Hitting coronavirus replication in its core by combined inhibition of the nsp12 polymerase and associated proteins of the replication-transcription complex

A Data Management Plan created using DMPonline.be

Creators: Nikolai Bakirtzoglou, Lieve Naesens

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Lieve Naesens

Grant number / **URL:** 1S21223N

ID: 197244

Start date: 01-11-2022

End date: 30-10-2026

Project abstract:

Though vital to control the pandemic, the COVID-19 vaccines offer only short-term protection due to waning immunity and virus variability. Antiviral drugs are crucial to treat fragile persons at risk of severe COVID-19, and to build stockpiles against future coronavirus (CoV) outbreaks. Avoiding resistance calls for drug combinations. This project focusses on EDA-1, a novel nucleoside analogue with dual activity against CoV and influenza virus. We synergistically combine this inhibitor of the CoV non-structural protein (nsp) 12 polymerase with an inhibitor of nsp15 endoribonuclease, a core protein of the replication-transcription complex (RTC). From these two lead molecules, analogues with optimized pan-CoV activity will be developed. Using dedicated virological and biochemical assays, the inhibitory mechanism towards the virus and target protein will be unraveled. For EDA-1, this involves assessing how termination of CoV RNA synthesis and avoidance of virus mutagenesis are achieved, and which modifications prevent excision by the nsp14 exonuclease. Next, we will produce the nsp proteins to reveal the cryo-EM structure of the RTC core, including the central role of nsp15 in shaping this complex. Finally, we aim to provide the first cryo-EM structure of native RTCs isolated from infected cells. Hence, besides delivering innovative CoV inhibitors with relevance for drug development, our findings will offer unique structural insight in the machinery performing CoV RNA synthesis.

Last modified: 29-04-2023

Hitting coronavirus replication in its core by combined inhibition of the nsp12 polymerase and associated proteins of the replication-transcription complex
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)

During this project I will be generating new data. I will not be working with personal data.
The type of data that I will be collecting is experimental and they some are included in the list below:

- Antiviral compound screening experiments will produce excel files (.xlsx) with numerical data (1-100 MB)
- Virus titration will produce excel files (.xlsx) with numerical data (1-100 MB)
- RT-qPCR will produce numerical data with Software 7500 MB v2.0.6 (.SDS) with numerical data (1-100 MB)
- Polymerase assays will produce GEL files (.gel) using a typhoon imager

Note: Cryo-EM will be performed by collaborators. They will be responsible for management of this structural data.

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. All complete protocols containing data, or other data files, will be stored. Myself (Nikolai Bakirtzoglou) and my supervisor Professor Lieve Naesens has access to our OneDrive folders and will keep them for as long as necessary, and for a minimum of 5 years. Nikolai Bakirtzoglou and Lieve Naesens are the responsible people for data preservation and can be contacted by the FWO with questions of long-term preservation of the data. After finishing the PhD project, the folders are archived on our shared driver. They will stay there for a minimum of 10 years.
2. For storage of all our data we use cloud software (OneDrive, using a KULeuven license) for short to midterm storage and to share files with team members and collaborators. Our Onedrive has automatic back-up procedures to ensure that my data is not lost. For long term storage/ar2chive we use a shared drive with automatic back-up procedures, maintained by the ICTS managers of the Rega institute. As far as storage capacity is concerned Onedrive has a 2 terabyte storage capacity which is more than enough for the data sets that I am storing as most files our between 1-100 MB. All my data will also be backed up to the shared drive also which also contains back-up procedures (managed by ICTS) and has a very high storage capacity if necessary.
   Finally our Onedrive has a strong security policy for data storage with a multifactor authentication system ensuring the safety of our data. The shared drive is also very well secured by the ICTS managers of the Rega Institute. Myself (Nikolai Bakirtzoglou) as well as my supervisor Professor Lieve Naesens have access to this data.

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)

At this point and time I have no reason to deviate from the principle of preservation of data and of the minimum preservation term of 5 years.

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)

There are no such issues concerning my research data as I will not work with personal data, human participants, animals, or data with dual use/misuse potential.

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

I have no other issues related to the data management that are relevant to mention.

Hitting coronavirus replication in its core by combined inhibition of the nsp12 polymerase and associated proteins of the replication-transcription complex
DPIA

DPIA

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

- Not applicable

Hitting coronavirus replication in its core by combined inhibition of the nsp12 polymerase and associated proteins of the replication-transcription complex
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Hitting coronavirus replication in its core by combined inhibition of the nsp12 polymerase and associated proteins of the replication-transcription complex
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
Antiviral data	Compound screening experiments will produce excel files with numerical data using qPCR or MTS quantification	New data	Digital	Experimental	.xlsx	<100MB	
Mitochondrial toxicity assay data	Mitochondrial toxicity assessment experiments will produce excel files with numerical data	New data	Digital	Experimental	.xlsx	<100MB	
Mutagenesis assay data	Mutagenesis assessment experiments will produce excel files with numerical data	New data	Digital	Experimental	.xlsx	<100MB	
Drug synergy data	Drug synergy experiments will produce excel files with numerical data	New data	Digital	Experimental	.xlsx	<100MB	
Virus titration data	Virus titration experiments will produce excel files with numerical data	New data	Digital	Experimental	.xlsx	<100MB	
RT-qPCR data	RT-qPCR assays will produce numerical data with Software 7500 MB v2.0.6	New data	Digital	Experimental	.SDS	<100MB	
Polymerase enzyme assay data	Polymerase enzyme assays will produce GEL files using a typhoon imager	New data	Digital	Experimental	.gel	<100MB	
Other enzyme assays data	Other enzyme assays will produce GEL files using a typhoon imager	New data	Digital	Experimental	.gel	<100MB	
Note: Cryo-EM will be performed by collaborators. They will be responsible for management of this structural data.							
Note: Synthesis of optimized analogues of compounds will be performed by collaborators. They will be responsible for management of this data.							

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:

I will reuse data of protein and gene sequences from databanks such as ncbi and uniprot.

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

The project may lead to a novel compound development, for which IP rights may be acquired in the future.

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.

- No

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).

For each experiment a protocol is prepared in excel containing:
1. An introduction section with background information about the experiment.
2. An experimental section explaining clearly every step of the experimental process and noting any variation in the conditions of the experiment at any point, as well as dates, incubation times, cell densities and any other detail necessary for experimental reproducibility
3. A results sheet depicting all the results of the experiment in detail
4. Finally a conclusion sheet which mentions a summary and explanation of the results as well any troubleshooting steps needed to further optimise the assay.
The results of the experiment are also saved in our Onedrive server for data safety, for easy comparison between experiments and so our team members and collaborators can easily access the information.

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

Same as above.

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

Where will the data be stored?

For short to midterm storage we use the cloud software with a KU Leuven license. This will also be used to share files with team members and collaborators.
For long term storage/archive: we use a shared drive with automatic back-up procedures, maintained by the ICTS managers of the Rega institute.

How will the data be backed up?

Our Onedrive server contains automatic backup procedures.
Our shared drive also contains backup procedures and is managed by the ICTS.

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

Our cloud software has 2 terabyte storage capacity which is sufficient for the files that we use during the project.

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

We use OneDrive for data storage, which has an excellent security policy (multi-factor authentication).
The shared drive is also very well secured by the ICTS managers of the Rega Institute.

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

The costs for OneDrive for Business are incurred by the KU Leuven.

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 complete protocols containing data, or other data files, will be stored. My supervisor Professor Lieve Naesens has access to our OneDrive folders and will keep them for as long as necessary, and for a minimum of 5 years. After finishing the PhD project, the folders are archived on our shared driver. They will stay there for a minimum of 10 years.

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

The data will be stored for at least 10 years on the shared drive with automatic back-up procedures, managed by the ICTS managers of the Rega institute of the KU Leuven.

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

The costs for data preservation are arranged centrally and are incurred by the KU Leuven.

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.

- Other, please specify:

Yes all generated data will be used in published articles and in my PhD thesis.

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

Access may be restricted for some unpublished data such as the structure of compounds that have not yet been published. Access may be restricted to anyone who hasn't signed a 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 in the comment section per dataset or data type where appropriate.

- Yes, Intellectual Property Rights

Access may be restricted for some unpublished data such as the structure of compounds that have not yet been published. Access may be restricted to anyone who hasn't signed a non disclosure agreement.

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

Data will be made available via the RDR, the KU Leuven institutional repository.

When will the data be made available?

At the time of publication in a journal, or at the time of a conference.

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

The use of specific data usage licenses is not yet known.

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

A DOI will be available through RDR, but is not yet available

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

The costs of data sharing are expected to be very small as the RDR is covered for KU Leuven staff.

6. Responsibilities

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

Nikolai Bakirtzoglou and Lieve Naesens

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

Nikolai Bakirtzoglou and Lieve Naesens and the ICTS managers of the Rega Institute

Who will manage data preservation and sharing?

Lieve Naesens and the ICTS managers of the Rega Institute

Who will update and implement this DMP?

Lieve Naesens, bears the end responsibility of updating and implementing this DMP.