policies regarding intellectual property, record retention, and data management.

The PI will work with Tisch Library at Tufts University to identify data repositories, such as PubChem, to determine the best way to archive and share data that is not published as part of supporting information. This includes unpublished data that is part of Ph.D. theses. Such data will be made available upon request as described above.

The data to be acquired through the proposed research does not involve human or animal subjects.

1. **Evaluation Data from Broader Impacts**

All information collected for evaluation of the Broader Impacts section of this proposal will be intended solely for making annual improvements to the program to increase its effectiveness. Accordingly, this data will not be made publically available in any format. After each evaluation, the data will be summarized and the original responses to surveys will be destroyed. The PI will maintain a copy of these summaries for his personal records during the duration of the outreach program. If for any reason these programs are terminated all information regarding the surveys will be destroyed.