

DMP TITLE

HETEROTYPIC AMYLOID INTERACTIONS AS MODULATORS OF SELECTIVE CELLULAR VULNERABILITY

ADMIN DETAILS

Project Name: Heterotypic amyloid interactions as modulators of selective cellular vulnerability

Project Identifier: G0C3522N

Project type: Senior research project fundamental research

Principal Investigator / Researcher: Frederic Rousseau

Project Data Contact: Béla Z Schmidt

Description: It is currently not known what determines the selective neuronal vulnerability of amyloids in neurodegenerative pathologies: we do not understand how amyloids interact with other cellular components, whether these interactions are specific, and how these interactions result in cell-specific toxic phenotypes.

Our lab has played a crucial role in elucidating how aggregation-prone regions drive the self-assembly of amyloid aggregates. We hypothesise that the same mechanism that drives the homotypic assembly of amyloids might also drive heterotypic interactions of amyloids with homologous sequence segments in other proteins. Such heterotypic aggregation could interfere with the normal function of these proteins and could explain the specific spatio-temporal emergence and spreading of aggregates in the brain. The combination of experimental limitations and the lack of a structural-molecular hypothesis directing the search has made finding interaction partners of amyloids challenging.

We propose to create a new structure-based method to search for specific interaction partners of amyloids and validate the method in vitro and in vivo. Protein aggregation plays a crucial role in neurodegenerative diseases, therefore identifying heterotypic aggregation partners of amyloids will provide new angles at understanding and eventually treating these conditions.

Institutions: KU Leuven

1. GENERAL INFORMATION**Name applicant**

Frederic Rousseau

FWO Project Number & Title

Application number: G0C3522N

English Title: Heterotypic amyloid interactions as modulators of selective cellular vulnerability

Dutch Title: Heterotypische amyloid interacties als modulatoren voor selectieve cellulaire kwetsbaarheid

Affiliation

- KU Leuven

2. DATA DESCRIPTION

Will you generate/collect new data and/or make use of existing data?

- Generate new data
- Reuse existing data

Describe the origin, type and format of the data (per dataset) and its (estimated) volume, ideally per objective or WP of the project. You might consider using the table in the guidance.

WP	Dataset	Purpose	New/ Existing (source)	Data type	Data subtype	Data format	Size	Unit	Comment
1	Human brain proteome	assembly of all proteins present in the brain to search modifier candidates from	Existing data (UniProt)	Canonical_data	Protein_sequences	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Mark-up Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	5	GB	download from Uniprot
1	Computational pipeline for searching modifier proteins	search the proteome for APRs similar to those in tau and TDP-43	Existing data (Louros Nature Comm 2022)	Derived_and_compiled_data	Algorithms_and_scripts	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	100	MB	
1	TANGO algorithm	scan for aggregation propensity and identify APRs	Existing data	Derived_and_compiled_data	Algorithms_and_scripts	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	1	MB	
1	WALTZ algorithm	scan for aggregation propensity and identify APRs	Existing data	Derived_and_compiled_data	Algorithms_and_scripts	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	1	MB	
1	FoldX Suite	homology modelling and energy calculations	Existing data	Derived_and_compiled_data	Algorithms_and_scripts	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	100	MB	
1	APR hits for	list of	New data	Derived_a	Research_docu	Text files: Rich Text Format (.rtf),	10	KB	

	tau (up to 2 mutations per 7 AA stretch			nd_compiled_data	mentation	plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;		
1	APR hits for TDP-43 (up to 2 mutations per 7 AA stretch		New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	10	KB
1	Homologous peptides		New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	10	MB
1	Homology models used to model heterotypic interactions	to calculate interaction energies	New data	Simulation_data	Protein_structures	Protein structures: Protein Data Bank format (.pdb / .pdbx);	500	MB
1	List of calculated interaction energies of homologous peptide sequences against tau APR hits	to calculate interaction likelihood	New data	Derived_and_compiled_data	Research_documentation	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	1	GB
1	List of calculated interaction energies of homologous peptide sequences	to calculate interaction likelihood	New data	Derived_and_compiled_data	Research_documentation	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	1	GB

	against TDP-43 APR hits								
1	Cytoscape software	to perform network analysis	Existing data	Derived_and_compiled_data	Algorithms_and_scripts	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	100	MB	
1	Lists of likely modifiers in the brain	generate a target list for WP2-	New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	1	GB	
2	Synthesized tau APR peptides	in vitro screening of APRs similar to tau, study modulatory effect of peptides on amyloid formed by tau APRs and full-length proteins	New data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	15-ml vials	ether stocks stored at -20 C, one vial per peptide
2	Synthesized TDP-43 APR peptides	in vitro screening of APRs similar to TDP-43, study modulatory effect of peptides on amyloid formed by TDP-43 APRs and full-length proteins	New data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	15-ml vials	ether stocks stored at -20 C, one vial per peptide
2	Synthesized peptides (modifier candidates for tau)	in vitro screening of APRs similar to tau, study modulatory effect of peptides on	New data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	250	15-ml vials	ether stocks stored at -20 C, one vial per

		amyloid formed by tau APRs and full-length proteins							peptide
2	Synthesized peptides (modifier candidates for TDP-43)	in vitro screening of APRs similar to TDP-43, study modulatory effect of peptides on amyloid formed by TDP-43 APRs and full-length proteins	New data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	250	15-ml vials	ether stocks stored at -20 C, one vial per peptide
2	Thioflavin T	reporter dye for amyloid fibril formation and structure	Existing data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	100	vial of 20 ul	
2	LCOs	reporter dye for amyloid fibril formation and structure	Existing data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	20	vial of 20 ul	
2	Aggregation kinetics	monitor amyloid formation	New data	Experimental_data	Biophysics_data	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	100	MB	
2	Fluorecent spectroscopy data	monitor mophological differences	New data	Experimental_data	Spectroscopy_data	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	1	GB	
2	AFM images	morphological analysis of aggregates formed by synthetic	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	20	GB	

		peptides similar to tau or TDP-43 APRs						
2	FTIR spectroscopy data	structural analysis of aggregates formed by synthetic peptides similar to tau or TDP-43 APRs	New data	Experimental_data	Spectroscopy_data	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	10	MB
2	Transmission electron microscopy images	morphological analysis of aggregates formed by synthetic peptides similar to tau or TDP-43 APRs	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	50	GB
2	Seeds made from recombinant TDP-43	reference sample of seeds for DIGAS and cellular assays	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	60	vial of 20 ul
2	Seeds made from recombinant tau	reference sample of seeds for DIGAS and cellular assays	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	60	vial of 20 ul
2	Recombinant plasmid tau	produce tau protein in bacteria	Existing data	Experimental_data	Vectors	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	1	vial of 20 ul
2	Recombinant plasmid TDP-43	produce TDP-43 protein in bacteria	Existing data	Experimental_data	Vectors	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	2	vial of 20 ul
2	E. coli competent cells	transform with plasmids for protein/peptide	Existing data	Experimental_data	Cell_lines	Biological and chemical samples: live animals, frozen samples in cryovials, samples	10	vial of 20 ul

		production				stored at 4°C.			
2	Glycerol stocks of bacteria transformed with recombinant plasmids	stocks to produce recombinant tau	New data	Experimental_data	Cell_lines	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	5	vials of 200 ul	
2	Glycerol stocks of bacteria transformed with recombinant plasmids	stocks to produce recombinant TDP-43	New data	Experimental_data	Cell_lines	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	5	vials of 200 ul	
2	Recombinant tau protein, several isoforms	strain and isoform analysis	New data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	50	vials of 1 mg	lyophilized vials of 1 mg each
2	Recombinant methionine TDP-43 protein	strain and isoform analysis	New data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	50	vials of 1 mg	lyophilized vials of 1 mg each
3	Tau RD P301S FRET Biosensor cell line (ATCC CRL-3275)	reporter cell line for high content imaging cellular assays	Existing data	Experimental_data	Cell_lines	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vials	2 million cells per vial, used as needed
3	TDP-43 seeding cell line	reporter cell line for high content imaging cellular assays	Existing data	Experimental_data	Cell_lines	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vials	2 million cells per vial, used as needed
3	Seeds made from recombinant TDP-43	determine seeding efficiency of seeds made from modifier peptides	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vial of 20 ul	

3	Seeds made from recombinant tau	determine seeding efficiency of seeds made from modifier peptides	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vial of 20 ul	
3	Microscope images from high-content screening	determine seeding efficiency of seeds made from modifier peptides	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	100	GB	
3	Seeding dose-response curves	determine seeding efficiency of seeds	New data	Derived_and_compiled_data	Research_documentation	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	20	MB	
3	Plasmids encoding modifier candidates	Over-expressing modifier candidates in seeding cell lines to see their effect on tau/TDP-43 seeding	New data	Experimental_data	Vectors	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	200	vials of 20 ul	stored frozen
3	RNAi	Knocking down modifier candidates in seeding cell lines to see their effect on tau/TDP-43 seeding	New data	Experimental_data	Vectors	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vials of 20 ul	stored frozen
3	Microscope images from high-content screening	determine the effect of expressed modifiers on morphology of tau/TDP-43 fibrils	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	100	GB	
3	CellTiter Blue	measure toxicity to cells	Existing data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen	100	vial of 20 ul	

						samples in cryovials, samples stored at 4°C.			
3	Cytotoxicity datasets	determine toxicity to cells	New data	Experimental_data	Cytometry_data	Quantitative tabular data: comma-separated value files (.csv), tab-delimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	10	MB	
3	Membranes containing filter-trapped material	staining with aggregate-specific A11 antibody	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	30	membranes	
3	Membranes containing filter-trapped material	staining with aggregate-specific phospho-tau specific-antibodies	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	100	membranes	
3	Scans of immunoblots	document binding of aggregate-specific A11 antibody or phospho-tau-specific antibodies	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	2	GB	
3	List of validated modulators	potential therapeutic targets	New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	1	MB	
3	Section of treated cells	distinguish strains and isoforms using FTIR and TEM	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	30	sections	stored frozen
3	Extracts of treated cells (seeded cells)	distinguish strains and isoforms using AFM-IR, fluorescence	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.			stored frozen

		spectral analysis						
3	AFM images	morphological analysis of tau or TDP-43 aggregates formed in cells	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	20	GB
3	FTIR spectroscopy data	structural analysis of tau or TDP-43 aggregates formed in cells	New data	Experimental_data	Spectroscopy_data	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	10	MB
3	Transmission electron microscopy images	morphological analysis of tau or TDP-43 aggregates formed in cells	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	50	GB
3	Fluorecent spectroscopy data	monitor morphological differences between tau or TDP-43 aggregates formed in cells	New data	Experimental_data	Spectroscopy_data	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	1	GB
3	List of most likely modifiers, based on in vitro data, cellular data, and intracellular interactions	shortlist and prioritize most likely modifiers to be detected/tested during in vivo validation	New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	10	MB
4	Homologous peptides	candidate modifiers	New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format	10	MB

						(.pdf), LaTeX (.tex) format;			
4	Homology models used to model heterotypic interactions	to calculate interaction energies	New data	Simulation_data	Protein_structures	Protein structures: Protein Data Bank format (.pdb / .pdbx);	500	MB	
4	List of calculated interaction energies of homologous peptide sequences against tau APR hits	to calculate interaction likelihood	New data	Derived_and_compiled_data	Research_documentation	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	1	GB	
4	List of calculated interaction energies of homologous peptide sequences against TDP-43 APR hits	to calculate interaction likelihood	New data	Derived_and_compiled_data	Research_documentation	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	1	GB	
4	Database of human proteins enriched in tau/TDP-43 deposits in patient brains	perform enrichment analysis	New data	Derived_and_compiled_data	Research_documentation	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);	10	MB	compiled from data reproted in the literature
4	Results of enrichment analysis	cross-reference with list of modifier candidates	New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	10	MB	

4	List of modifier candidates from WP2 and WP3	cross-reference with list results of enrichment analysis	New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	10	MB	
4	Shortlist of modifier candidates to be tested in vivo in mice	In vivo validation of modifier candidates	New data	Derived_and_compiled_data	Research_documentation	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	10	MB	
4	PS19 tau mouse line	In vivo validation of modifier candidates	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	1	colony	maintained in animal house
4	TDP-43 knock-in mouse	In vivo validation of modifier candidates	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	1	colony	maintained in animal house
4	Lentivirus or adeno-associated virus stocks encoding modifier candidates	expression of full-length modifier candidates in mouse brain	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	15-ml vials	stored frozen
4	Mouse brain tissue sections	immunohistochemical, immunofluorescent staining	New data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	200	sections	
4	Microscope images of stained tissue sections	analysis of the effect of modifier expression on the pathology in mouse brains	New data	Experimental_data	Digital_images	Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), JPEG 2000 (.jp2), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif;	100	GB	
4	Quantitation	analysis of the	New data	Derived_and_compiled_data	Research_documentation	Quantitative tabular	10	KB	

	results of aggregate size and frequency	effect of modifier expression on the pathology in mouse brains		nd_compiled_data	mentation	data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb);		
4	Manuscript	summarizing results	New data	Derived_and_compiled_data	Manuscripts	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word (.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;	100	MB

3. LEGAL & ETHICAL ISSUES

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to the file in KU Leuven's Record of Processing Activities. Be aware that registering the fact that you process personal data is a legal obligation.

- No

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- Yes

Vertebrate animals (mice) will be used in the last year of the project. We will seek the approval of the Ethical Committee for Animal Experimentation (ECD) for the study protocol.

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- Yes

The goal of the project is to find proteins that interfere with amyloid formation of tau and TDP43. Both proteins are relevant for neurodegenerative disease and both modifying proteins that promote amyloid formation and those that slow down amyloid formation by tau or TDP43 would be of great interest as potential therapeutic targets. We do hope that the proposed work will lead to tech transfer and valorisation of the research data. VIB and KU Leuven 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. Further research beyond the scope of this project may be necessary for developing a strong IP portfolio.

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. DOCUMENTATION & METADATA

What documentation will be provided to enable reuse of the data collected/generated in this 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) 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 (see more details below). The data will be generated following standardized protocols. Clear and detailed descriptions of these protocols will be stored in our lab protocol database, and published along with the results.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

❖ The following metadata standards will be used for certain datasets

- Nucleotide sequence files (vectors and sequencing) : GenBank Sequence Format (<https://fairsharing.org/FAIRsharing.rg2vmt>)
- Protein structures will be saved in Protein Data Bank Format (PDB) (<https://fairsharing.org/FAIRsharing.9y4cqW>)
- For sharing computer code, we use the Zenodo format (<https://zenodo.org/>)
- ❖ For instrument-specific datasets, additional metadata will be associated with the data file as appropriate.
- ❖ For other datasets, the metadata will include the following elements:
 - 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 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.

5. DATA STORAGE & BACK UP DURING THE FWO PROJECT

Where will the data be stored?

Digital files will be stored either on KU Leuven servers or in shared laboratory folders of an off-site online backup service. The researchers working on the project will have copies of the data files as well as of the derived and compiled data stored on their personal computers.

The Switch Lab has a professional subscription to an off-site online backup service with unlimited space, version control and roll-back capability, which will be used for storage during the project and after. There is a secondary on-campus physical backup of the online storage which synchronizes with the online content with a one-day delay.

Algorithms, scripts and software: All the relevant algorithms, scripts and software code driving the project will be stored in a private online git repository from the GitHub account of the department (<https://github.com/vibcbd>).

The screening core has a database system in place to handle the data stream from the high content imaging screen, including archiving facilities and will store the data during the project. Representative images and the quantitation of the images will be transferred to the Switch laboratory storage for long term storage.

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 bacteria 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.

Cell lines: Newly created cell lines will be stored locally in the laboratory in liquid nitrogen storage and will be deposited in the UZ Leuven-KU Leuven Biobank.

Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.

How is back up of the data provided?

The Switch Lab has a professional subscription to an off-site online backup service with unlimited space, version control and roll-back capability, which will be used for storage during the project and after. There is a secondary on-campus physical backup of the online storage which synchronizes with the online content with a one-day delay.

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

The Switch Lab has a professional subscription to an off-site online backup service with unlimited space, which will be used for storage during the project and after.

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

Data storage and backup costs are included in general lab costs. The Switch Lab has a yearly subscription to an off-site online backup service paid from the general budget of the laboratory. The yearly cost of the service is 5500 Euros. This cost includes unlimited data storage, not only the data belonging to the present project.

Electricity costs for the -80° and -20° freezers and refrigerators present in the labs as well as the cost of liquid nitrogen cryostorage are included in general lab costs.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

All notebooks and physical data are stored in the labs. Entry to the lab requires ID-card and key. Access to the digital data is u-number and password controlled.

6. DATA PRESERVATION AFTER THE FWO PROJECT

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

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

Where will the data be archived (= stored for the longer term)?

For the datasets that will be made openly accessible, we will use, whenever possible, the 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 will be stored on storage space of an online data-backup service. -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). -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 retention period of 5 years? How will the costs be covered?

Electricity costs for the -80° and -20° freezers and refrigerators present in the labs as well as for in liquid nitrogen cryostorage are included in general lab costs. The cost of the laboratory's professional subscription to the online data backup service is 5500 Euros per year (27 500 Euros for 5 years). This cost includes unlimited data storage, not only the data belonging to the present project. Data storage and backup costs are included in general lab costs.

7. DATA SHARING AND REUSE

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?

Participants to the present project are committed to publish research results to communicate them to peers and to 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). 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 data repository if requested by the journal and following deposit advice given by the journal. Depending on the journal, accessibility restrictions may apply. Physical data (e.g. cell lines) will be distributed to other parties if requested.

Where/how will the data be made available for reuse?

- The data will be shared upon request by mail.
- Possible ways of sharing the generated data:
 - nucleic acid sequences: GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>)
 - protein sequences: UniProt KB (<https://www.uniprot.org/>)
 - vectors: AddGene (<http://www.addgene.org/depositing/start-deposit/>)
 - cell lines: direct mailing on dry ice
 - microscope images: Image Data Resource (<http://idr.openmicroscopy.org/about/>)
 - proteomics data: PRIDE (<https://www.ebi.ac.uk/pride/>)
 - manuscripts: bioRxiv (<https://www.biorxiv.org/>)
 - other digital data: Zenodo data repository (<https://zenodo.org/>)

When will the data be made available?

- Upon publication of the research results

Generally, research outputs 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.

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 be 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, a Creative Commons Attribution (CC-BY) 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.

What are the expected costs for data sharing? How will the 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.

The receiving party will pay for sharing physical data (e.g. cell lines).

8. RESPONSIBILITIES

Who will be responsible for data documentation & metadata?

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 be responsible for data storage & back up during the 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 be responsible for ensuring data preservation and reuse ?

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 bears the end responsibility for updating & implementing this DMP?

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