TARGETING LEDGF/P75 FOR MIXED LINEAGE LEUKEMIA AND CHEMORESISTANT MALIGNANCIES

A Data Management Plan created using DMPonline.be

Creators: Thatcher Akele, 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.

Grant number / **URL:** 1S11223N

ID: 198446

Start date: 01-11-2022

End date: 21-10-2026

Project abstract:

Chromosomal translocation of the KMT2A gene leads to KMT2A fusions and the development of MLLrearranged (MLL-r) leukemia which mainly affects children and is linked with poor prognosis. Lens Epithelium Derived Growth Factor/p75 (LEDGF/p75) is a transcriptional co-activator that tethers multiple proteins to chromatin and is found to be essential for the initiation of MLL-r leukemia, but dispensable for hematopoiesis. As a result, blocking the LEDGF/p75-chromatin interaction should suppress KMT2A function and halt malignant transformation. Recently, my host group reported that LEDGF/p75 contributes to chemoresistance in KMT2A rearranged (KMT2A-r) cell lines. Here, I propose to investigate the role of LEDGF/p75 in chemoresistance of different cancer types and validate it as a drug target for MLL-r and chemoresistant malignancies. I will explore how the expression level of LEDGF/p75 affects chemoresistance in MLL-r, non-MLL-r leukemia, prostate and breast cancer. For cancer types where LEDGF/p75 is found to contribute to chemoresistance, I will elucidate the exact mechanism how LEDGF/p75 is involved in chemoresistance. Moreover, in collaboration with medicinal chemists and structural biologists we will develop a LEDGF/p75 inhibitor. I will characterize the profile of the inhibitors using biochemical and cellular assays. Our final aim is to obtain at least one LEDGF/p75 inhibitor that could be used for the treatment of MLL-r leukemia and LEDGF/p75 driven chemoresistant cancers.

Last modified: 26-04-2023

TARGETING LEDGF/P75 FOR MIXED LINEAGE LEUKEMIA AND CHEMORESISTANT MALIGNANCIES
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)

- Spreadsheets created from observational data (.xlsx, .csv)
- Images from observational data (.png, TIFF, CBF)
- Text notes (.docx, .csv, .txt, .ape, .scf, one note lab book)
- Presentations of observational data (.ppt, .pdf)
- biochemical experimentation readouts (western blotting) - (.png, TIFF, CBF)
- biochemical experimentation readouts (Alphascreen, TR-FRET)- (docx, .csv, .txt)

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. Joris Van Asselberghs, senior technician in our lab, guarantees the safeguarding of all data on a common storage drive with automatic back-up.
2. All data is saved on a J drive with automatic back-up (capacity of 1 TB). Furthermore, for storage of my personal data, a OneDrive made available by KU Leuven with a capacity of 2 TB (can be expanded upon my request) can be used. In addition, printouts of my research can be saved in the laboratory's archive, under responsibility of Brigitte Verheyden. Furthermore, all samples will be preserved in the appropriate freezers of -20°C and -80°C whereas cell lines are stored in liquid nitrogen. After my research, data will be stored on the OneDrive made available by KU Leuven

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)

NA

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)

NA

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

NA

TARGETING LEDGF/P75 FOR MIXED LINEAGE LEUKEMIA AND CHEMORESISTANT MALIGNANCIES
DPIA

DPIA

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

- Not applicable

TARGETING LEDGF/P75 FOR MIXED LINEAGE LEUKEMIA AND CHEMORESISTANT MALIGNANCIES
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

TARGETING LEDGF/P75 FOR MIXED LINEAGE LEUKEMIA AND CHEMORESISTANT MALIGNANCIES
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.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Cell line	generating specific LEDGF knockdown cell lines	new	physical	NA	NA		Generated cell lines are stored in liquid nitrogen (LN2) in MolMed
Biochemical experimentation readout	Western blot	new	digital	experimental	.png, .Tiff	<100GB	
Biophysical assay	DSF	new	digital	Experimental	.pcrd, .csv	<1GB	
Biophysical assay	Alpha screen	new	digital	Experimental	.txt	<1GB	
Biophysical assay	TR-FRET	new	digital	Experimental	.txt	<1GB	
Biological assay	nano BRET	new	digital	Experimental	.txt, .csv	<1GB	
cell cycle and apoptosis assay	Flow cytometry	new	digital	Experimental	.csv, .pdf, .fcs	<100GB	
Biological assay	Colony formation assay	new	digital	Experimental	.csv, .jpg	<100GB	
Cell proliferation assay	WST-1	new	digital	Experimental	.csv	<1GB	
Biological assay	RT-qPCR	new	digital	Experimental	.csv, .png		
Imaging	confocal microscopy images	new	digital	Experimental	.png, .Tiff	<100GB	
Biological assay	RNA seq	new	digital Physical	Experimental	.csv, .txt, .bam, .fastq,	<TB	RNA samples will be stored in LN2 at MolMed, KU Leuven

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:

NA

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

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 the industry. The biophysical assays (TR-FRET, DSF, Alpha screen) and biological profiling (nano BRET, colony forming assays) of the compounds will be part of the data that will be potentially patented.

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

Data on active lead compounds will not be shared due to confidentiality. Research data will be published in open access papers.

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.

- No

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).

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 (Muluembet Akele (Thatcher Akele)) KU Leuven Onedrive. For example: PhD
2022 > WB > WB001 > Date of data > Specific data.

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.

- 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 the different types of experiments and subfolders. For example: for each western blot experiment in an Excel sheet numbering is provided with the specific date (i.e 20230410_WB024_Thp1.xlsx) with a description of the most important notes (sample loading sequence on the gel, the aim of the experiment, antibody used for detection).

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

Where will the data be stored?

Data will be stored at the FWO fellow's (Muluembet Akele) lab PC, The data on the fellow's lab PC is automatically synchronized onto the fellow's KU Leuven Onedrive Ultimately, data will be stored in the archive drive in KU Leuven servers to ensure read-only mode.

How will the data be backed up?

Data will be backed up in the J drive of the molecular medicine laboratory at the KU Leuven server.

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 than could be expanded upon request. The molecular medicine laboratory has a comment storage (J drive) with a capacity of 1 TB that is automatically backed up.

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

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

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

The price of storage will be 608,08 €/TB/year. FWO grant 1S11223N

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...).

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

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

All obtained data will be archived at the molecular medicine laboratory at KU Leuven. Digital data will be archived in the Large volume storage server of KU Leuven.

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

The price of storage will be 608,08 €/TB/year. FWO grant 1S11223N

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, …)

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 of molecular virology and gene therapy. In this way, data can still be accessed on demand.

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

Members of Molecular Medicine will have access to the research data on the server. Hit compounds and will only be accessible to team members working on the project (restricted acces through passwords). After publication, data will be available upon request

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

Data on hit compounds will not be shared due to confidentiality. Research data will be published in open-access papers. The biophysical assays (TR-FRET, DSF, Alpha screen) and biological profiling (nano BRET, colony forming assays) of the hit compounds will be part of the data that will be restricted

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

All obtained data will be archived at the L drive (Large Volume storage) of the Laboratory of Molecular Virology and Gene Therapy at the KU Leuven. Everyone will be able to request access to published data by mail.

When will the data be made available?

Upon publication of the research results or upon request by mail.

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

NA

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

Not available yet

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

The expected costs for data sharing will be largely based on the publication of papers. The price of the publication is estimated at 3000 euros. This will be covered by the allocated project budget (FWO) and the Laboratory of molecular virology and gene therapy.

6. Responsibilities

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

Muluembet Akele

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

Muluembet Akele

Who will manage data preservation and sharing?

Zeger Debyser (PI) & Muluembet Akele

Who will update and implement this DMP?

Muluembet Akele