Development of KRAS/ PDE6D molecular glues to inhibit KRAS by nucleo-cytoplasmic sequestration

A Data Management Plan created using DMPonline.be

Creator: Shehab Ismail Mohamed

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: G042824N

ID: 208988

Start date: 01-01-2024

End date: 31-12-2026

Project abstract:

KRAS is a high-priority drug target in cancer. Its obligate localization to the plasma membrane is aided by the trafficking chaperone PDE6D.
We previously showed that inhibition of PDE6D can block KRAS activity. However, current PDE6D inhibitors suffer from unfavorable properties and by design lack a KRAS selectivity. We recently published that stabilization of the KRAS/ PDE6D complex can sequester KRAS via its specific C-terminus to the cytoplasm to inhibit it. We have identified small molecules, which could serve as starting points for the development of a KRAS/ PDE6D stabilizer. Such 'molecular glues' are currently of highest interest in drug development.
We here propose the following objectives:
1. To understand the cancer cell biological consequences of KRAS/ PDE6D complex stabilization.
2. To develop proof-of-concept KRAS/ PDE6D-molecular glues and characterize them biochemically and structurally.
3. To determine the effect of identified molecular glues on inhibiting KRAS-driven oncogenicity.
This BE-LU collaboration between two RAS/ PDE6D expert groups, with highly complementary expertise in cancer cell- and structural-biology, will generate the first PDE6D/ KRAS complex stabilizing molecular glues with specific efficacy in KRAS-mutant cancers.

Last modified: 04-07-2024

Development of KRAS/ PDE6D molecular glues to inhibit KRAS by nucleo-cytoplasmic sequestration
DPIA

DPIA

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

Question not answered.

Development of KRAS/ PDE6D molecular glues to inhibit KRAS by nucleo-cytoplasmic sequestration GDPR

GDPR

Have you registered personal data processing activities for this project?

Question not answered.

Development of KRAS/ PDE6D molecular glues to inhibit KRAS by nucleo-cytoplasmic sequestration 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)

we strongly support open access of resulting data, codes,
and publications with strict adherence to FAIR-data principles after securing valuable intellectual property (IP). Early and open access to peer-reviewed publications is ensured by dissemination under the green and the gold open access model. Manuscripts will be preprint server deposited. Data originate from cell biological and biochemical experiments e.g. microscopic imaging, plate reader screening data or specialized data from protein structure analysis.
Data and metadata file formats include .txt, .tiff, .xlsx,.pdb, .fasta, .fcs,.pzfx, and related.25 TB expected

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. Shehab Ismail Mohamed
2. The data (25 TB expected) will be stored for 5 years after the project ends on the university's central servers with sufficient capacity and automatic daily back-up procedures.

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)

Question not answered.

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)

Question not answered.

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

Question not answered.

Development of KRAS/ PDE6D molecular glues to inhibit KRAS by nucleo-cytoplasmic sequestration
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
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Scans and pictures of gels	gels to check protein protein interactions and stabilisation	Generate new data	Digital	Experimental	.jpg	<100GB	NA
Platereader data	Fluorescence direct and anisotropy for protein protein and protein peptide interactions	Generate new data	Digital	Experimental	.xlsx	<1 TB	NA
High Throughout	HTS results	Generate new data	Digital	Experimental	.xlsx	< 1 TB	NA
Microscopy images	imaging of live cell and IF experiments for molecules testing	Generate new data	digital	Experimental	.tiff	<5TB	NA
Plasmid maps	cloning of plasmids and constructs	Generate new data	Digital	Software	files on benchling platform	1GB	NA
Synchrotron data	X ray structures of protein complex with fragments and small molecules	Generate new data	Digital	Experimental	.mtz, .pdb	<5TB	NA

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:

Plasmid maps in Benchling files of existing plasmids that are purchased from Addgene

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

NA

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

NA

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

The project will generate hit compounds to stabilise the interaction of KRAS-PDE6D these hit compounds will be further optimized either by commercial collaboration or spin off. Data will be generated from cell assays and biochemical experiments e.g, microscopic imaging, plate reader screening data and crystal structures. Data and metadata file format include .txt, .tiff, .xlsx, .pdb, .fcs, .pzfx

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

This project is part of WEAVE, a collaboration agreement is being drafted and signed between KU Leuven and the collaborator from Luxembourg University

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

NA

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).

A detailed electronic lab notebook will be generated in OneNote for Windows10. Protocols will be available in this ELN, and general protocols will also be available in the lab's Benchling space. For images of agarose- and SDS-PAGE gels, they will be named yyyymmdd_lane-contents. For platereader data, they will be named yyyymmdd_protein_type of experiment with specific concentrations and calculations as well as excitation/emission wavelength, bandwidths, and G factor determination type will be noted down in the
.xlsx file. Microscopy data will be named yyyymmdd_protein_type of experiment and specific parameters will be attached in the documentation. All other data will be also named yyyymmdd_protein_type of experiment with summary details and stored in folders specific for the experiment that is carried out

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

Metadata will be recorded by the instrument software acquiring the data and in lab books.
The following metadata will be collected: i) name of data creator, ii) date of creation, iii) description of resources used for data creation, iv) sample ID used to identify the data, v) the format of data and vi) where and how the data can be accessed by other researchers.

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

Where will the data be stored?

Data will be stored on OneDrive for business on KU Leuven account (2TB storage space) larger data will be stored as File storage. File storage is all data storage where data is stored and managed as files within folders.
File storage is offered via the SMB protocol or the NFS protocol. The NFS protocol is only for use from servers in the central KU Leuven datacenters, while the SMB protocol is also offered to clients outside the data centers (typically for user shares such as the I-disk).
This is a highly scalable and flexible solution; the chosen performance level and volume size can be adjusted to your needs at any time.

How will the data be backed up?

Backup procedures are in place

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

External drives and 2TB storage space from oneDrive is already available and sufficient currently. ICTS will be contacted once data expansion is foreseen to arrange for data storage

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

KUL multifactor authentication is in place which insures security. furthermore no personal nor patient information involved

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

around 1k per year which will be covered by the grant

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...).

The data (25 TB expected) will be stored for 5 years after the project ends on the university's central servers with sufficient capacity and automatic daily back-up procedures.

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

Published data will be made available, other data will be archived on the KU Leuven network drive and external harddrives

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

the expected costs for the data will be around 1k per year which will be covered by the my budget

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 an Open Access repository

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

NA

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

In case of generating stablising hits. PDB codes and structures of those molecules will be subject to intellectual property rights

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

published paper will be open access, structures will be available on protein data bank. others are not known yet

When will the data be made available?

Structures not subject to IP will be deposited and released on the date of publication. Papers will be published as soon as possible

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

we strongly support open access of resulting data, codes,
and publications with strict adherence to FAIR-data principles after securing valuable intellectual property (IP). Early and open access to peer-reviewed publications is ensured by dissemination under the green and the gold open access model. Manuscripts will be preprint server deposited.

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

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

PDB structures will be deposited in protein data bank which is free. Early and open access to peer-reviewed publications is ensured by dissemination
under the green and the gold open access model which will depend on the journal. All costs will be covered by the funding

6. Responsibilities

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

Shehab Ismail Mohamed

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

Shehab Ismail Mohamed

Who will manage data preservation and sharing?

Shehab Ismail Mohamed

Who will update and implement this DMP?

Shehab Ismail Mohamed