LAPLACE

A Data Management Plan created using DMPonline.be

Creators: Thomas Nicolaï, n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n. https://orcid.org/0000-0003-4451-1964, n.n. n.n.

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n. https://orcid.org/0000-0003-4451-1964

Data Manager: Stijn Spaepen https://orcid.org/0000-0001-5465-8287, Thomas Nicolaï, n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n., Samuel Eyley https://orcid.org/0000-0002-1929-8455

Project Administrator: Stijn Spaepen https://orcid.org/0000-0001-5465-8287, Thomas Nicolaï, n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n., n.n. n.n., Samuel Eyley https://orcid.org/0000-0002-1929-8455

Grant number / URL: S004624N

ID: 202211

Start date: 01-10-2023

End date: 30-09-2027

Project abstract:

The LAPLACE project aims to transform polylactic acid (PLA) production by addressing critical challenges in four key areas. By exploring lignocellulose as a feedstock, optimizing fermentation processes, streamlining lactide production, and enhancing material properties with self-reinforced PLA composites, LAPLACE seeks to advance a more sustainable, cost-effective, and environmentally friendly generation of PLA for broader industrial adoption.

Last modified: 28-03-2024

LAPLACE
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

Type of data	Format	Volume (max)	WP's involved
Fractionation of biomass and catalyst development for RCF	.xlsx .pzfx .PDF	1 GB	WP 1
Saccharification experiments	.xlsx .pzfx	1 GB	WP 1 WP 2
Data generated from purification development	.xlsv .csv .PDF	1 GB	WP 4
Catalyst design of gas phase reactor for lactide production: property-activity relationship, kinetic modelling and mechanism elucidation	.xlsx .PDF	1 GB	WP 5
Process modelling gas phase reactor	.apw .xlsx	1 - 5 GB	WP 5
Nanocellulose characterization	.csv .ods .txt .vms .edf .tiff .h5	10 GB	WP6
PLA nano cellulose composite characterization	.tiff .csv .txt .edf .ods .vms .h5	10 GB	WP6
Process modelling fermentation and downstream processing	.spf .xlsx	1 - 5 GB	WP 7
TEA data and reports	.spf .xlsx .pdf	1 GB	WP 7
Phenotypic data yeast screening and fermentations	.xlsx .tiff .txt	1 GB	WP 2 WP 3
whole-genome sequencing	.fastq .bam	100 GB	WP 2
Sanger sequencing	.ab1 .fasta	10 GB	WP 2 WP 3
metabolite analysis	.txt .xlsx	10 GB	WP 2 WP 3
crystal structure data	.mtz .pdb	50 GB	WP3
protein design data	.pdb .csv .fasta	10 GB	WP3
protein expression data	.fasta .csv	1 GB	WP3

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

1. Designation of responsible person
    1. WP1: Thomas Nicolaï, Quinten Mintiens, Ahmed Taleb, Bert Sels
    2. WP2: Quinten Deparis, Stijn Spaepen, Kevin Verstrepen, Jolien Smets
    3. WP3: Quinten Deparis, Stijn Spaepen, Kevin Verstrepen, Arnout Voet, Giovanni Desiderati
    4. WP4: Michiel De Middelaer (Automation Engineer BBEPP), Brecht Van der Beken (ICT support BBEPP)
    5. WP5: Thomas Nicolaï, Ahmed Table Al-Qathmi, Bert Sels
    6. WP6: Wim Thielemans, Samuel Eyley
    7. WP7: Michiel De Middelaer (Automation Engineer BBEPP), Brecht Van der Beken (ICT support BBEPP)
2. Storage capacity/repository
    - During the research: Onedrive, KU Leuven servers, KU Leuven ManGO, BBEPP servers (at BBEPP online data are stored on a Hyper-V server (RAID-10, capacity 3 TB) within the domain controller (DC) Virtual Machine)
    - After the research: Onedrive, KU Leuven servers, KU Leuven ManGO, KU Leuven RDR, BBEPP servers (at BBEPP online data are stored on a Hyper-V server (RAID-10, capacity 3 TB) within the domain controller (DC) Virtual Machine.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

/

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

/

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

/

LAPLACE
FWO DMP (Flemish Standard DMP)

---

1. Research Data Summary

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

Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
RCF of biomass and saccharification	Analysis of all fractions origination from RCF and saccharification of the pulps with enzyme cocktails	New	• Digital (analysis) • Physical (lignin oil, water fraction and pulp)	• Experimental • Compiled data	.xlsx, .csv, .prism, .txt	<100 GB	Per experiment pulp: 30g - 100g Lignin oil: 20g - 40g Water fraction: 220 mL - 440 mL
Catalyst design for RCF	Design of new catalyst for RCF	New	• Digital (analysis) • Physical (catalyst)	• Experimental • Compiled data	.xlsx, .csv, .prism, .txt	<100 GB	per experiment: 0.2g - 13g
Catalyst design for production of lactide + kinetics	Catalyst design of gas phase reactor for lactide production: property-activity relationship, kinetic modelling and mechanism elucidation	New	• Digital (analysis) • Physical (catalyst)	• Experimental • Compiled data • Simulation data	.xlsx, .csv, .prism, .txt	<100 GB	0.1-1g
Purification development	Process and analytical data and reports	New	• Digital • Physical	• Observational • Experimental • Compiled data	.xlsx, .csv, .txt, .pdf	< 100 GB	Per experiment (0.2-20g)
Process modelling	SPD process models and parameters	New	• Digital	• Simulation data	.spf, .xlsx	< 100 GB	
TEA	TEA data and reports	New	• Digital	• Compiled data • Simulation data	.spf, .xlsx, .pdf	< 100 GB	
Characterization data (WP6 - SusMat)	FTIR, NMR, XPS, TGA, EA, SWAXS, SEC	New	Digital	Experimental	.jdx (FTIR, NMR), .vms (XPS), .txt (TGA), .ods (EA), .edf, .h5, .dat (SWAXS)	<100 GB	
Phenotypic and genotypic data set on K. marxianus strains	WGS, plate screening, metabolite profiles	New	Digital	Experimental Compiled data	.xlsx .txt .fastq	< 100 GB	
Engineered K. marxianus / S. cerevisiae strains	Design and construction of optimized enzymes and pathways	New	Digital Physical	Experimental Compiled data	.fasta		100's of yeast and E. coli strains in microwell plates or cryo vials
Protein design and expression data	Different designs of proteins and expression in relevant organisms	New	Digital	Experimental Compiled data	.pdb .csv .fasta	< 100 GB	
Metabolite production of engineered yeast strains	Comprehensive metabolite analysis under relevant fermentation conditions	New	Digital	Experimental	.xlsx .txt	< 10 GB	

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:

1) The literature study. Data will be compiled from existing data available through journal publishers websites (DOI or URL) and exported as RIS, TXT or PDF files to online (or desktop) library manager or reference manager (Google Scholar, EndNote, Mendeley or Zotero). Processed literature data will be summarized in a

review document (.docx, .xls, . pptx)
2) physical samples (liquid and solid) available from preliminary experiments (e.g., RCF streams) can be reused for analysis and testing.

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

- No

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

- No

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

- Yes

This work might have potential for tech transfer and valorization. Therefore, there will be restrictions for data disclosure as it may contain IP-sensitive information. All data will be subjected for their patentability prior to any publication. If applicable, patent applications will be filled. IP management will be conducted as described in the collaboration agreement SWOK: Patent applications will be made for relevant innovative procedures/products and the costs will be shared according to ownership percentage that shall be subject to negotiation on a pro rata basis. The IP can be filled by one partner on exclusive project results or by multiple partners to protect common project results (joined IP). Following the SWOK, the intention to file a patent application should be reported to the consortium at least two weeks prior to the filing date. Furthermore, to avoid publication of IPR sensitive results, all partners will be responsible for transfer of anticipated publications.

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

- Yes

For this FWO project a research collaboration agreement is drawn up (i.e. SWOK). Furthermore, material product samples will be transferred between partners. Common project results and data will be shared in (bilateral) technical meetings (.pdf). Future (joined) anticipated publications or patent applications will be discussed in consortium meetings. The agreements concerning IPR are, briefly, reported in the previous question and described in the SWOK.

In addition, dissemination of data and related experiments with material obtained from industrial partners, are subjected to the agreements that are made bilaterally between project partners and the industrial partners. These are defined in the SWOK and additional documents (e.g., MTA or NDA) signed by both parties. Examples of these are, but not limited to, enzymes provided by Metgen and catalysts provided by Johnson Matthey.

For all base strains used in this project, Verstrepen lab will check, together with René Custers (Regulatory & Responsible Research Manager at VIB) whether we need to comply with the Nagoya protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization. In case we need to comply, we will meet certain diligence obligations under EU Regulation (EU) No 511/2014 (apply for ‘Prior Informed Consent’ document and negotiate ‘Mutually Agreed Terms” with Competent National Authority of the provider country).

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

- Yes

Intellectual property rights and ownership are defined in the SWOK and, possibly, additional documents (e.g., MTA or NDA) signed by partners within the project and industrial partners. Examples of these are, but not limited to, enzymes provided by Metgen and catalysts provided by Johnson Matthey.

2. Documentation and Metadata

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

Documentation levels and types:

- **Project-Level Documentation**: Comprehensive documentation outlining the overall project objectives, methodologies, and key milestones. (.docx, .ppt, .xlsx)
- **Experiment-Level Documentation**: Detailed records for each experiment, including the purpose, methods, variables, and results embedded in documents of each experiment. (.docx, .ppt, .xlsx)

Collaborative Documentation:

- **Collaboration Platforms**: Utilizing collaborative platforms (e.g., Google Workspace, Microsoft Teams) to facilitate real-time communication, shared documentation and overview of sample flows among team members.
- **Regular Meetings**: Holding regular meetings to discuss and update project documentation collaboratively, ensuring alignment with project goals.

The project coordinator (CSCE) has established a shared folder in the Teams application. This folder contains an overview table that needs to be completed for each sample shared among partners to maintain clarity, along with a brief description. Specific numbers will be assigned to each sample shared, which will be used as references in future discussions.

At the Verstrepen lab (KU Leuven - VIB), data (digital files) generated in this project will be stored in a Dropbox Business Advanced account for processing and analyses; following secure data transfer, modern data encryption standards, and encrypted block storage (256-bit AES and SSL/TLS encryption). For more details see:
https://www.dropbox.com/business/trust. Additionally, project data and sequencing data will be backed up to KU Leuven servers.
Digital data files will be accompanied with a read me text file that contains relevant metadata for understanding and re-use of data.
All the relevant scripts driving the project will be stored on a secure Dropbox account. Scripts used for analysis will also be stored in Jupyter notebook (jupyter.org - an open source web application to store and share scripts), in github or in the GitLab service of KU Leuven.

At BBEPP, batch records with standardized outline and formatting will be prepared and completed with description of novel process technologies and optimized conditions, including a sample plan, analytical procedures and non-conformities during trials. Analytical data of obtained product streams from HPLC analysis will be stored as data files obtained directly from the Chemstation software (version C.01.05).

SusMat (KU Leuven): all experimental notes are stored in an on-site electronic lab notebook server. Data is linked through a serial number to the appropriate experimental notes. Additional metadata is extracted from data files on a per technique basis and added to the ManGO (storage platform) metadata catalogue to increase findability. The technique and project identifier are also added to all data during upload in order to increase findability.

At the Voet LAb (KUL LBMD) all data is double backed up in sinology NAS storage with 132TB capacity during the duration of the project and linked to benchling notebooks. All folders have README files to link files to the digital labbook. CSV files link the physical sample locations with the digital information. Scripts are stored in jupyter notebooks and Github repositories.

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

- Yes

The metadata for the data in this project entails:

- Creator of the dataset
- Name of the dataset
- File type of the dataset (depending on the employed software different file formats will be generated).
- Date of generation
- Data type (experimental or modelled)
- Software employed to generate the data (in case of modelled)

SusMat (KU Leuven): Experiment identifier, technique and project ID metadata is attached to all data in the ManGO storage platform, along with technique dependent data (extracted from the header of the data files). Data published to RDR will adhere to the DataCite standard.

Verstepen lab (KU Leuven - VIB): Since there is no formally acknowledged metadata standard specific to our discipline, Dublin Core Metadata will be used. Moreover, we will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards that are more specific to our data.

Voet lab (KUL LBMD) : There is no official METAdata standard in our protein engineering discipline but at least the Dublin Core Metadata information will be included. For crystal derived data these are embeddedalready in the mtz format.

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

Where will the data be stored?

All data will be stored in a cloud service offered by the KU Leuven university or others used by partners as specified below:

- J-drive is used at CSCE for raw data collection and transfer
- Personal I-drives will be used by CSCE members working on the project to have a personal back-ups of their files
- OneDrive of KUL is used as storage of all data related to project
- BBEPP online data server
- SusMat will use KU Leuven irods (SET zone): ManGO and ELN server within KU Leuven
- Verstrepen lab (KU Leuven VIB): data (digital files) generated in this project will be stored in a Dropbox Business Advanced account for processing and analyses; following secure data transfer, modern data encryption standards, and encrypted block storage (256-bit AES and SSL/TLS encryption. For more

details see: https://www.dropbox.com/business/trust
Sequencing data will be stored on an internal lab server (present in host lab) as well as on a secure Dropbox Business account for processing and analyses.
All the relevant algorithms, scripts and software code driving the project will be stored on a secure Dropbox account. Scripts used for analysis will also be stored in Jupyter notebook (jupyter.org - an open source web application to store and share scripts), in github or in the GitLab service of KU Leuven.
Upon publication, all sequences supporting a manuscript will be made publicly available via repositories such as the GenBank database or the European Nucleotide Archive (nucleotide sequences from primers / new genes / new genomes), NCBI Gene Expression Omnibus (microarray data / RNA-seq data / CHIPseq data), the Protein Database (for protein sequences), the EBI European Genome-phenome Archive (EGA) for personally identifiable (epi)genome and transcriptome sequences.
- Voet lab (KUL LBMD) will store data on a double backed up (physical) NAS of 132TB of which 16TB is dedicated to this project (working data in protein desing) as well as the OneDrive for final datasets.

How will the data be backed up?

- At KU Leuven, the data will be stored on the university's central servers with automatic daily back-up procedures.
- ManGO is mirrored between two physically separated datacenters and protected by snapshots.
- ELN server is stored on KUL server file storage with mirroring and snapshotting.
- At BBEPP online data are stored on a Hyper-V server (RAID-10, capacity 3 TB) within the domain controller (DC) Virtual Machine.
- Verstrepen lab (KU Leuven - VIB): data (digital files) are automatically backed up by the secure Dropbox Business Advanced account cloud backup services. Additionally, project data and sequencing data will be backed up to KU Leuven servers.
- Voet lab (KUL LBMD) : all data of the workstations is backed up in the NAS server of the lab, documents of manuscripts reports , final solved crystal structures and designed models are also backed-up on the Onedrive of the involved researchers.

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

- Each KU Leuven employee has 2 TB on storage facility
- At BBEPP online data are stored on a Hyper-V server (RAID-10, capacity 3 TB) within the domain controller (DC) Virtual Machine.
- Verstrepen lab (KU Leuven - VIB): Dropbox Business offers unlimited storage and back-up capacity in their clouds.
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 controllers with an Ontap 9.1P9 operating system.
- The NAS at the Voet lab has sufficient space allocated for this project aside form the KU Leuven employee storage

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

- The KU Leuven central storage is secured and the access to the data can be strictly authorized to only researcher involved directly in project execution.
- Both ManGO and the ELN are protected through KU Leuven central login with mandatory MFA. Access to the data is limited to those strictly required to have access.
- BBEPP: The data is stored on a Hyper-V server (RAID-10) within the domain controller (DC) Virtual Machine. This virtual machine is password protected and can only be managed by users that are member of the “Admin Function Group” within the BBEU-domain.
- Verstrepen lab (KU Leuven - VIB): Access to data stored on the Dropbox Business Advanced cloud is granted based on role based access control and all access requires layers of authentication that includes strong passwords, SSH keys, 2 factor authentication, and one time passcodes. Dropbox safeguards data with document watermarking, granular content permissions and policies, document watermarking, and legal holds.
Both the “L-drive” and “J-drive” KU Leuven 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.
- The NAS as the Voet lab cannot be accessed outside of the physical network of the workstations at the laboratory.

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

during the course of the project - there are no additional costs related to the data storage and backup. KU Leuven and BBEPP's data storage costs are both covered by each partner.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some

data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

Retained data (for each WP responsible):

- Publications (experimental manuscripts, review papers, PhD, master thesis)
- Regular progress and final reports (as ppt and doc files)
- Final processed experimental data (as xls files)
- All the biological material will be saved as well as all the experimental results that have potential for commercial valorization and/or required as a back-up for the research papers developed during the project.

- Crystal structures which will be publisehd will be deposited as PDB and MTZ data in the PDB database. Other confidential structurres will be ratined as PDB and MTZ on the OneDrive

Not retained data:

- Raw experimental data (as csv, txt, xls,.. format) - easy and low cost reproducibility

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

- At KU Leuven, the data will be stored on the university's central servers for at least 10 years, conform the KU Leuven RDM policy (e.g., K-drive of CSCE: K:\SET-CSCE-Archive-Data-D0771)
  Biological data: yeast and bacterial strains will be stored locally in the laboratory (-80°C). Other biological and chemical samples: storage at 4°C and/or as frozen samples as appropriate.
- At BBEPP, after the research project has ended, all data will be kept on the Hyper-V server (RAID-10, capacity 3 TB) within the domain controller (DC) Virtual Machine

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

- K-drive long-term storage costs are covered by project budgets. Running projects covers the costs of finished projects.
- For BBEPP, all costs regarding data storage and backup are fully covered by BBEPP.
- SusMat: 950 EUR/Tb (at current rates, based on 10 y storage) for archive storage. RDR storage is free for published datasets.

5. Data sharing and reuse

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

- Yes, in a restricted access repository (after approval, institutional access only, …)

Due to confidentiality, proprietary considerations, or other constraints, access to the data will be restricted and accessible to only KU Leuven employees and specific collaborators (VIB, BBEPP) to ensure that it is shared exclusively with those who participated in the project.

Data will be available only in a format of publications (PhD, IP, master thesis) or strictly for internal use within research group internally.

WP6 (SusMat): data will be made available openly through RDR if not prohibited due to IP.

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

1. In a restricted access repository
2. Upon request by mail and approval by responsible PI (main data owner)
3. Via publications, patents

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

- Yes, Intellectual Property Rights

As reported in question 1.5, this work might have potential for tech transfer and valorization. Therefore, there will be restrictions for data disclosure as it may contain IP-sensitive information. Note, this will not be applicable on the reference dataset, however it might be applicable on the experimental, analytical and sample storage dataset.

In the section 'article 6 Publication' of the collaboration agreement, the procedure of publication of results generated in this project covers the following requirements:

- In all forms of dissemination the collaboration of the consortium should be mentioned.
- A partner can not publish results which he/she does not own.
- The type of dissemination of results will be discussed in the steering group, taking into account the valorization potential of these results.
- A publication should be shared with the collaborating partners at least 30 days before submission, to allow them to register objection within 15 days after reception.
- If there is any objection, the involved partners will discuss this a meeting and try to find, as soon as possible, a solution, or tackle this problem by adapting the manuscript. If there is no objection, the publication can be submitted.

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

The experimental and analytical data will be stored in a restricted access repository. This data and insights will be, later, made available via publications or patents, which are accessible via LIMO (KU Leuven) and search engines like patentscope. More detailed information and data can be shared upon request by mail (corresponding author).

For KU Leuven, data related to published manuscripts will be published in RDR (subject to restriction due to IP).

As a general rule, datasets will be made openly accessible via existing platforms that support FAIR data sharing (www.fairsharing.org). Sharing policies for specific research outputs are detailed below:
Biological data: Bacteria and yeast strains will be shared upon simple request following publication, unless we identify valuable IP. In this case, we will first protect commercial exploitation, either through patenting or via an MTA that restricts the material from commercial use. Plasmids will be made available via Addgene (non-profit plasmid repository), unless we identify valuable IP.
Datasets will be deposited in open access repositories.
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 deposited in lab notebook are accessible to the PI and the research staff involved in the project, and will be made available upon request.
Manuscripts: We opt for open access publications where possible. Publications will be automatically listed in our institutional repository, Lirias 2.0, based on the authors name and ORCID ID.
Nucleic acid and protein sequences: Upon publication, all sequences supporting a manuscript will be made publicly available via repositories such as the GenBank database or the European Nucleotide Archive (nucleotide sequences from primers / new genes / new genomes), NCBI Gene Expression Omnibus (microarray data / RNA-seq data / CHIPseq data), the Protein Database (for protein sequences). The PDB database will be used to deposit the experimental crystal structures. Designed models with substrates will be deposited online in a Github account to gether with the scripts which yielded the structures.

When will the data be made available?

Upon publication of the research results of the associated manuscripts and/or vetting by LRD.

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

For SusMat: CC-BY.

For Verstrepen lab (KU Leuven - VIB): Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support FAIR data sharing. Metadata will contain sufficient information to support data interpretation and reuse, and will be conform 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 License, 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.

For others: N/A.

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

- Yes

Once the research results will be published, the DOI will be linked to the dataset including the project results. Other data stored (e.g., at KU Leuven archive), will not.

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

Publications in specific sources might be a subject of additional costs that will be paid form running projects.

6. Responsibilities

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

The postdocs and PhD students will collectively manage their own data and data documentation during the project based on mutual agreements about data sets (see part 2). They will be assisted by professor and project manager.

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

Project manager will take responsibility to organize data collection (set-up One-drive), coordinate of data management activity (incl. manual back-ups) and assist people involved in the project in proper handling of collected data (documentation of data, sharing data etc.) Professor will take a helicopter view on data management (assign roles, specify access permissions etc.) Project manager of current project (Thomas Nicolaï) will be always backed-up by other group members involved in project management (IOF valorization manager Bert Lagrain, PI Bert Sels)

Who will manage data preservation and sharing?

Research manager (Thomas Nicolaï), main PI/professor (Bert Sels)

Who will update and implement this DMP?

The end responsibility for updating and implementing the DMP is with the supervisor (promotor) and project manager.

LAPLACE
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

LAPLACE
DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable