DATA MANAGEMENT PLAN (DMP)

Project number: G046824N

1. General Project Information	
Name Grant Holder & ORCID	*Lucía Chávez-Gutiérrez & 0000-0002-8239-559X*
Contributor name(s) (+ ORCID) & roles	Prof. Martin. Zacharias (0000-0001-5163-2663) & co-applicant, Prof. Peter Verwilst (0000-0002-4673-2050) & partial collaborator, Prof. Rouslan Efremov (0000-0001-7516-8658) & close collaborator, Dr. Joao Carvalho & collaborator
Project number [1] & title	G046824N, Structure-based design of stabilizers of γ-secretase-APP interactions for Alzheimer's disease
Funder(s) GrantID [2]	G008023N
Affiliation(s)	KU Leuven VIB – KU Leuven Center for Brain and Disease Research *Laboratory of Proteolytic Mechanisms mediating Neurodegeneration*
Please provide a short project description	GSEC modulators (GSMs) are small compounds that enhance GSEC processivity, leading to generation of shorter Aβ peptides, while preserving the overall proteolysis. GSMs spare essential GSEC-mediated signalling cascades. Specifically, they increase the efficiency of the sequential Aβ processing along one or both GSEC product lines by stabilizing GSEC-Aβn interactions10 and preventing the premature dissociation and release of longer Aβs13. As a result, GSMs enhance the cleavage of longer and toxic Aβ42 and Aβ43 into shorter Aβ37 and Aβ38 peptides. These shorter Aβs have been shown to be non-toxic, or even attenuate the Aβ toxicity in vivo in the Drosophila eye. Of note, recent clinical studies have linked increased CSF Aβ38 levels to slower cognitive decline in AD patients. These findings collectively indicate that the mechanism of action of GSMs is inherently safe and support the notion that GSM-based therapy might be beneficial, at least for AD prevention. we propose that generating solid structural and mechanistic understanding of GSM-GSECAPP (drug-target) interactions will facilitate structure-based discovery of chemically diverse scaffolds securing GSEC-APP complexes and thereby lowering production of longer and toxic Aβ peptides. These novel scaffolds can then serve as valuable lead compounds in the development of next-generation GSMs (acting as GSEC stabilizers)selectively and potently targeting APP processing.

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

2. 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 [3].

				ONLY FOR PHYSICAL DATA
Dataset Name	**Description**	**New or Reused**	**Digital or Physical**	**Physical Volume**
Experimental dataset 1 - Cell lines	Bacteria strains, insect cell lines, mammalian cell lines	☒ Generate new data ☒ Reuse existing data	☐ Digital ☒ Physical	200 ml
Experimental dataset 2 -Vectors	Vectors for protein expression in bacteria, mammalian and insect cells	☒ Generate new data ☒ Reuse existing data	☐ Digital ☒ Physical	5 ml
Experimental dataset 3- Recombinant proteins and compounds	Human proteins (proteases, their substrates, newly discovered interactors) produced in insect or human cell lines, nanobodies expressed in bacteria, chemical inhibitors and modulators of proteases	☒ Generate new data ☒ Reuse existing data	☐ Digital ☒ Physical	5 L

				ONLY FOR DIGITAL DATA	*ONLY FOR DIGITAL DATA*	*ONLY FOR DIGITAL DATA*
Dataset Name	**Description**	**New or Reused**	**Digital or Physical**	**Digital Data Type**	**Digital Data Format**	**Digital Data Volume (MB, GB, TB)**
Experimental dataset 4 -Canonical data	Nucleotide sequencing data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☒ Software	- Raw sequence data trace (.ab1) - Text-based format (.fasta/.fa) and accompanying QUAL file (.qual) - Genbank format (.gb/.gbk) - Sequence alignment data: (.sam), (.bam)	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA

[3] Add rows for each dataset you want to describe.

				☐ Other:		
Experimental dataset 5 - Mass spectrometry data	Mass spectrometry datasets	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☒ Software ☒ Other:	Raw data (.raw)	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☒ > 5 TB ☐ NA
Experimental dataset 6 -Digital images	Microscopy pictures, gel scans, graphs, illustrations, figures	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☒ Textual ☒ Model ☒ Software ☐ Other:	- Digital images in raster formats: uncompressed TIFF (.tif/.tiff), JPEG (.jpg), Adobe Portable Document Format (.pdf), bitmap (.bmp), .gif, .lif (Leica), .czi (Zeiss), .gel (ImageQuant) - Digital images in vector formats: scalable vector graphics (.svg), encapsulated postscript (.eps), Scalable Vector Graphics (.svg), Adobe Illustrator (.ai)	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA
Experimental dataset 7 – Cryo - EM	Cryo – EM datasets and illustrations	☒ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☒ Textual ☒ Model ☐ Software ☐ Other:	Uncompressed TIFF (.tif/.tiff), Images stacks (.mrc), 3D volumes (.mrc), Atomic coordinates (.pdb), Text files (.star)	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☒ > 5 TB ☐ NA
Derived and compiled dataset 1 – Research documentation	Documentation generated by the research and technical staff or collected from	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual	−Text files: Rich Text Format (.rtf), MS Word (.doc/.docx), Adobe Portable Document Format (.pdf) −Quantitative tabular data: comma-separated value files (.csv), MS Excel	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB

	online sources and from collaborators, including ethical approval documents, laboratory notes and protocols.			☐ Model ☐ Software ☐ Other:	(.xls/.xlsx), MS Access (.mdb/.accdb)	☐ NA
Derived and compiled dataset 2 -Manuscripts	Manuscripts resulting from the project	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	–Text files: Rich Text Format (.rtf), MS Word (.doc/.docx), Adobe Portable Document Format (.pdf)	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

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.	We will reuse part of the physical datasets that were generated in our laboratory in previous projects, such as bacteria strains, mammalian cell lines and vectors for expression. These datasets are included as part of the Experimental dataset 1, 2, 3 as described above. We will also reuse part of the digital datasets that were generated in our laboratory in previous projects (https://doi.org/10.1038/s41380-022-01518-6, https://doi.org/10.15252/embj.2022111084, https://doi.org/10.1101/2023.09.09.556900) as well as published such as Cryo – EM structural data. These datasets are included as part of Experimental dataset 7 and described above.

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, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☐ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☒ No Additional information:
Will you process personal data[4]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☐ Yes (provide PRET G-number or EC S-number below) ☒ No Additional information:
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 ☐ No If yes, please comment: 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 to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain: No third-party agreement restricts dissemination or exploitation of the data or strains generated from this project. In particular, existing agreements between VIB and KU Leuven do not restrict publication of data.

[4] See Glossary Flemish Standard Data Management Plan

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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. 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). *RDM guidance on documentation and metadata*.	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. 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-lab journal) and/or in hard copy lab notebooks that refer to specific datasets. Cryotubes of biological samples (bacterial strains, mammalian cell lines) stored at -196°C will be labelled with a reference number that links to an entry or database. All datasets will be accompanied by a README.txt file containing all the associated metadata (see more details below).

Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☒ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: While specific data types might require particular metadata, as a general rule the metadata are based on a generalized metadata schema such as Dublin Core or DataCite, including 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. For specific datasets, additional metadata will be associated with the data file as appropriate. Mass spectrometry metadata will be generated at the VIB Proteomics Core facility. A copy of the results (data) will be stored at the VIB-KU Leuven Proteolytic Mechanisms mediating Neurodegeneration laboratory. Cryo – EM data will be generated at the CryoEM facility BECM (VIB – VUB, Brussels) and stored and analyzed there on a server owned by the applicant (Chávez-Gutiérrez) laboratory. Additionally, a back-up copy of the data will be stored on datatape and maintained for 10 years.

4. Data Storage & Back-up during the Research Project	
Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☒ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☐ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☒ Large Volume Storage ☐ Digital Vault ☒ Other: Digital files are stored on KU Leuven servers. Other types of data are stored in different forms as follows; - Vectors: Vectors will be preserved in a form of purified DNA (in -20°C freezer). - Cell lines: Cryo preserved cell lines will be stored locally in the laboratory in liquid nitrogen tank. At least two vials per cell line derived from independent freezings will be stored. - Other biological and chemical samples: storage at 4°C and/or as frozen samples as appropriate. - Nucleic acid and protein sequences: All nucleic acid and protein sequences generated during the project will be stored on KU Leuven servers. - Proteomics data: Proteomics data generated during the project will be stored on KU Leuven servers. - All other digital data (research documentation and manuscripts) will be stored on KUL servers.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify) 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.

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 ☐ No There is sufficient storage and back-up capacity on all KU Leuven servers: - The "L-drive" is an easily scalable system, built from General Parallel File System (GPFS) cluster with NetApp eseries storage systems, and a CTDB samba cluster in the front-end. - The "J-drive" is based on a cluster of NetApp FAS8040 controlers with an Ontap 9.1P9 operating system.
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	Security of digital datasets in KU Leuven drives: Both the "L-drive" and "J-drive" servers are accessible only by laboratory members, and are mirrored in the second ICTS datacenter for business continuity and disaster recovery so that a copy of the data can be recovered within an hour. Security of the physical datasets that are stored in the laboratory unit are ensured by allowing access only to the laboratory members through KU Leuven ID card access.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The total cost of data storage during the project is projected to be max. 15 000 € per year. This includes the following costs: The costs of digital data storage are as follows: 174 € /TB per year for the "L-drive" and 519 € /TB per year for the "J-drive" and 17.7 €/TB per year for data tape storage. These costs will be covered by the VIB dotation budget assigned to the laboratory of Lucía Chávez-Gutiérrez.

5. 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...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)
Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☒ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☒ Other (specifiy): As a general rule, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org), at the latest at the time of publication. For all other datasets, long term storage will be ensured as follows: - Digital datasets: files will be stored on the "L-drive". - Vectors: vectors are preserved in the form of purified DNA (in -20°C freezer). - Cell lines: cell lines are stored locally in the laboratory (liquid nitrogen). - Other biological and chemical samples: storage either at 4°C and/or as frozen, as appropriate, locally in the laboratory of Proteolytic Mechanisms mediating Neurodegeneration.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The total estimated cost of data storage during the 10 years after the end of the project is approx. 40 000€ These costs will be covered by the VIB dotation budget assigned to the laboratory of Lucía Chávez-Gutiérrez.

6. Data Sharing and Reuse

Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *Note that 'available' does not necessarily mean that the data set becomes openly available, conditions for access and use may apply. Availability in this question thus entails both open & restricted access. For more information: https://wiki.surfnet.nl/display/standards/info-eu-repo/#infoeurepo-AccessRights*	☒ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify: All research outputs supporting publications will be made openly accessible. 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. Sharing policies for specific research outputs are detailed below: - Physical experimental datasets described above (cell lines, vectors, recombinant proteins and compounds) will be stored as appropriate in the laboratory. Within availability, they will be reused during and after the project. - Digital experimental datasets (nucleic acid and protein sequences, images, and spectrometry data): they are stored on the servers and can be made available for reuse. - Research documentation: All protocols used to generate published data will be described in the corresponding manuscript(s), and the related documentation will be included as supplementary information. These data and all other documents (daily logs, raw data) deposited in the E-Notebook are accessible to the PI and the research staff, for reuse during/after project. - Manuscripts: All scientific publications will be shared openly. At the time of publication, research results will be summarized on the VIB website and post-print pdf versions of publications or links to them will be made available there if allowed by copyright agreements, possibly after an embargo as determined by the publisher. Before the end of the embargo or in cases where sharing the post-print is not allowed due to copyright agreements, a pre-print version of the manuscript will be made available.
If access is restricted, please specify who will be able to access the data and under what conditions.	X

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 per dataset or data type where appropriate.	☐ Yes, privacy aspects ☒ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☐ No Structural data and novel compound structures, identified in this project, may be subjected to IP rights protection.

Where will the data be made available? If already known, please provide a repository per dataset or data type.	☐ KU Leuven RDR ☒ Other data repository (specify) ☒ Other (specify) As a general rule, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org), at the latest at the time of publication. Sharing policies for specific research outputs are detailed below: - Vectors: vectors are stored as appropriate in the laboratory. Within availability, they will be shared with interested researchers upon request. - Cell lines: cell lines are stored as appropriate in the laboratory. Within availability, they will be shared with interested researchers upon request. - Other digital datasets that support publications (including image files, and mass spectrometry data) are stored on the servers and can be made available upon request. - Antibodies, synthetic and recombinant compounds: samples are stored as appropriate in the laboratory. Within availability, they will be shared with interested researchers upon request. - Research documentation: All protocols used to generate published data will be described in the corresponding manuscript(s), and the related documentation will be included as supplementary information. These data and all other documents (daily logs, raw data) deposited in the E-Notebook are accessible to the PI and the research staff, and will be made available upon request. - Manuscripts: All scientific publications will be shared openly. At the time of publication, research results will be summarized on the VIB website and post-print pdf versions of publications or links to them will be made available there if allowed by copyright agreements, possibly after an embargo as determined by the publisher. Before the end of the embargo or in cases where sharing the post-print is not allowed due to copyright agreements, a pre-print version of the manuscript will be made available. Publications will also be automatically listed in our institutional repository, Lirias 2.0, based on the authors name and ORCID ID. - Nucleic acid and protein sequences: they are stored on the KU Leuven servers and can be made available upon request. - Data that do not support publication will be made available upon request by email.

When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify) As a general rule all 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.
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☒ CC-BY 4.0 (data) ☒ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify) Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. 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). We will follow the conditions of the repository that the data is submitted to regarding the usage licences. 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.
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No

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.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the electronic laboratory notebook (E-notebook) that refer to specific datasets.
Who will manage data storage and backup during the research project?	The research and technical staff will ensure data storage and back up, with support from René Custers and Alexander Botzki for the electronic laboratory notebook 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 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.