FWO DMP Template - Flemish Standard Data Management Plan

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**Yulia Ivanova 0000-0002-6131-5693**
Contributor name(s) (+ ORCID) & roles	**Wim Dehaen promotor** **Steven De Jonghe co-promotor**
Project number[1] & title	Synthesis of fused isothiazole derivatives as inhibitors of cyclin G associated kinase (GAK) and phosphoinositide kinase FYVE-type zinc finger (PYKfyve)
Funder(s) GrantID[2]	1S04023N
Affiliation(s)	x KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR[3] identifier when possible: 05f950310
Please provide a short project description	Previous research in our lab led to the discovery of 6-aryl-isothiazolo[4,3-b]pyridines as potent and selective inhibitors of either cyclin G associated kinase (GAK) or phosphoinositide kinase FYVE-type zinc finger (PYKfyve). Both kinases are promising drug targets for the development of broadspectrum antiviral agents. Using the current chemistry, 3,6-disubstituted isothiazolo[4,3-b]pyridines can easily be prepared. In this project, we will focus on the establishment of synthetic procedures to have access to novel fused isothiazolo derivatives, which will allow us to broaden chemical space that can be studied as GAK and PYKfyve inhibitors. The design will be guided by molecular modeling and biochemical kinase assays. Promising compounds will be investigated for activity against various viruses, such as the dengue virus and the venezuelan equine encephalitis virus.
2. Research Data Summary	

---

[1] "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

[3] Research Organization Registry Community. https://ror.org/

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[4].

| | | | | *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
OBSERVATIONAL DATA AND PROCEDURES ON CHEMICAL SYNTHESIS	Detailed notes on the used amount of reagents, procedure, observations, and the obtained amount of product, written down in analog lab notebooks or the ELN.	☒ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☒ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☒ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☒ .pdf ☒ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .docx. .cdx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Several analog lab notebooks will be generated.
CHEMICAL COMPOUNDS	Vials containing several mg of final compounds.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	NA	NA	NA	Several boxes of compounds stored in Leuven Chem&Tech in stock room 01.186
NMR	Data folders generated by the spectrometer containing raw and processed	☐ Generate new data ☐ Reuse existing	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/	☐ .por ☐ .xml ☐ .tab	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB	NA

[4]Add rows for each dataset you want to describe.

data.	data			aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .csv ☒ .pdf ☒ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .exec, .png, .docx ☐ NA	☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
IR	Data folders generated by the spectrometer containing raw and processed data.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☒ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: dpt, .png, .docx ☐ NA	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	NA
UV-VIS-NIR	Data folders generated by the spectrometer containing raw and processed data.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software	☐ .por ☐ .xml ☐ .tab ☐ .csv ☒ .pdf ☐ .txt ☐ .rtf	☐ < 100 MB ☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB	NA

				□ Other □ NA	□ .dwg □ .tab □ .gml ☒ other: dsw, .xsl, .png, .docx □ NA	□ > 50 TB □ NA	
MS RELATED CHROMATOGRAPHY	Data folders generated by the spectrometer containing raw and processed data.	☒ Generate new data □ Reuse existing data	☒ Digital □ Physical	□ Observational ☒ Experimental □ Compiled/ aggregated data □ Simulation data □ Software □ Other □ NA	□ .por □ .xml □ .tab □ .csv ☒ .pdf □ .txt □ .rtf □ .dwg □ .tab □ .gml ☒ other:.png, .docx □ NA	□ < 100 MB ☒ < 1 GB □ < 100 GB □ < 1 TB □ < 5 TB □ < 10 TB □ < 50 TB □ > 50 TB □ NA	NA
MPLC	Data folders generated by the chromatography containing raw and processed data.	☒ Generate new data □ Reuse existing data	☒ Digital □ Physical	□ Observational ☒ Experimental □ Compiled/ aggregated data □ Simulation data □ Software □ Other □ NA	□ .por □ .xml □ .tab □ .csv ☒ .pdf ☒ .txt □ .rtf □ .dwg □ .tab □ .gml ☒ other:.png	□ < 100 MB ☒ < 1 GB □ < 100 GB □ < 1 TB □ < 5 TB □ < 10 TB □ < 50 TB □ > 50 TB □ NA	NA

				Simulation data / Software / Other / NA	File formats	Digital data volume
generated final compounds.				☒ Simulation data ☐ Software ☐ Other ☐ NA	☐ .pdf ☒ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other:.mol, .pdb, .png ☐ NA	☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

GUIDANCE:

DATA CAN BE DIGITAL OR PHYSICAL (FOR EXAMPLE BIOBANK, BIOLOGICAL SAMPLES, …). DATA TYPE: DATA ARE OFTEN GROUPED BY TYPE (OBSERVATIONAL, EXPERIMENTAL ETC.), FORMAT AND/OR COLLECTION/GENERATION METHOD.

EXAMPLES OF DATA TYPES: OBSERVATIONAL (E.G. SURVEY RESULTS, SENSOR READINGS, SENSORY OBSERVATIONS); EXPERIMENTAL (E.G. MICROSCOPY, SPECTROSCOPY, CHROMATOGRAMS, GENE SEQUENCES); COMPILED/AGGREGATED DATA[5] (E.G. TEXT & DATA MINING, DERIVED VARIABLES, 3D MODELLING); SIMULATION DATA (E.G. CLIMATE MODELS); SOFTWARE, ETC.

EXAMPLES OF DATA FORMATS: TABULAR DATA (.POR,. SPSS, STRUCTURED TEXT OR MARK-UP FILE XML, .TAB, .CSV), TEXTUAL DATA (.RTF, .XML, .TXT), GEOSPATIAL DATA (.DWG,. GML, ..), IMAGE DATA, AUDIO DATA, VIDEO DATA, DOCUMENTATION & COMPUTATIONAL SCRIPT.

DIGITAL DATA VOLUME: PLEASE ESTIMATE THE UPPER LIMIT OF THE VOLUME OF THE DATA PER DATASET OR DATA TYPE.

PHYSICAL VOLUME: PLEASE ESTIMATE THE PHYSICAL VOLUME OF THE RESEARCH MATERIALS (FOR EXAMPLE THE NUMBER OF RELEVANT BIOLOGICAL SAMPLES THAT NEED TO BE STORED AND PRESERVED DURING THE PROJECT AND/OR AFTER).

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.	SAR study is based on the previous discovery of potent compounds, ideally leading to obtaining novel compounds with enhanced properties. Previous studies were conducted within the group or supervised by the co-promotor. Examples of the previous SAR studies: 10.1021/jm501759m, 10.1016/j.ejmech.2021.113158, 10.1021/acs.jmedchem.8b00613.

---

[5] These data are generated by combining multiple existing datasets.

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, please describe these issues further and refer to specific datasets or data types when appropriate.	☐ Yes, human subject data ☐ Yes, animal data ☐ Yes, dual use ☒ No If yes, please describe:
Will you process personal data[6]? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☐ Yes ☒ No If yes: - Short description of the kind of personal data that will be used: - Privacy Registry Reference:
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 ☐ No If yes, please comment: - BIOLOGICAL EVALUATION DATA: the described compounds are foreseen to be active in treating the following human pathologies - (re-) emerging viral infections, Parkinson's disease, non-Hodgkin lymphoma, and various neurodegenerative diseases. Therefore, the valorization potential of this work is high with the biological evaluation data playing a key role in the assessment of how potent obtained compounds are. - PROCEDURES ON CHEMICAL SYNTHESIS: synthetic procedures toward the potent compounds. - CHARACTERIZATION DATA: all the data proving the structure of compounds (NMR, HRMS, IR, X-ray data, etc.) being essential for the development of further application.

---

[6] See Glossary Flemish Standard Data Management Plan

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 to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:
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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. 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).	Two types of data will be gathered during this project. Chemical data from experiments (reaction conditions, experimental sequence, observations, amount of reagents used) will be stored by the researcher in Microsoft OneDrive (2 TB), MBook ELN(Electronic Lab Notebooks) and analog lab notebooks. Corresponding structural identification data for the obtained compounds (raw and processed) will be uploaded in the ELN and linked to the correct experiment making data traceable. Biological data will include information on kinase inhibition and antiviral activity and will be preserved digitally and shared among all collaborators. All project data will be shared with the PI and other collaborators via Microsoft OneDrive and via the ELN.

| Will a metadata standard be used to make it easier to **find and reuse the data**?

If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse.

REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS. | ☐ Yes
☒ No
If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:

If no, please specify (where appropriate per dataset or data type) which metadata will be created:

Metadata will be saved in the OneDrive storage and will be reported in the ELN. Filenames of the data generated (raw and processed) will be annotated in the ELN as well. |

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	The active data used during the project will be will be stored by the researcher in Microsoft OneDrive (2 TB), MBook ELN(Electronic Lab Notebooks) and analog lab notebooks. Data that has to be actively used by KU Leuven or external collaborators will also be shared via Microsoft OneDrive. All project data will be shared with the PI and collaborators via Microsoft OneDrive and via the ELN. Biological data will include information on kinase inhibition and antiviral activity and will be preserved digitally as manuscripts, structure drawings, image files, and modelling information. Analog lab notebooks will be stored by the researcher during the project. Samples will be stored in the chemical inventory system.

How will the data be backed up? *What storage and backup procedures will be in place to prevent data loss? Describe the locations, storage media and procedures that will be used for storing and backing up digital and non-digital data during research.[7]* *Refer to institution-specific policies regarding backup procedures when appropriate.*	For the active data, the data is automatically backed up via OneDrive. The electronic lab book provides a back-up for the analog lab notebooks and is backed-up automatically in the ELN's cloud service. The PI stores the paper copies of the analog lab notebooks and a scanned version in the online archival storage as PDF.
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 ☐ No If yes, please specify concisely: KU Leuven OneDrive provides space of 2 TB which is regularly backed up and can be extended if needed. If no, please specify:
How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *Clearly describe the measures (in terms of physical security, network security, and security of computer systems and files) that will be taken to ensure that stored and transferred data are safe. [7]*	OneDrive is not publicly accessible and password protected. Access needs to be granted by the researcher and will be limited to the PI (and his delegate responsible for data storage). For collaborations, accessibility to the relevant files will be granted if described in a non-disclosure agreement. Archival storage at KU Leuven is not publicly accessible and only people with permission (PI and his delegate responsible for data storage) can access the data.

---

[7] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The use of OneDrive is free of charge if the capacity of 2 TB is not exceeded. Archival data storage is centrally offered via KU Leuven at 270 Euro/TB/Year. MBook ELN and the cloud service are offered by Mestrelab at 10 Euro/month for the subscription and a one-time license fee of 120 Euro. These costs are being covered by the general operating budget of the research group, or by the individual bench fee of researchers.

5. 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 data gathered by the researchers during the project that is stored in Microsoft Onedrive will be copied to archival storage rented at KU Leuven. The data stored in the MBook ELN remains accessible to the PI. The analog lab book will be stored in a physical archive. Samples will be stored for 10 years in our depository as long as the long-term stability and quantity of the compound allow it.
Where will these data be archived (stored and curated for the long-term)?	The data will be stored on the university's central servers (with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy. The ELN data remains available to the PI via the cloud subscription to the ELN. All publications accompanied with the corresponding datasets will be deposited in LIRIAS. Also, the data that is disclosed in the supporting information of publications will be stored permanently.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Data archival at KU Leuven is currently offered at 270 Euro/TB/year. Since no large datasets are generated in this project, one TB of storage will be sufficient and the price for storage for 10 years is 2700 Euro. These costs are covered by the general operating budget of the research group or by the individual bench fee of researchers.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☐ Yes, in an Open Access repository ☒ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☐ Other, please specify:
If access is restricted, please specify who will be able to access the data and under what conditions.	In general, data generated in projects will be made publicly available at the time of publication. The information that is not publicly available (incomplete datasets, those can be used for the set-up of new projects and the continuation of the group's research), will be stored in the OneDrive account of the individual team members. The group leader Prof. Dehaen and co-promotor Steven De Jonge will have co-ownership in this folder. Access to data, that is not publicly available, can be provided after signing the non-disclosure agreement.
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 per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify:

Where will the data be made available? If already known, please provide a repository per dataset or data type.	Data papers describing and promoting the research will be published in peer-reviewed journals. The final thesis will be stored in the digital repository for KU Leuven Association research (LIRIAS; https://limo.libis.be/).
When will the data be made available? *THIS COULD BE A SPECIFIC DATE (DD/MM/YYYY) OR AN INDICATION SUCH AS 'UPON PUBLICATION OF RESEARCH RESULTS'.*	Upon publication of the research results.
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *EXAMPLE ANSWER: E.G. "DATA FROM THE PROJECT THAT CAN BE SHARED WILL BE MADE AVAILABLE UNDER A CREATIVE COMMONS ATTRIBUTION LICENSE (CC-BY 4.0), SO THAT USERS HAVE TO GIVE CREDIT TO THE ORIGINAL DATA CREATORS." [8]*	Data from the project that can be shared will be made available under a creative commons attribution license (CC-BY 4.0).
Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes ☐ No If yes: Upon publication of the research results, each publication will be associated with a unique DOI.

---

[8] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

What are the expected costs for data sharing? How will these costs be covered?	No costs are related to depositing data in the LIRIAS repository of KU Leuven. Also depositing data in Mendeley Data is free. No charges apply to the publication of supporting information related to publications. Exceptionally, data will be published at a publication charging a publication fee (around 1500 Euro).

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The researcher is responsible for collecting all relevant data files and for entering the observational data in the ELN and or analog Lab notebook.
Who will manage data storage and backup during the research project?	The researcher is responsible for storing all relevant data in Microsoft OneDrive (backed up regularly), MBook ELN (backed up regularly) and analog lab notebooks.
Who will manage data preservation and sharing?	The PI (supported by a delegate data manager) is the end responsible for ensuring data preservation and reuse.
Who will update and implement this DMP?	The DMP is updated by the researcher in agreement with the PI