

## DMP title

**Project Name** My plan (FWO DMP) - DMP title

**Grant Title** G0F7621N

**Principal Investigator / Researcher** Xing Yang

**Project Data Contact** xing.yang@kuleuven.be

**Description** This research aims to establish the next generation platform for ultrafast purification in bioprocessing. Firstly, limitations in mass transfer and ligand distribution will be overcome by designing 3- dimensional membrane matrix. Secondly, breakthrough in binding kinetics entails the incorporation of stimuli-responsive functionality into the membrane to achieve selective separation with reduced consumption of process materials. Lastly, computational modeling is used to bring mechanistic understanding on mass transport and binding kinetics, facilitating technology scale-up. The obtained results will directly impact efforts in bioprocessing, healthcare and broadly product purification.

**Institution** KU Leuven

### 1. General Information

**Name applicant**

Xing Yang

**FWO Project Number & Title**

**Project number:** G0F7621N

**Title:** Next Generation Membrane Platform for Ultra-fast Purification in Bioprocessing

**Affiliation**

- KU Leuven

### 2. Data description

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

- Generate new data

**Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).**

type of data	Format	Volume	How created
Microscopy images	.tif	200-500GB	microscopic images of the structure of membranes
Fourier-transform infrared spectroscopy	Uncompressed tiff	2-5GB	Qualitative surface chemistry analysis of membrane samples
X-ray photoelectron spectroscopy	uncompressed tiff	10-20GB	quantitative elemental analysis of membrane samples
Other experimental data	lab books, .txt. csv	1-2GB	The original data may consist of the experimentally acquired values during filtration testing.
computational codes/programs	.txt, .matlab	1-5GB	numerical simulation of chromatographic membrane processes

The expected types of data generated from this research are publishable research manuscripts and patents. The original data may consist of the experimentally acquired values (e.g., lab books, .txt. csv) and computational codes/programs (e.g., matlab, .txt) in the chemical engineering domain, these however may be reused for future research development. Nevertheless, not all data will be kept beyond the lifetime of the proposed project, such as invalid data (that is not reproduceable), and data that can contribute to the public knowledge domain, e.g., as journal publications and for education training purposes. Overall, there is no personal data involved.

### 3. Legal and ethical issues

**Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.**

- No

Privacy Registry Reference:

Short description of the kind of personal data that will be used:

**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)**

- No

**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

Our study may produce industrially relevant membrane separation materials that can be exploited as important separators to capture therapeutic biomolecules such as antibodies. In case an interesting the design of the membranes and separation devices will be identified, a patent application may be submitted in collaboration with KU Leuven Research & Development (LRD). However, further research outside of this FWO project will be needed to support a potential patent application.

**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 and metadata**

**What documentation will be provided to enable reuse of the data collected/generated in this project?**

Digital data:

For each WP, the (e-)labbooks will contain information on experimental design, protocols, sample evaluation (WP1 and 2), abbreviations used, structure of the data (including link with physical storage of data), and steps involved in data analysis and relevant analysis scripts (matlab or CFD scripts). A clear coding for all data files related to the project will be used. These will have the form: WPX\_TaskY\_yyyymmdd\_NameExperiment. Templates for such documentation will be provided by the PI to everyone associated to the project to allow consistent documentation. In the concluding stage of the project, a master index file containing the combined information for all WPs will be compiled which will be archived and also stored on the personal harddrives/PC of the PI. Altogether, this should allow any secondary analyst to use the data accurately and effectively.

Physical data:

Samples taken from the field or from experiments will be documented and stored in freezers for up to 1 years after the end of the project, and a file with test details (in Access) (identification/source of origin/main characteristics/storage medium/revival guide/location in the chemical storage) will be maintained. Hard copy notebooks will be stored for at least 5 years after the end of the research.

**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.**

- No

Metadata with experimental procedures (including fabrication materials, samples, evaluation performance, date and handling person, etc.), preparation information, storage, etc. will be created manually and saved to read and interpret the data for other users in the future. Persons that perform the experiments and generate the data will document this information.

During the research, the data will be stored in the university's central network drives as they are safe and automatically backed up. Sensitive data will not be carried on unprotected personal devices. When the research project ends, there are a few options to preserve the data for minimum of 5 years, such as KU Leuven/UZ Leuven servers, archiving and/or share/publish data. Also, we can make use of the general repositories such as Harvard's Dataverse and internal KU Leuven repository (i.e., Lirias).

#### **5. Data storage and backup during the FWO project**

**Where will the data be stored?**

The time-stamped master copy of the data will be kept on our research unit central storage facility (KU Leuven server) and BOX. In addition, copies will be made and kept on personal devices.

**How is backup of the data provided?**

The data will be stored on the university's central server with automatic daily back-up procedures. The general repositories such as Harvard's Dataverse and internal KU Leuven repository (i.e., Lirias) can be used. Further, data will be stored on an external hard drive.

**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

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

Costs will be covered by the project or lab budget whenever needed.

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

The access to the folder containing all project related data is restricted to the project researchers and access can only be granted by the PI. Furthermore, data on the general repositories such as Harvard's Dataverse and internal KU Leuven repository (i.e., Lirias) will be protected by a password. For extra safety we will (for the smaller data sets) also have back-ups on external hard disks that are kept at the office of the PI.

## **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, ...).**

Digital data:

KU Leuven policy on data management will be followed which entails a preservation term of 5 years. Sequencing data will be submitted to public databases, where they will be permanently archived to preserve access to the public. Other data accompanying published papers will be archived in repositories such as Lirias.

Physical data:

Refrigerated stocks of protein solutions in the lab at 0-4°C and will be available upon request. Membranes samples after evaluation runs and would be kept for further morphological or aging studies will be stored in freezers (-20°C) for up to 5 years after the end of the project.

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

The data will be stored on external hard disks and the university's central servers (with automatic back-up procedures) for at least 5 years, conform the KU Leuven RDM policy. For most published data sets, data will also be stored in data repositories like Lirias.

**What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?**

For most data repositories there are no extra costs. Whenever needed (e.g for additional external hard drives), the project or lab budget will be used to cover the costs incurred.

## **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)?**

- Yes. Specify:

If effective devices with beyond the SoA performance and unique configurations are designed, a patent application could be filed. In such case, it will not be possible to share these data immediately after generating the data.

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

All published data will be readily available. Unpublished data that are important for future project applications (FWO, internal KU Leuven funds, Vlaio,...) will not be made available immediately. In general, these data will be made available within an embargo period of 3 years.

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

- In an Open Access repository

Open access data repositories and research papers. We aim to publish in open

access journals to enhance visibility of our research.

#### **When will the data be made available?**

- After an embargo period. Specify the length of the embargo and why this is necessary

Upon publication of the research results, data will be made available either upon publication of the research results or after an embargo period of max. 3 years, e.g. when future project applications are considered based on the generated data.

#### **Who will be able to access the data and under what conditions?**

Only researchers participating in the project will be able to access the data before publishing. Upon publication everyone can access the data (e.g., Zenodo, DRYAD). Creative Commons Licenses (CC BY) will be attached to the data deposited to enable researchers to access, mine and reproduce our data.

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

Publishing costs will be covered by the FWO project, or the lab. No expected costs for public data repositories, except for DRYAD. In the latter case, costs will be covered by the project fund or the lab's budget.

### **8. Responsibilities**

#### **Who will be responsible for data documentation & metadata?**

The PI, Xing Yang, will be responsible for data documentation & metadata.

#### **Who will be responsible for data storage & back up during the project?**

The PI, Xing Yang, will be responsible for data storage & back up during the project.

#### **Who will be responsible for ensuring data preservation and reuse ?**

The PI, Xing Yang, will be responsible for ensuring data preservation and reuse.

#### **Who bears the end responsibility for updating & implementing this DMP?**

The PI Xing Yang bears the end responsibility of updating & implementing this DMP.