Data Management and Sharing Plan

Element 1: Data Type:

A. Types and amount of scientific data expected to be generated in the project:

Summarize the types and estimated amount of scientific data expected to be generated in the project.

This project is estimated to generate a significant amount of scientific data across several categories:
- Design, production, characterization, and purification protocols for 30 Compound A analogs.
- In vitro and cell-based assay results for these 30 compounds.
- Substance stability data, as well as in vitro and in vivo toxicology results for 10 selected compounds, involving 6 wild-type (WT) mice per dose across 5 doses.
- Pharmacokinetics (ADME) data for 3-4 compounds tested in 10 5xFAD mice per dose across 4 doses.
- Preclinical positron emission tomography (PET) imaging data for 2-3 compounds, using 15 WT or 5xFAD mice per compound.

B. Scientific data that will be preserved and shared, and the rationale for doing so:

Describe which scientific data from the project will be preserved and shared and provide the rationale for this decision.

All scientific data generated from this project, including protocols, assay results, toxicology findings, pharmacokinetics data, and PET imaging data, will be preserved and shared. The rationale is to contribute meaningfully to the research community by providing comprehensive insights into Compound A analogs' potential therapeutic applications, especially in aging-related diseases. Sharing these data will facilitate collaboration, accelerate discovery, and potentially lead to breakthroughs in understanding and treating age-associated conditions.

C. Metadata, other relevant data, and associated documentation:

Briefly list the metadata, other relevant data, and any associated documentation (e.g., study protocols and data collection instruments) that will be made accessible to facilitate interpretation of the scientific data.

To facilitate interpretation of the scientific data, the following metadata and documentation will be made accessible:

- Detailed protocols for compound synthesis and characterization.

- Study designs and methodologies used for in vitro, cell-based assays, toxicology studies, pharmacokinetics analyses, and PET imaging experiments.

- Data collection instruments and software used.

- Statistical analysis plans and software codes.

Element 2: Related Tools, Software and/or Code:

State whether specialized tools, software, and/or code are needed to access or manipulate shared scientific data, and if so, provide the name(s) of the needed tool(s) and software and specify how they can be accessed.

Specialized tools, software, and code are needed to access or manipulate shared scientific data. Specifically:

- Computational chemistry software for analyzing compound structures and predicting properties.

- Statistical software (e.g., R, Python libraries) for data analysis and visualization.

- Imaging analysis software for PET imaging data processing.

These tools will be specified in the project's methods section and made available through open-source repositories or cited publications.

Element 3: Standards:

State what common data standards will be applied to the scientific data and associated metadata to enable interoperability of datasets and resources and provide the name(s) of the data standards that will be applied and describe how these data standards will be applied to the scientific data generated by the research proposed in this project. If applicable, indicate that no consensus standards exist.

To enable interoperability of datasets and resources, common data standards will be applied to the scientific data and associated metadata. Specifically:

- The Minimum Information About a Bioactive Entity (MIABE) for chemical structures and bioactivity data.

- The Investigation/Study/Assay (ISA) framework for describing experimental designs and workflows.

- Standardized metadata formats for imaging data, such as those recommended by the National Institute of Mental Health (NIMH) Data Archive.

These standards will be applied during data collection, processing, and sharing to ensure consistency and facilitate integration with other datasets.

Element 4: Data Preservation, Access, and Associated Timelines:

A. Repository where scientific data and metadata will be archived:

Provide the name of the repository(ies) where scientific data and metadata arising from the project will be archived.

The scientific data and metadata arising from the project will be archived in publicly accessible repositories such as the National Center for Biotechnology Information (NCBI) PubChem for chemical compounds, the NIH-supported data repository for imaging data, and other discipline-specific databases recommended by the NIA.

B. How scientific data will be findable and identifiable:

Describe how the scientific data will be findable and identifiable, i.e., via a persistent unique identifier or other standard indexing tools.

The scientific data will be made findable and identifiable through the use of persistent unique identifiers (e.g., DOIs) for datasets, publications, and other research outputs. Standard indexing tools such as PubMed for literature and database-specific search functions will also facilitate discovery.

C. When and how long the scientific data will be made available:

Describe when the scientific data will be made available to other users (i.e., no later than the time of an associated publication or end of the performance period, whichever comes first) and for how long data will be available.

The scientific data will be made available to other users no later than the time of publication of the main findings or at the end of the performance period, whichever comes first. The data will remain accessible

for a minimum of 10 years after the project's completion to allow for long-term research and validation studies.

Element 5: Access, Distribution, or Reuse Considerations:

A. Factors affecting subsequent access, distribution, or reuse of scientific data:

NIH expects that in drafting Plans, researchers maximize the appropriate sharing of scientific data. Describe and justify any applicable factors or data use limitations affecting subsequent access, distribution, or reuse of scientific data related to informed consent, privacy and confidentiality protections, and any other considerations that may limit the extent of data sharing.

While maximizing data sharing, considerations for limiting access include protecting sensitive information related to compound synthesis that could have commercial applications. However, all shared data will be anonymized and de-identified where applicable to protect privacy and confidentiality.

B. Whether access to scientific data will be controlled:

State whether access to the scientific data will be controlled (i.e., made available by a data repository only after approval).

Access to certain datasets, particularly those involving proprietary compound information or pre-publication research findings, may be controlled through a data use agreement (DUA) to ensure appropriate use and protection of intellectual property rights.

C. Protections for privacy, rights, and confidentiality of human research participants:

If generating scientific data derived from humans, describe how the privacy, rights, and confidentiality of human research participants will be protected (e.g., through de-identification, Certificates of Confidentiality, and other protective measures).

Since this project primarily involves animal models and in vitro studies, protections for human subjects are not directly applicable. However, any future extension involving human data will strictly adhere to HIPAA guidelines, IRB approvals, and informed consent protocols to protect participant privacy and rights.

Element 6: Oversight of Data Management and Sharing:

Describe how compliance with this Plan will be monitored and managed, frequency of oversight, and by whom at your institution (e.g., titles, roles).

The implementation of these plans will be overseen by the project's principal investigator in collaboration with research team members. Regular meetings and progress updates will ensure that data management, sharing, and preservation activities are carried out as outlined.