Investigating the Cell Autonomous and Non-cell Autonomous Factors Driving Neuronal TAU Pathology and Degeneration in Alzheimer's Disease

A Data Management Plan created using DMPonline.be

Creator: Nóra Baligács

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: n.n. n.n.

Grant number / **URL:** 11PSA24N

ID: 204456

Start date: 01-11-2023

End date: 31-10-2027

Project abstract:

Alzheimer’s disease (AD) is the most common neurodegenerative disorder, characterized by deposition of amyloid-β (Aβ) plaques, neuronal TAU tangles, neuroinflammation, and neuron loss. Mutations that predispose to develop Aβ plaques are sufficient to trigger downstream TAU pathology and neuronal degeneration. Deposition of Aβ induces neuroinflammatory responses from glia, including microglia. However, the mechanism by which TAU pathology and neurodegeneration are triggered in neurons remains unclear.

In this study, I aim to investigate whether human neurons cultured in vitro develop TAU and necroptosis pathology autonomously when exposed to long-term amyloid plaque exposure. Advancements in stem cell derived human neurons and modelling amyloid plaque-like deposits in vitro allow me to answer this question.

I will also explore whether the initiation of TAU and necroptosis pathology is not autonomous in a xenograft model of AD. Human xenografted neurons in the brains of mice with Aβ plaques develop TAU pathology and neurodegeneration. This is accompanied by gliosis mediated by the host. By depleting microglia from this model, I will test whether microglia contribute to TAU pathology.

Following identification of whether TAU induction is autonomous, a screening to identify candidate molecules that link the pathologies in AD will be conducted. The identified molecules will be validated experimentally in the appropriate model and investigated for their expression in human samples.

Last modified: 02-04-2024

Investigating the Cell Autonomous and Non-cell Autonomous Factors Driving Neuronal TAU Pathology and Degeneration in Alzheimer's Disease
FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Tissue samples	Biological and chemical samples: Fixed and fresh frozen brain samples from mice, human iPSC and NPC in cryovails in liquid nitrogen tank	Generate new data	Physical	Experimental	NA	NA	200-500 samples
RNA/Protein samples	Extracted RNA/protein from collected tissues and biological samples	Generate new data	Physical	Experimental	NA	NA	300-600 tubes
Microscopy slides	Stained tissues and cells mounted on microscopy slides	Generate new data	Physical	Experimental	NA	NA	500-800 microscopy slides
DNA plasmids	tubes containing DNA plasmids	Generate new data	Physical	Experimental	NA	NA	3-15 tubes

Microscopy images	Confocal Microscope, Electron Microscope, and Axioscans from mouse brains and in vitro cells	Generate new data	Digital	Experimental	.tif .ZEN .nd2	<1TB	
ELISA/MSD data	Reads from the MSD reader from mouse brain extracts or in vitro cell lysates, and analysis in Microsoft excel and graph pad prism	Generate new data	Digital	Experimental	.txt .xlsx .prism	<100MB	
Western blot images, PCR images and analysis	Numerical data. Quantification of western blot images, performed using imageJ, in microsoft excel and graph pad prism	Generate new data	Digital	Experimental	tif .xlsx .prism	<100MB	
qPCR data	qRT-PCR data files and statistical analysis	Generate new data	Digital	Experimental	.txt .xlsx .prism	<100MB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

NA

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, animal data

All the experiments related to the mouse are approved by the KU Leuven ethical committee 117/2022 and 032/2024.

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- Yes

We might possibly identify new targets to treat Alzheimer's disease.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

1. Physical data: Sample source, sample type, concentration, storage location, date of collection/generation, methodology, and protocol will be described in detail in the Electronic lab notebook.
2. Digital data: the following information will be noted when applicable:
   dimensions, image type, bit-depth, pixel sizes, and microscope settings. The methodology and protocol will be described in detail in the Electronic lab notebook. The file name will contain all the relevant information about the samples, e.g. sample number and antibodies used.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

Type of Data	Metadata standard
Confocal images of mouse brain tissue	OME-TIFF
PCR gel images	OME-TIFF
Western blot images	OME-TIFF

Where no metadatastandard exists, metadata will be stored based on the Dublin core standard. The following information will be stored:
This document was generated by DMPonline (http://dmponline.dcc.ac.uk) 4 of 8
Title: free text
Creator: Last name, first name, organization

Date and time reference
Subject: Choice of keywords and classifications
Description: Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data, the software(s) (including version number) used to produce and to read the data, the purpose of the experiment, etc.
Format: Details of the file format,
Resource Type: data set, image, audio, etc.
Identifier: DOI (when applicable)
Access rights: closed access, embargoed access, restricted access, open access.

The final dataset will be accompanied by this information under the form of a README.txt document. This file will be located in the top level directory of the dataset and will also list the contents of the other files and outline the file-naming convention used. This will allow the data to be understood by other members of the laboratory and add contextual value to the dataset for future reuse.

3. Data storage & back-up during the research project

Where will the data be stored?

All digital data: will be stored on KU Leuven servers (L-drive), with regular on-site backup and mirroring, or stored on a cloud-based service offered by KU Leuven (OneDrive).
All samples will be stored as appropriate: -80°C for nucleic acids, and protein samples,
will be stored with appropriate backup copies of the stock in the Prof. De Stroopers lab at KU Leuven. Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both under the form of purified DNA (in -20°C freezer). One copy will be submitted to the lab plasmid database system.
All published vectors and the associated sequences will be sent to the non-profit plasmid repository Addgene, which will take care of vector storage and shipping upon request. Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.

How will the data be backed up?

KU Leuven drives are backed up according to the following scheme:
- data stored on the “L-drive” is backed up daily using snapshot technology, where all incremental changes in respect of the previous version are kept online; the last 14 backups are kept.
- data stored on the “J-drive” is backed up hourly, daily (every day at midnight) and weekly (at midnight between Saturday and Sunday); in each case the last 6 backups are kept.

Data on OneDrive is synced approximately every 10 minutes.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

There is sufficient storage and back-up capacity on all KU Leuven servers:
- the “L-drive” is an easily scalable system, built from General Parallel File System (GPFS) cluster with NetApp eseries storage systems, and a CTDB samba cluster in the front-end.
- the “J-drive” is based on a cluster of NetApp FAS8040 controlers with an Ontap 9.1P9 operating system.

Onedrive provides 1TB of storage data which should be sufficient for the data generated in the project.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Both the “L-drive” and “J-drive” servers are accessible only by laboratory members and are mirrored in the second ICTS datacenter for business continuity and disaster recovery so that a copy of the data can be recovered within an hour.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

The costs of digital data storage are 173.78€/TB/Year for the “L-drive”. The total estimated cost of data storage of 1-2TB data during the 4-year project is maximally (174€/TB/Year * 2TB * 4years =) 1392€. These costs will be met partially by the benchfee provided by the FWO and partially by existing lab grants.
Maintaining a mouse colony alive costs about 1,200 euro per year (for 6 cages), excluding the costs of genotyping. When no experiment is planned with a particular mouse strain, and in compliance with the 3R’s rule (https://www.nc3rs.org.uk), cryopreservation will thus be used to safeguard the strain, prevent genetic drift, loss of transgene, and potential infections or breeding problems. Cryopreservation of sperm/embryos costs about 500 to 700 euro per genotype, plus a minimal annual storage fee (25 euros per strain for 250 to 500 embryos). Frozen specimen are kept in two separate liquid nitrogen tanks at two different sites on campus. When necessary, the costs of revitalization from cryopreserved sperm/embryos are about 1,100/600 euro.
Electricity costs for the -80° freezers and physical data storage are included in the general lab costs.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

The minimum preservation term of 5 years after the end of the project will be applied to all datasets.

Where will these data be archived (stored and curated for the long-term)?

As a general rule, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org), at the latest at the time of publication.
For all other datasets, long term storage will be ensured as follows:

- Digital datasets: files will be stored on the “L-drive”.
- Tissue samples: Tissues will be stored locally in the laboratory.
- Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both in the form of purified DNA (in -20°C freezer) and as a bacteria glycerol stock (-80°C).
- Cell lines: cell lines will be stored locally in the laboratory (-80°C).
- Genetically modified organisms: lines that are not actively used for experiments will be cryopreserved.
- Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

The total estimated cost of data storage during the 5 years after the end of the project is about 4500€ (for 5 TB). This estimation is based on the different costs described before.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

We are committed to publishing research results to communicate them to peers and a wide audience. All research outputs supporting publications will be made openly accessible. Depending on their nature, some data may be made available prior to publication, either on an individual basis to interested researchers and/or potential new collaborators, or publicly via repositories

(e.g. negative data).
Biological material will be distributed to other parties if requested.

If access is restricted, please specify who will be able to access the data and under what conditions.

Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. As detailed above, metadata will contain sufficient information to support data interpretation and reuse and will conform to community norms. These repositories clearly describe their conditions of use (typically under a Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication or an ODC Public Domain Dedication and Licence, with a material transfer agreement when applicable). Interested parties will thereby be allowed to access data directly, and they will give credit to the authors for the data used by citing the corresponding DOI. For data shared directly by the PI, a material transfer agreement (and a non-disclosure agreement if applicable) will be concluded with the beneficiaries in order to clearly describe the types of reuse that are permitted.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

As a general rule, datasets will be made openly accessible via existing platforms that support FAIR data sharing (www.fairsharing.org). Sharing policies for specific research outputs are detailed below:

- Vectors: Upon publication, all vectors supporting a manuscript will be made publicly available via the non-profit plasmid repository Addgene, along with the corresponding DNA sequences. Addgene in turn performs quality control on the DNA, curates the plasmids online with all relevant information (maps, sequences), and for a minimal cost ships the vectors upon simple request and signature of a material transfer agreement.
  The MTA will be prepared before depositing the vectors with the help of our organization's Tech Transfer office. For transfer between nonprofit or academic institutions, Addgene typically uses the Uniform Biological Material Transfer Agreement (https://www.addgene.org/terms/1047/). All non-published vectors and the associated documentation will be shared by the PI upon request and after signature of a material transfer agreement, at no cost except the cost of shipment.
- Cell lines: All human pluripotent cell lines supporting publications will be registered in hPSCreg, the European human embryonic stem cell registry supported by the European Commission (https://hpscreg.eu/). Information about the deposited lines (including donor information, derivation method, availability, and characterization) will also be made accessible. Registration of cell lines in hPSCreg will provide visibility, confirm ethical procurement, and facilitate comparison with other hPSC lines. The PI will remain the distributor of the pluripotent cell lines.

All other cell lines supporting publications will be deposited in the American Type Culture Collection (ATCC) database (https://www.atcc.org/), which is a private, non-profit biological resource center. This will provide a secure backup for this material. Investigators can purchase cell lines from the ATCC database upon signature of a material transfer agreement (https://www.lgcstandards-atcc.org/~/media/PDFs/MTA_2.ashx) and, in some cases, of a Limited Use/Label License (e.g. for CRISPR products or iPSC materials) and/or a Customer Acceptance of Responsibility (for potentially highly pathogenic materials). Information about the cell lines (including organism, cell type, tissue, biosafety level, and disease if applicable) will also be made accessible.

- Genetically modified organisms: All genetically modified organisms used in publications will be made available to researchers upon request at the time of publication.
- Antibodies, synthetic and recombinant compounds: samples will be stored as appropriate in the laboratory. Within availability, they will be shared with interested researchers upon request.
- Research documentation: All protocols used to generate published data will be described in the corresponding manuscript(s), and the related documentation will be included as supplementary information. These data and all other documents (daily logs, raw data) deposited in the E-Notebook are accessible to the PI and the research staff and will be made available upon request.
- Manuscripts: All scientific publications will be shared openly. Manuscripts submitted for publication will be deposited in a pre-print server such as bioRxiv, arXiv, Nature Precedings, or ASAPbio). At the time of publication, research results will be summarized on the PI's website (add website address) and post-print pdf versions of publications will be made available there if allowed by copyright agreements, possibly after an embargo as determined by the publisher. Before the end of the embargo or in cases where sharing the post-print is not allowed due to copyright agreements, a pre-print version of the

manuscript will be made available. Publications will also be automatically added to our institutional repository, Lirias 2.0, based on the author's name and ORCID ID.
- Nucleic acid and protein sequences: All nucleic acid and protein sequences generated during the project will be stored on KU Leuven servers. Upon publication, all sequences supporting a manuscript will be made publicly available via repositories such as the GenBank database or the European Nucleotide Archive (nucleotide sequences from primers / new genes / new genomes), NCBI Gene Expression Omnibus (microarray data / RNA-seq data / CHIPseq data), the Protein Database (for protein sequences).
- Data that do not support publication will be either deposited in an open-access repository or made available upon request by email.

When will the data be made available?

The data will be made available as a pre-print when submitting the manuscript and as a post-print upon publication. No embargo will be foreseen unless imposed e.g. by pending publications, or potential IP requirements. In those cases, datasets will be made publicly available as soon as the embargo date is reached.

Which data usage licenses are you going to provide? If none, please explain why.

The provided data usage license will depend on the journal/repository in which the project will be published. This is the standard creative commons license is CC-BY.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

A permanent identifier is added to the data upon deposit in a repository.

What are the expected costs for data sharing? How will these costs be covered?

It is the intention to minimize data management costs by implementing standard procedures e.g. for metadata collection and file storage and organization from the start of the project, and by using free-to-use data repositories and dissemination facilities whenever possible. Data management costs will be covered by the laboratory budget.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the electronic laboratory notebook (E-notebook) that refer to specific datasets.

Who will manage data storage and backup during the research project?

The research and technical staff will ensure data storage and back up, with support from René Custers and Alexander Botzki for the electronic laboratory notebook (ELN) and from Raf De Coster for the KU Leuven drives.

Who will manage data preservation and sharing?

The PI is responsible for data preservation and sharing, with support from the research and technical staff involved in the project, from René Custers and Alexander Botzki for the electronic laboratory notebook (ELN) and from Raf De Coster for the KU Leuven drives.

Who will update and implement this DMP?

The PI is ultimately responsible for all data management during and after data collection, including implementing and updating the DMP.