

## **DMP TITLE**

A HIGH-THROUGHPUT PLATFORM FOR SCREENING AMYLOID STRAINS AND THEIR SEEDING CAPACITY AT SINGLE-PARTICLE RESOLUTION

## **ADMIN DETAILS**

**Project Name:** A high-throughput platform for screening amyloid strains and their seeding capacity at single-particle resolution

**Project Identifier:** 12P0922N

**Project type:** Senior Postdoctoral Fellowship

**Principal Investigator / Researcher:** Nikolaos Louros

**Project Data Contact:** Béla Z Schmidt

**Description:** The development of recent methods, such as cryo-electron microscopy, has had a great impact on our understanding of amyloid aggregates. Despite this recent influx of structural information, however, we still do not fully understand how differential intrinsic properties of amyloid strains/conformers, such as their morphology, co-aggregation potential and transmissibility barriers relate to selective cellular/tissue toxicity and aggregation spreading patterns or the role they play in the progression of neurodegenerative disorders. This complexity increases even further when considering systemic or other localised forms of amyloidosis.

We desperately need fresh approaches to diagnose and potentially treat amyloid diseases, therefore, I propose to develop a high-throughput screening platform for the analysis of strain-specific propagation rates and structural features of patient-derived amyloids at an unprecedented resolution. I will integrate cutting-edge single-molecule techniques (microfluidics, AFM-FTIR, cryo-electron microscopy, and high-content cellular screens) and will take advantage of the availability of patient-derived material to build a comprehensive database of disease-associated diagnostic markers of major proteinopathies. I will use this database and the integrated infrastructure to find cell-autonomous aggregation modulators and for the fast and accurate screening of novel amyloid-specific inhibitors with therapeutic potential.

**Institutions:** KU Leuven

## **1. GENERAL INFORMATION**

### **Name applicant**

Nikolaos Louros

### **FWO Project Number & Title**

Application number: 12P0922N

**English Title** A high-throughput platform for screening amyloid strains and their seeding capacity at single-particle resolution

**Dutch Title** Een high-throughput platform voor het screenen van amyloïdstammen en hun seedingcapaciteit met moleculaire resolutie

### **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.**

Please see data table in the following pages.

WP	Dataset	Purpose	New/ Existing (source)	Data type	Data subtype	Data format	Size	Unit	Comment
1	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	
1	Recombinant plasmid methionine amyloid beta	produce methionine amyloid beta peptide 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	
1	E. coli competent cells	transform with plasmids for protein/peptide production	Existing data	Experimental_data	Cell_lines	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vial of 20 ul	
1	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	
1	Glycerol stocks of bacteria transformed with recombinant plasmids	stocks to produce recombinant abeta	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	
1	Recombinant tau protein	use as a reference sample and to make	New data	Experimental_data	Recombinant_c	Biological and chemical samples: live animals, frozen	50	vials of 1	lyophilized

		seeds			ompoun ds	samples in cryovials, samples stored at 4°C.		mg	vials of 1 mg each
1	Recombinant methionine ab peptide	use as a reference sample and to make seeds	New data	Experiment al_data	Recomb inant_c ompoun ds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	50	vials of 1 mg	lyophili zed vials of 1 mg each
1	Seeds made from recombinant ab	reference sample of seeds for DIGAS and cellular assays	Existing data	Experiment al_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	60	vial of 20 ul	
1	Seeds made from recombinant tau	reference sample of seeds for DIGAS and cellular assays	Existing data	Experiment al_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	60	vial of 20 ul	
1	Human brain tissue	extract aggregates/generate seeds and tissue sections	Existing data	Observatio nal_data	Tissue_s amples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	20	sampl es	tissue cubes of approx. 50 mg each
1	Human seeds/extrac ts	to be tested in DIGAS, cell assay, etc.	New data	Experiment al_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	40	vial of 20 ul	
1	Human tissue sections	map using AFM-IR	New data	Experiment al_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	100	sectio ns	
1	Human tissue sample log	keeping track of use of human tissue	New data	Derived_an d_compiled _data	Researc h_docu mentati on	Quantitative tabular data: commaseparated value files (.csv), tabdelimited file (.tab), delimited text (.txt), MS	10	MB	Excel file

						Excel (.xls/.xlsx), MS Access (.mdb/.accdb);			
2	Recombinant tau protein	reporter for DIGAS, PMCA, and RT-QuIC	Existing data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	30	vials of 1 mg	
2	Recombinant methionine ab peptide	reporter for DIGAS, PMCA, and RT-QuIC	Existing data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	30	vials of 1 mg	
2	Seeds made from recombinant ab	reference sample of seeds for DIGAS and cellular assays	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	40	vials of 20 ul	
2	Seeds made from recombinant tau	reference sample of seeds for DIGAS and cellular assays	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	40	vials of 20 ul	
2	Human seeds/extracts	reference sample of seeds for DIGAS, PMCA, RT-QuIC, and cellular assays/high-content imaging	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	15	vial of 20 ul	
2	Thioflavin T	reporter dye for DIGAS, PMCA, and RT-QuIC	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 DIGAS, PMCA, and RT-QuIC	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	Optimized microfluidic protocol	Determine WoD and LLOD, automate readout	New data	Derived_and_compiled_data	Research_documentati	Text files: Rich Text Format (.rtf), plain text data (Unicode, .txt), MS Word	5	MB	

					on	(.doc/.docx), eXtensible Markup Language (.xml), Adobe Portable Document Format (.pdf), LaTeX (.tex) format;			
2	Fluorescence spectroscopy data	output of DIGAS, PMCA, and RT-QuiC	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;	15	MB	
2	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
2	Abeta1-42-mCherry seeding cell line (HEK293)	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
2	Microscope images	determine seeding efficiency of seeds	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	
2	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	20	MB	

						(.mdb/.acddb);			
3	Seeds made from recombinant ab	AFM-IR optimization	Existing 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	AFM-IR optimization	Existing 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	Human seeds/extracts	AFM-IR analysis	Existing 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	Human tissue sections	AFM-IR and cryoEM analysis	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	100	sections	
3	AFM images	morphological analysis of aggregates in seed samples or tissue samples	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	Infrared imaging	structural analysis of aggregates in seed samples or tissue samples	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	Infrared hyperspectral data	structural analysis of aggregates in seed samples or tissue samples	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	5	GB	

						(.mdb/.accdb);			
3	cryoEM images	determination of structure of aggregates in seed samples	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	Aggregate structures	determination of structure of aggregates in seed samples	New data	Derived_and_compiled_data	Molecular_models	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;	5	MB	
4	Recombinant plasmids	cellular expression	Existing data	Experimental_data	Vectors	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	100	vials	plasmid prep ordered from Twist
4	Glycerol stocks of bacteria transformed with recombinant plasmids	make plasmids stocks for expression of modulator proteins in mammalian cells	Existing data	Experimental_data	Genetically_modified_organisms	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	100	vials of 200 ul	store at -80 in vials
4	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
4	Abeta1-42-mCherry	reporter cell line for high content imaging	Existing data	Experimental_data	Cell_lines	Biological and chemical samples: live animals, frozen	10	vials	2 million



	seeding cell line (HEK293)	cellular assays				samples in cryovials, samples stored at 4°C.			cells per vial, used as needed
4	Seeds made from recombinant ab	screening aggregation modulators	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vial of 20 ul	
4	Seeds made from recombinant tau	screening aggregation modulators	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vial of 20 ul	
4	Human seeds/extracts	validating aggregation modulator hits	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	5	vial of 20 ul	
4	Microscope images	determine seeding efficiency of seeds	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	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	
4	The list of 5 strongest modulators	identifying hits	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	1	MB	

						(.pdf), LaTeX (.tex) format;			
4	Recombinant plasmids	bacterial expression of 5 strongest modulators	Existing data	Experimental_data	Vectors	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	5	vials	plasmid prep ordered from Twist
4	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.	15	vials of 200 ul	
4	Recombinant modulator protein	validating 5 strongest modulators using DIGAS, AFM-IR	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	10 vials of each protein
4	Recombinant tau protein	reporter for DIGAS	Existing data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vials of 1 mg	stored frozen
4	Recombinant methionine ab peptide	reporter for DIGAS	Existing data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vials of 1 mg	stored frozen
4	Human seeds/extracts	reference sample of seeds for DIGAS and AFM-IR	Existing data	Experimental_data	Samples	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	5	vial of 20 ul	store at -80 in vials
4	Thioflavin T	reporter dye for DIGAS	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	
4	LCOs	reporter dye for	Existing	Experiment	Syntheti	Biological and chemical	20	vial of	

		DIGAS	data	al_data	c_compound	samples: live animals, frozen samples in cryovials, samples stored at 4°C.		20 ul	
4	Optimized microfluidic protocol	validate modulators using DIGAS	Existing 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;	5	MB	
4	Fluorescence spectroscopy data	output of DIGAS	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;	15	MB	
4	AFM images	morphological changes induced by modulators	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	
4	Infrared imaging	structural changes induced by modulators	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	
4	Infrared hyperspectral data	structural changes induced by modulators	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	2	GB	

						(.mdb/.accdb);			
4	List of validated modulators	potential therapeutic targets	New data	Derived_and_d_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	
5	Predicted structural models	structure-based inhibitor designs	New data	Derived_and_d_compiled_data	Molecular_models	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	
5	Calculated interaction energies	structure-based inhibitor designs	New data	Derived_and_d_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	
5	List of predicted peptide inhibitors	structure-based inhibitor designs	New data	Derived_and_d_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	
5	Synthesized peptides	screening peptide inhibitors	New data	Experimental_data	Synthetic_compound	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	100	15-ml vials	ether stocks stored at -20 C, one vial per

									peptide
5	Recombinant tau protein	reporter for DIGAS	Existing data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vials of 1 mg	
5	Recombinant methionine ab peptide	reporter for DIGAS	Existing data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	10	vials of 1 mg	
5	Human seeds/extracts	reference sample of seeds for DIGAS and AFM-IR	Existing data	Experimental_data	Recombinant_compounds	Biological and chemical samples: live animals, frozen samples in cryovials, samples stored at 4°C.	5	vial of 20 ul	
5	Thioflavin T	reporter dye for DIGAS	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	
5	LCOs	reporter dye for DIGAS	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	
5	Optimized microfluidic protocol	validate cappers using DIGAS	Existing 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;	5	MB	
5	Fluorescence spectroscopy data	output of DIGAS	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	15	MB	

						(.pdf), LaTeX (.tex) format;			
5	AFM images	morphological changes induced by cappers	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	
5	List of experimentally validated cappers	potential therapeutic agents	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. 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

Human tissue samples will be used in the framework of study protocol S63759 approved by the Ethical Committee of Research of UZ/KU Leuven.

**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 develop a method that has great diagnostic potential and we 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 softwares: 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](http://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.