Multi-modal chemical probes as diagnostic tools

A Data Management Plan created using DMPonline.be

Creator: Marcin Skoreński

Affiliation: KU Leuven (KUL)

Funder: European Commission (Horizon)

Template: Horizon 2020 FAIR DMP +

Principal Investigator: Marcin Skoreński

Grant number / URL: 101109192 — MULTIGLOW

ID: 205196

Start date: 01-10-2023

End date: 30-09-2025

Project abstract:

The overarching objective of the project is the development of a detection method of active TMPRSS2 in blood samples and to determine TMPRSS2 as a useful diagnostic marker of colorectal and prostate cancer. The project is divided into two connected parts. The goal of the first, chemical part of the project is to design and synthesize “multi-modal” imaging probes for the study, detection and visualization of serine proteases. The obtained smart activity-based probes will be used to complete the goal of the second, biomedical part of the project which is centered around TMPRSS2, a membrane bound, but often secreted serine protease implicated in various cancer types. We will detect and image this protease in relevant models and patient samples, providing evidence that these novel probes will be useful diagnostic tools.

Last modified: 04-03-2024

Multi-modal chemical probes as diagnostic tools
Horizon 2020 FAIR DMP +

Version information

Version number

v1.0

Description

Update of the initial DMP (submitted in month 5) for the first periodic evaluation of the project.

Date of first version

01/09/2022

Date of last update

-

1. Data summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?

The primary purpose of data collection is to gather information that directly supports project research objectives. Data collection ensures that my findings during the project are based on empirical evidence. By collecting relevant data, I can validate the conclusions drawn from your research and enhance the credibility of your project. Data collection allows you to track the progress of my project and demonstrate its impact. By documenting the data collected and the outcomes achieved, I can provide evidence of the project's success and effectiveness.

What types and formats of data will the project generate/collect?

The data will be exported from the relevant software to formats commonly used in this field of research, e.g. to CSV tables, TIFF images, PDF files.

Will you re-use any existing data and, if so, how?

No

What is the origin of the data?

Research data will be acquired from various equipment: nuclear magnetic resonance, spectrofluorometer, high performance liquid chromatography, molecular imager and mass spectrometer, in a format matching the data.

What is the expected size of the data (if known)?

I expect a total data volume of approximately 100-200 Gb.

To whom might the data be useful ('data utility')?

Some of the collected data will be published (also as rare data) and could be used further by researchers in my area to compere and evaluate their results

2.1 FAIR data: Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata?

During microscopy experiments OME-XML (Open Microscopy Environment eXtensible Markup Language) standard will be used to save the information.

Are the data produced and/or used in the project identifiable and locatable by means of a standard identification mechanism?

Each dataset will have a unique identifier and label that distinguishes it from other datasets. This identifier will be a combination of project code and version number and will refer to the appropriate page in the laboratory notebook.

What naming conventions do you follow?

In organic chemistry each reaction is named as: MSL_01, MSL_02...
During biological experiment I number my experiments as MSL_EXP_01, MSL_EXP_02....

Will search keywords be provided that optimize possibilities for re-use?

The experiments/ synthesis of the project will begins for "MSL"

What is your approach for clear versioning?

For the same experiment but different attempts I use numeration after coma (eg. MSL_01.01, MSL_01.02...)

What metadata will be created?

OME-XML (file format for biological image data)

2.2. FAIR data: Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If some data is kept closed provide a rationale for doing so.

I will be openly available except for personal information of the patients from whom samples will be studied during the project.

How will the data be made accessible?

By publish a data paper describing and promoting dataset in a peer-reviewed data journal.

What methods or software tools are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?

Where is it possible data will be stored in the format that could be open by standard software (.txt, .pdf). If is it not possible (eg. rare NMR data), the source data will be provide without any relevant software to open it.

Where will the data and associated metadata, documentation, and code be deposited? Have you explored appropriate arrangements with the identified repository?

The articles will be archived in the KU Leuven repository 'Lirias'. I will upload, describe, and share my research data to Research Data Repository of KU Leuven (RDR). I also will include all the necessary documentation that will help others understand my data and make it fully reusable.

If there are restrictions on use, how will access be provided?

If someone would like to use stored data, they should obtain copyright permissions

2.3. FAIR data: Making data interoperable

Are the data produced in the project interoperable? What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

The most of produced data will be saved in the standard format that could be open by standard software used in the field. If is possible data will be transformed to the basic text file format (.txt, .csv) to assure the data nteroperable.

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability? In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Obtained date will be mainly numbers, The data will be marked using standard vocabularies.

2.4. FAIR data: Increase data re-use (through clarifying licenses)

How will the data be licensed to permit the widest re-use possible?

The work identified as being free of known restrictions under copyright law, including all related and neighboring rights (Public Domain Mark).

When will the data be made available for re-use? If applicable, specify why and for what period a data embargo is needed.

No embargo for data usage is needed.

Are the data produced and/or used in the project usable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

The obtained data could be used as reference for other researchers working in the similar area.

How long is it intended that the data remains re-usable?

All relevant research data will be re-usable for at least 10 years.

Are data quality assurance processes described?

Most of the measurements made during the project are conducted in duplicates or triplicates. All instruments and small equipment (pipettes, scales) are regularly calibrated.

3. Allocation of resources

What are the costs for making data FAIR in your project? How will these costs be covered?

Costs related to open access to research data are eligible as part of the Horizon 2020 grant. The possible cost differ among different journals.

Who will be responsible for data management in your project?

The principal investigator will be responsible for the data management.

What are the costs and potential value of long term preservation?

Once data are published, they will be available as long as the publisher provides access to them. Simultaneously, the data will be available from the original sources at the Chemical Biology Laboratory (KUL) repositories, where they are normally available for the next 20 years.

4. Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

The data are stored in three different, independent locations. One is the hard drive of the computer where the data were collected, the second is the hard drive of the computer where the data were analyzed. The third location is the OneDrive - KU Leuven cloud drive. This ensures data safety. Access to each location is possible only after logging in using personal KU Leuven authorization.

5. Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing?

N/A

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

I use RDM Policy at KU Leuven

Multi-modal chemical probes as diagnostic tools
GDPR Record

GDPR record

Have you registered personal data processing activities for this project?

- Not applicable

Multi-modal chemical probes as diagnostic tools
DPIA

DPIA

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

- Not applicable