

DMP title

Project Name Rational discovery of drug compounds targeting the chromatin-reading function of LEDGF/p75 towards the treatment of MLL-rearranged leukemia - DMP title

Project Identifier 1SC9822N

Principal Investigator / Researcher Thibault Vantieghem

Project Data Contact 0478328832 thibault.vantieghem@kuleuven.be

Description This project is a drug discovery project focusing on rational design of compounds targeting the PWWP domain of LEDGF/p75. This project includes computer-aided drug design followed by in vitro affinity and activity assays to verify the compounds' potency. Eventually the compounds will also be tested in cellular assays.

Institution KU Leuven

1. General Information

Name applicant

Thibault Vantieghem

FWO Project Number & Title

1SC9822N - Rational discovery of drug compounds targeting the chromatin reading function of LEDGF/p75 towards the treatment for acute MLL-rearranged leukemia

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 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).

Collected/generated new data in different work packages:

- Spreadsheets created from observational data (.xlsx, .csv)
- Images from observational data (.png, TIFF, CBF)
- Protein-compound complex structures and simulations (.pdb, .mtz, .pse, .pze, SMILES, .cif, .cms, .eaf)
- Text notes (.docx, .csv, .txt, .ape, .scf, lab book)
- Presentations of observational data (.ppt, .pdf)

WP1: Fragment screening and fragment growth

Source

1. X-ray diffraction data from soaked protein crystals
2. Chemical structures of compounds
3. Molecular docking and molecular dynamics simulations

Type

1. X-ray diffraction patterns (synchrotron), data collection sheets, images from crystals and text notes
2. Visualization of chemical structures
3. Protein-compound complex structures and corresponding analysis

Format

1. CBF, TIFF, .png, .jpg, .csv, .pdb, .mtz, .pse, .pze, .xlsx, .pdf, lab book
2. SMILES, .cif, .png
3. .pse, .pze, .pdb, .dat, .png, .xlsx, .pdf, .eaf, .cms

Estimated volume

1. 10 TB
2. 500 MB
3. 1 TB

WP2: Biophysical in vitro profiling of potential leads and development of lead series

Source

1. Affinity measurements using various techniques including MST and ITC
2. In vitro inhibition assays TR-FRET and Alphascreen
3. QSAR

Type

1. Spreadsheets with original and processed data (MST: relative fluorescence - ITC: relative heat release/uptake); visualization of processed data
2. Fluorescence counts at specific wavelengths
3. QSAR analysis

Format

1. .csv, .png
2. .csv
3. .csv, .xlsx

Estimated volume

1. 10 GB
2. 1 GB
3. 3 GB

WP3: Biological profiling of lead compounds and their optimization

Source

1. NanoBRET assay
2. Colony formation assay
3. RT-qPCR

Type

1. Fluorescence counts measured at 618 nm
2. Cell counts and pictures of cells before and after treatment
3. Fluorescence counts throughout the amplification cycles

Format

1. .csv
2. .csv, .jpg
3. .csv, .png

Estimated volume

1. 1 GB
2. 1 GB
3. 1 GB

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

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

In this project we aim to design and validate compounds targeting the LEDGF/p75 PWWP domain. Active lead compounds will be patented and possibly licensed out to industry. To this purpose we collaborate with IOF manager Dr. Frauke Christ and with LRD to enable efficient valorization of lead compounds

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?

In general, in-house collected data will be noted in the lab book unless otherwise described in that lab book. Data collected not in-house or digital data will be assigned to a specific folder in the FWO fellow's (Thibault Vantieghem) KU Leuven Onedrive. For example: PhD Biocrystallography > WP2 > MST > Date of data > Specific data. One exemption here is all data collected at synchrotrons. These data sets will be stored at the NAS of the Laboratory for Biocrystallography and will be easily to find since there is a standard known by everyone working in the Laboratory for Biocrystallography.

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

A metadata standard will not be used. However, all data will be easy to find since an excel sheet will be present in which all dates of experiments are mentioned with a short description. In this way it should be easy to find all data since they are nicely divided into different WPs and subfolders. For example: for each session at a synchrotron, soaked crystals and corresponding collected data sets will be mentioned in an excel file (i.e. 2022.03.20_SLS_PXI.xlsx) with a description of the most important notes (protein, compounds, soaking time, diffraction resolution etc.). In this way it will be easy to find that specific dataset on the NAS of the Laboratory for Biocrystallography (XXX > XXX > 20220320_SLS PXI > PIN 3 > dataset 1).

5. Data storage and backup during the FWO project

Where will the data be stored?

Data will be stored at the FWO fellow's (Thibault Vantieghem) lab PC, personal external hard drive and the NAS of the Laboratory for Biocrystallography. The data on the fellow's lab PC is automatically synchronized onto the fellow's KU Leuven Onedrive. A shared OneDrive folder is used to share data with collaborators.

How is backup of the data provided?

The data on the fellow's lab PC is also automatically synchronized with the fellow's KU Leuven Onedrive. Data from the FWO fellows (Thibault Vantieghem) KU Leuven Onedrive will also be backed up every 3 months onto a personal external hard drive. Data on the NAS of the Laboratory for Biocrystallography is backup up on itself since it is a Synology hybrid RAID system.

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 KU Leuven Onedrive provides 2 TB of storage. The network-attached storage (NAS) of the Laboratory for Biocrystallography has currently 50 TB

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

The expected costs for data storage during the project will be around 1000 euro. Costs will be partially covered both by the research budget provided alongside with the FWO fellowship of Thibault Vantieghem and the Laboratory for Biocrystallography.

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

According to the FWO fellow's (Thibault Vantieghem) lab PC: an antiviral system is installed, no connection will be made with unknown networks and no illegal programs/software will be downloaded.

According to the personal external hard drive: this will be kept safely at the fellow's home.

According to the NAS of the Laboratory for Biocrystallography: access is only provided for strictly assigned persons and protected by a password.

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, ...).

All obtained data will be archived for at least 5 years.

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

All obtained data will be archived at the Laboratory for Biocrystallography at KU Leuven.

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

The expected costs for data preservation during those 5 years will be around 500 euro. Costs will be covered by the Laboratory for Biocrystallography.

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:

Before patenting, the structures of the compounds will only be shared between close collaborators when necessary.

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

Only data which were already published/mentioned in papers will be made available after the end of the project for public use. However, for a possible continuation of this research project, all unpublished data will be available in the Laboratory for Biocrystallography. In this way, data can still be accessed on demand.

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

- Upon request by mail

When will the data be made available?

- Upon publication of the research results

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

Everyone will be able to request access to published data by mail. Nevertheless, the PI (Sergei Strelkov) and FWO fellow (Thibault Vantieghem) will, under consultation, evaluate if the request will be accepted. For example, on going collaborations/collaborators will, by great change, be accepted while competitive labs will be denied.

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

The expected costs for data sharing will be largely based on the publication of papers. The price

of publication is estimated at 3000 euro. This will be covered by the allocated project budget (FWO) and the Laboratory for Biocrystallography.

8. Responsibilities

Who will be responsible for data documentation & metadata?

Thibault Vantieghem

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

Thibault Vantieghem

Who will be responsible for ensuring data preservation and reuse ?

Sergei Strelkov

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

Thibault Vantieghem