Tau-dependent synaptic remodeling in health and disease

A Data Management Plan created using DMPonline.be

Creators: Patrick Vandormael, Patrik Verstreken https://orcid.org/0000-0002-5073-5393

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Patrik Verstreken https://orcid.org/0000-0002-5073-5393

Project Administrator: Patrick Vandormael

Grant number / URL: G045724N

ID: 208301

Start date: 01-01-2024

End date: 31-12-2027

Project abstract:

Tau is a neuronal microtubule-associated protein under normal conditions. However, stressful triggers like toxic protein aggregates or traumatic brain injury can cause its detachment and invasion of synaptic terminals, leading to synapse loss. Interestingly, cold stress-induced torpor in hamsters also results in Tau detachment and massive synapse loss, indicating that Tau may be responsible for executing stress signals that cause synapses to retract. However, unlike in neurodegeneration, hibernating hamsters do not convert into overt neurodegeneration and they manage to quickly regenerate synapses upon emergence from hibernation. This proposal aims to identify the molecular determinants of this remarkable form of synaptic plasticity at single-cell resolution and identify factors that cause hibernation-induced synapse remodeling in hamster brains using new technology. We will then test the hypothesis that synaptic regeneration pathways that hamsters use, are able to counteract synapse loss in models of Tau-induced disease. Preliminary data already identified several synaptic proteins which are deregulated in hibernation and implicated in neurodegenerative disease, including Rab3a that we will study here. This research will not only uncover the role of Tau- intersecting pathways in a remarkable form of synaptic plasticity, but also how this can be used to counteract synapse loss in Tau-induced disease.

Last modified: 26-06-2024

Tau-dependent synaptic remodeling in health and 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:* •	*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	
Objective1 Task 2.1: snRNAseq	snRNAseq of hamster hippocampi during hibernation	new	digital	Experimental measurement	- Raw: binary base call format (.bcl) - textual data: FASTQ file (.fastq, zipped as .gz) - Sequence alignment data: .bam, - analysed data: tabular data (.xlsx, .tsv, .csv) - metadata: textual data (.rtf, .xml, .txt)		

Objective1 Task 2.1: mouse cell atlases	sn/scRNAseq of mouse brain cell types	reused (external)	digital	Experimental measurement	- textual data: FASTQ file (.fastq, zipped as .gz) - Sequence alignment data: (.bam) - analysed data: tabular data (.xlsx, .tsv, .csv) - metadata: textual data (.rtf, .xml, .txt)		
Objective1 Task 2.2: STOmics	spatial transcriptomics of hamster brains during hibernation	new	digital	Experimental measurement			
Objective1 Task 3: mass spectrometry P301S Tau mice	mass spec of hippocampal synaptosome of P301S Tau mice	reused (own)	digital	Experimental measurement			
Objective1 Task 3: mass spectrometry AD patients	mass spec of synapse-enriched fractions from pre-symptomatic vs symptomatic AD patients	reused (external)	digital	Experimental measurement			
Objective1 Task 4.1: RNAscope	mRNA location and abundance	new	digital	Experimental measurement	- raw data: Carl Zeiss Image files (.czi, .lsm), Nikon Image file (.nd2) - Digital images in raster formats: uncompressed TIFF (.tif/.tiff) - analysed data: tabular data (.xlsx, .tsv, .csv) -metadata: textual data (.rtf, .docx, .txt)		

Objective1 Task 4.2 and Objective 3 Task 1.1: fluorescence microscopy	immunofluorescence stainings of hamster hits, and hamster Tau, synaptoporin, and nectin-3	new	digital	Experimental measurement	- raw data: Carl Zeiss Image files (.czi, .lsm), Nikon Image file (.nd2) - Digital images in raster formats: uncompressed TIFF (.tif/.tiff) -metadata: textual data (.rtf, .xml, .txt)		
Objective2 Task 1: electroretinograms (ERGs)	electrophysiological assay to assess synaptic communication in the adult fly brain	new	digital	Experimental measurement	- Applied Biosystems Sequence Tracer Sequence Trace (.abf)		
Objective2 Task 1: fly lines	LOF fly lines, crossed with human Tau P301L-expressing flies	existing and new	physical	NA	NA		2 vials per line
Objective2 Task2: iPSC lines	Kolf2.1iPSCs with knock-in of TauP301L and dSpCas9-KRAB	existing and new	physical	NA	NA		Biological and chemical samples: samples stored at liquid nitrogen
Objective2 Task2: gRNA vectors	lentiguide-Puro vectors with sgRNA insertion	new	physical	NA	NA		Biological and chemical samples: samples stored at -20°C

Objective2 Task3.1: vesicle mobility	FRAP of synaptotagmin-GFP in Tau P301L-expressing flies, crossed with mediators of cluster formation	new and existing	digital	Experimental measurement	- raw data: Carl Zeiss Image files (.czi, .lsm), Nikon Image file (.nd2) - Digital images in raster formats: uncompressed TIFF (.tif/.tiff) - Digital video data: MPEG-4 High Profile (.mp4); Audio Video Interleave (.avi) - analysed data: tabular data (.xlsx, .tsv, .csv) -metadata: textual data (.rtf, .xml, .txt)		
Objective 2 Task 3.3 and Objective 3 Task 2.2: TEM	transmission electron microscopy of Drosophila synapses	new	digital	experimental measurement			

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:

Own datasets/tools

WP/Task	dataset/tool	source
Obj1 Task 3	mass spectrometry P301S Tau mice	unpublished, generated in-house

external datasets/tools

WP/Task	dataset/tool	source
Obj1 Task 2.1	sn/scRNAseq of mouse brain cell types	publicly available datasets linked to publications (Zeisel et al: https://doi.org/10.1016/j.cell.2018.06.021; Cembrowski et al: https://doi.org/10.7554/eLife.14997)
Obj1 Task 3	mass spectrometry AD patients	publicly available datasets linked to publications (Li, X. et al: https://doi.org/10.1002/alz.12345)

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, human subject data
- Yes, animal data

Our study involves the creation and re-use of data derived from both human bodily materials (HBM) and laboratory animal experiments. We are committed to ensuring all ethical considerations are addressed.

Human Bodily Materials and Data

This study will characterize 10 hibernation hits in a human in vitro neurodegeneration model using the Kolf2.1iPSC line with KI of TauP301L and dSpCas9-KRAB. Our lab has approval from the Ethics Committee Research UZ/KU Leuven for the use of the parental Kolf2.1iPSC ine (study S65730), and will seek approval via amendment for the use of the KI line before initiation of the described experiments.

Laboratory Animal Experiments

We will be conducting experiments on hamsters and will seek approval from the Ethical Committee for Animal Experimentation before their initiation.

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 do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. Ownership of the data generated belongs to KU Leuven and VIB in accordance with the framework agreement of both institutes. VIB has a policy to actively monitor research data for such potential. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that publications need not be delayed.

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

No third-party agreement restricts dissemination or exploitation of the data or strains generated from this project. Existing agreements between VIB and KU Leuven do not restrict publication of data.

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).

The organisation of research data follows a hierarchical **folder structure** based on model organism (eg. flies/mice/hamsters/cells). At the second level, these folders contain subfolders per experiment type (eg. western blotting, ERG, EM, immunofluorescence,...). At the third level, these folders contain subfolders per individual experiment (with date indication, see below). Each of these experiment folders contain subfolders for raw data, processed data, and analysed data.

Digital files and folders of individual experiments will be **named** following a standard naming convention. All file names in a given dataset will follow the same format: they will start with the date, followed by the project acronym, a concise descriptive name, and an optional version number. Whenever possible, names will be kept under 32 characters, containing only letters, numbers, and underscores. Dots will only be used

for version control indicators (minor revisions indicated by decimal numbers, and major revisions by whole numbers): YYYYMMDD_Project_Experiment_ version.format. Similarly, folder naming for individual experiments follows the following structure: YYYYMMDD(_Project)_Experiment

Data will be accompanied by **documentation** containing all contextual and descriptive features of the research data, which allow to understand and (re)use the data. Detailed procedures of each experiment are logged in the electronic laboratory notebook (**E-notebook)**, with reference of each experiment to the directory path of the folder containing raw and analysed data files. This also includes data collection methods, protocols, and code explanation. Documentation is stored at the study- and the data-level, providing data provenance from the original source data to specific datasets linked to publications. Data will be generated following standardized **protocols**. Clear and detailed descriptions of these protocols will be stored in our lab protocol database and E-notebook.

Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the E-notebook and/or in hard copy lab notebooks that refer to specific datasets.

To allow long term access and use of research data will be stored or converted to **open file formats** as much as possible.

• Containers: TAR, ZIP
• databases: XML(xlsx), CSV, JSON
• Statistics: DTA, POR, SAS, SAV
• Images: TIFF, JPEG 2000, PNG, GIF
• Tabular data: CSV, TXT
• Text: XML(docx), PDF/A, HTML, JSON, TXT, RTF
• Sequencing data: FASTA, FASTQ

We use **controlled vocabularies** or ontologies when applicable to provide unambiguous meaning, for example:

• Gene Onotology: molecular function, cellular component, and biological role of RNA seq
• ENSEMBL or NBCI identifiers: gene identity
• HUGO Gene Nomenclature Committee: names and symbol of human genes
• FlyBase: names and symbol of Drosophila genes
• UniProt protein accessions: protein identity

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

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) and/or in hard copy lab notebooks that refer to specific datasets. All datasets will be accompanied by a README.txt file containing all the associated metadata, which will include the following elements:

• Title: free text
• Creator: Last name, first name, organization
• Date and time reference
• Subject: Choice of keywords and classifications
• Structure: internal structure of the dataset, or the meaning of abbreviations (not necessary when it is clear from the in-file documentation).
• 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.

Additionally, we will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards more specific to our data type. For specific datasets, additional metadata will be associated with the data file as appropriate.

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

Where will the data be stored?

Digital data

- Primary storage for **active digital files** will be on KU Leuven servers. KU Leuven offers fast ("J-drive) and slower ("L-drive") storage that allows reading/writing/modification of non-confidential, confidential, and strictly confidential data.
- KU Leuven further offers the ManGO platform for storage and management of **large volumes of active research data**. This platform allows secure storage, manual and automated metadata coupling, data workflows, and file sharing.
- Data that is no longer active, can be **archived** on the KU Leuven K-drive, which allows reading of non-confidential, confidential, and strictly confidential data.

Physical samples

- Tissue samples: Tissues will be stored locally in the laboratory.
- Cell lines: Human cell lines will be stored locally in the laboratory in liquid nitrogen storage and will be deposited in the UZ Leuven-KU Leuven Biobank. Animal cell lines will be stored in liquid nitrogen of the laboratory.
- 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) and as a bacterial glycerol stock (-80°C). 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.
- Bacterial strains will be stored in a -80°C freezer.
- Fly line stocks are preserved as a minimum of two separate cultures, each maintained at 18°C on a 4-to-5-week generation cycle.
- Hamsters: Hamsters will be maintained in facilities of the Laboratory Animal Center of KU Leuven, which applies Standard Operation Procedures concerning housing, feeding, health monitoring to assure consistent care in accordance with European and national regulations and guidelines. All animals will be registered in the Leuven Animal Information System (LAIS) database, along with corresponding genotyping information, ethical approval documents and animal provider receipts.

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 stored in **ManGO**: Snapshots are made at regular intervals (hourly, daily and monthly) in case data needs to be recovered. The data itself is synchronized on two separate hardware storage systems, each 6 PB large, located at Leuven and at Heverlee (ICTS). The data is protected against calamities at either site by synchronizing it in real-time at hardware level.

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

KU Leuven servers offer sufficient storage for active data (J/L-drive, ManGO) and archived data (K-drive). Required data-storage volumes can be easily scaled up.

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

The buildings on our campus are restricted by badge system so only employees are allowed in and visitors are allowed under supervision after registration.
Access to KU Leuven servers is possible only through using a KU Leuven user-id and password, and user rights only grant access to their own data, or data that was shared to them. Data in these drives 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 as follows:

- 519€/TB/Year for the "J-drive" (fast storage, active data),
- 569,2€/5TB/Year for the "L-drive" (medium speed storage, active data),
- 35€/TB/Year for the ManGO platform (large volume storage, active data),

Data storage and backup costs are included in 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...).

According to KU Leuven RDM policy, relevant research data will be preserved on the university's servers for a minimum of 10 years. Such data include data that are at the basis of a publication, that can only be generated or collected once, that are generated as a result of a substantial financial or personal effort, or are likely to be reused within the research unit or in wider contexts.

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

As a general rule all research outputs (data, documentation, and metadata) related to publications will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org). We aim at communicating our results in top journals that require full disclosure upon publication of all included data, either in the main text, in supplementary material or in a separate data repository.
Other research data will be archived on KU Leuven servers as described above.

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

-The costs of digital data storage on KU Leuven servers are as follows:

- 113,84€/TB/Year for the “K-drive” (archive, inactive data).
- 569,2€/5TB/Year for the “L-drive” (medium speed storage, active data)
- 35€/TB/Year for the ManGO platform (large volume storage, active data)

Data storage and backup costs are included in general lab costs.

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

- Upon publication, datasets and metadata generated from **animal omics** will be stored in public repositories such as Zenodo or the NCBI Gene Expression Omnibus, where they will receive a unique and persistent identifier.
- Source data for **electrophysiology** experiments will be deposited on the online repository Zenodo, where they will be assigned a unique and persistent identifier.
- To ensure data findability, links and references these datasets, workflows and modes will be included in the **data availability statements** of the associated publication.

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

N/A

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.

- Yes, Intellectual Property Rights
- The researchers involved and the IP team of the VIB TechTransfer Office shall make the necessary arrangements in order to maintain an embargo on the public access of research data, at least until the essential steps in securing intellectual property (e.g. the filing of a patent application) have been taken. As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that publications need not be delayed.

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

- Upon publication, datasets and metadata generated from animal omics will be stored in public repositories such as Zenodo or the NCBI Gene Expression Omnibus, where they will receive a unique and persistent identifier.
- Source data for electrophysiology experiments will be deposited on the online repository Zenodo, where they will be assigned a unique and persistent identifier.

When will the data be made available?

All research outputs (data, documentation, code, and associated metadata) will be made openly accessible at the latest at the time of publication. No embargo will be foreseen unless imposed e.g. by pending publications, potential IP requirements – note that patent application filing will be planned so that publications need not be delayed - or ongoing projects requiring confidential data. 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.

Data is typically available under a Creative Commons CC0 1.0 Universal (CC0 1.0) Public Domain Dedication, a Creative Commons Attribution (CC-BY), or an ODC Public Domain Dedication and Licence, with a material transfer agreement when applicable. Software and code usually are available under a GNU General Public License or an Academic Non-commercial Software License.

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

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?

The researchers who generate the data are responsible for managing data, documentation, and metadata.

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

The researchers who generate the data are responsible for storage and backup, 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, with support from Patrick Vandormael.