Dual targeting of nsp12 and nsp15: a powerful concept to suppress coronavirus replication and transcription

A Data Management Plan created using DMPonline.be

Creators: Lieve Naesens, n.n. n.n., n.n. n.n., n.n. n.n.

Affiliation: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Principal Investigator: n.n. n.n., n.n. n.n.

Project Administrator: n.n. n.n.

Grant number / **URL:** C24E/22/035

ID: 197923

Start date: 01-10-2022

End date: 30-09-2026

Project abstract:

To treat SARS-CoV-2 infected persons at risk of severe COVID-19 and prepare against future coronavirus (CoV) outbreaks, effective antiviral drugs are highly needed. These drugs should have broad CoV coverage and a high barrier for resistance. We here develop dual targeting of two conserved components of the viral replication-transcription complex (RTC): non-structural protein 12 (nsp12) and nsp15. Starting from lead compounds
identified prior to this work, we will optimize the pan-CoV activity and disclose the inhibitory mechanism towards the virus and target protein, i.e. nsp12 polymerase or nsp15 endoribonuclease. This includes the potential of nsp15 inhibitors to boost the interferon response in infected cells. The synergy of nsp12-nsp15 inhibitor combinations will be examined. Also, the crystallographic and cryo-EM structures of the RTC core and
native RTCs will be resolved to enhance anti-CoV drug development and contribute to a better structural understanding of the machinery performing CoV RNA synthesis.

Last modified: 31-03-2023

Dual targeting of nsp12 and nsp15: a powerful concept to suppress coronavirus replication and transcription

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.

Dataset name / ID	Description	New or reuse	Digital or Physical data	Data Type	File format	Data volume	Physical volume
		Indicate: ***N****(ew data) or* ***E****(xisting data)*	Indicate: **D**(igital) or **P**(hysical)	Indicate: **A**udiovisual **I**mages **S**ound **N**umerical **T**extual **M**odel **SO**ftware Other (specify)		Indicate: <1GB <100GB <1TB <5TB >5TB NA	
Coronaviruses	WT and mutant CoV stocks produced in cell culture	**N and E**	**P**	**Biological material**	Not applicable	Not applicable	-80°C, storage managed with FreezerPro software
Cells	Cell lines permissive to CoVs	**N and E**	**P**	**Biological material**	Not applicable	Not applicable	Liquid N2, storage managed with FreezerPro software
Reverse genetics	Plasmids with CoV genome fragments	**N and E**	**P**	**Biological material**	Not applicable	Not applicable	-80°C, storage managed with FreezerPro software
Virus sequencing	Data provided by Macrogen	**N and E**	**D**	**I**	.ab1 .phd.1 .pdf	< 1 GB	None
Microscopy	Scores of viral CPE and cell morphology	**N**	**D**	**N**	.pdf	< 1 GB	None
MTS	OD values after MTS staining	**N**	**D**	**N**	.xls	< 1 GB	None
RT-PCR	RT-PCR curves and Ct values	**N**	**D**	**N**	.sds .xls	< 1 GB	None
HCI	High-content imaging of stained cells	**N**	**D**	**I and N**	.xls .tif	< 1 GB	None
Nsp12 RdRp	Primer extension assay	**N**	**D**	**I**	.pdf .gel .tif	< 1 GB	None
Nsp15 EndoU	FRET cleavage assay	**N**	**D**	**N**	.xls	< 1 GB	None
Compound binding	Data from SPR or thermal shift assays	**N**	**D**	**I**	.blr .eds	< 1 GB	None
Purified CoV proteins	E. coli-expressed and purified proteins (e.g. nsp12, nsp15)	**N**	**P**	**Biological material**	Not applicable	Not applicable	-80°C, storage managed with FreezerPro software
Protein analysis	Data from dynamic light scattering and size exclusion chromatography of CoV proteins and their complexes	**N**	**D**	**I**	.xls .png	< 1 GB	None
Cryo-EM images	Raw cryo-EM micrographs of CoV proteins	**N**	**D**	**I**	.mrc	<5TB	600TB large storage device available in-house
Cryo-EM maps and structures	Electron density map reconstructions and structures of CoV proteins	**N**	**D**	Map, Model	.mrc .pdb	<1GB	600TB large storage device available in-house
Small molecules	Obtained by chemical synthesis	N and E	P	Chemical material	Not applicable	Not applicable	1 mg to 10 g stored at 4 or -20 °C
Compound analysis	NMR	N and E	D and P	I and N	.fid .pdf	<100GB	None
Compound analysis	MS	N and E	D and P	I and N	.pdf	<100GB	none

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:

We use stocks of viruses, cell lines and plasmids generated earlier in the lab. The archiving and storage of these materials at -80°C or liquid N2 is managed with FreezerPro software. Some of the chemical molecules were synthesized earlier; these compounds were continuously stored at 4°C.

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, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.

- No

/

Will you process personal data? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S

number).

- 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

If we discover interesting anti-CoV activity for newly synthesized and structurally unique compounds, IP protection for these new inhibitors will be attained (via LRD).

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or 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

To validate the target and resistance profile of the new nsp12/nsp15 inhibitors, we will generate mutant CoVs by reverse genetics. The plasmid materials required for this work are subject to an MTA, which may need to be amended if we use the mutant viruses to attain IP. The MTA does not prohibit to do this, but requires approval from the providing institute.

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

/

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).

Team of Naesens and Stevaert: all experiments are organized in excel files. Each experiment file contains several worksheets with all details on the procedure and materials used, results and calculations. These .xls files are named by date and organized in folders by assay. The corresponding raw data (e.g. MTS, PCR, …) are also named by date, and archived in the same folder.
Team of Abhi Singh: lab notebooks are maintained to capture details of CoV protein expression and purification procedures. Data generated from protein purification steps are stored electronically, categorized with sample name and date. Details of cryo-EM experiments are organized in excel files.
Team of Elisabetta Groaz: notebooks contain information on research methods, protocols and experimental results (hard copy of images, HPLC profiles, Mass and NMR spectra including a reference code to trace back raw data in digital database).

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.

- No

We use a consistent and transparent structure to organize the datafiles.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- Shared network drive (J-drive)
- OneDrive (KU Leuven)

Team of Naesens and Stevaert: all team members involved in the project have access to the shared OneDrive folders, in which the files are organized as explained above. Via OneDrive for Business software, automated storage of all file versions is assured. For heavy files that are very demanding in terms of storage capacity (e.g. high-content imaging), the raw data files are stored on the Rega Institute's J-drive.
Team of Abhi Singh: Experimental files concerning protein expression, purification, and characterization are stored in Rega Institute's J-drive and shared across team members involved in the project. Large size files originating from cryo-EM experiments are stored on a 600TB storage server available locally in the lab and accessible to concerned team members. Files which are produced intermittently during cryo-EM data processing are stored on our GPU-enabled computers as well as on 10TB space allocated on VSC HPC clusters.
Team of Elisabetta Groaz: the time-stamped master copy of the data is kept on our research unit central storage facility. Copies can be made and kept on personal devices. OneDrive is used to share the data during the project.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution
- Other (specify below)

Team of Naesens and Stevaert: as explained above, most files are on our shared OneDrive folders, ensuring automatic backup service via the OneDrive provider. Also the Rega Institute's J-drive

is subject to KU Leuven ICTS-ensured backup service.
Team of Abhi Singh: files on the J-drive space are backup ensured by KU Leuven while large files on the 600TB server, as mentioned above, are periodically backed-up on 8TB USB expansion drives.
Team of Elisabetta Groaz: idem as Naesens and Stevaert

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

Team of Naesens and Stevaert: the total storage capacity for the OneDrive managed by L. Naesens is 2 TB. Currently, only 21% of this maximum is used. Though it is highly unlikely that the storage limit will be reached, L. Naesens can ask ICTS to increase her OneDrive limit to 5 TB.
Team of Abhi Singh: current space available on the J-drive and local storage server is sufficient.
Team of Elisabetta Groaz: KU Leuven covers 2TB on the OneDrive, this should be sufficient.

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

Team of Naesens and Stevaert: L. Naesens is the owner of the OneDrive folder and decides who gets personalized access to this folder. The access is limited to her direct team members who login via their KU Leuven account. As explained above, OneDrive assures access to older file versions in case an accidental modification would be made. In case a file would be accidentally deleted, OneDrive stores the file for 90 days in the recycle bin. Also access to the Rega Institute's J-drive is limited to Rega staff, who need to login via their u-account.
Team of Abhi Singh: files stored on our 600TB storage disk can only be accessed by designated persons. Similarly, access to our local data processing computers as well as HPC cluster is restricted to registered users authorized to do so. Access is granted by ICTS/VSC HPC only after explicit permission is given by the project/lab head.
Team of Elisabetta Groaz: idem as Naesens and Stevaert

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

Team of Naesens and Stevaert - use of OneDrive for Business software: the cost is financed centrally for KU Leuven staff. Use of Rega Institute J-drive: financed centrally for staff of the Rega Institute.
Team of Abhi Singh: use of Rega Institute J-drive: financed centrally for staff of the Rega Institute. Finances for other storage areas mentioned previously are covered by our other projects.
Team of Elisabetta Groaz: idem as Naesens and Stevaert

Data Preservation after the end of the Research Project

Which data will be retained for 10 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 will be preserved for 10 years according to KU Leuven RDM policy

All teams: all data are kept for at least 10 years (in reality much longer).

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

- Large Volume Storage (longterm for large volumes)
- Shared network drive (J-drive)

Naesens has a standard procedure to move archived data (= no longer used in an active manner) from the shared OneDrive folder to the Rega Institute's J-drive, for which a backup program is in place.
Team of Abhi Singh: Raw micrographs will be stored in individual back-up disks for long-term.

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

Use of Rega Institute J-drive: financed centrally for staff of the Rega Institute.

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.

- No (closed access)

/

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

Team of L. Naesens and A. Stevaert: access is limited to their team only. Hence, new team members acquire access to the data during or after execution of the project.
Team of Abhi Singh: only team members, current or future, assigned to the project will have access to the data.

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, intellectual property rights

Since we work with chemically unique molecules, uncontrolled open access to their biological (i.e. anti-coronavirus) activity is not warranted. Even if some of these molecules would not be IP-protected during the course of the project, there is a possibility that IP protection may be considered at later times.

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other (specify below)

Team of Abhi Singh: Electron density map reconstructions and determined protein structures originating from cryo-EM datasets will be deposited in EMDB and RCSB protein data bank (PDB) respectively at appropriate time (e.g., after journal publication). The deposited files will become freely accessible to the public after a release instruction is given by the depositor.

When will the data be made available?

- Upon publication of research results

Team of Abhi Singh: Electron density map reconstructions and determined protein structures originating from cryo-EM datasets will be deposited in EMDB and RCSB protein data bank (PDB) respectively at appropriate time (e.g., after journal publication). The deposited files will become freely accessible to the public after a release instruction is given by the depositor

Which data usage licenses are you going to provide?

If none, please explain why.

- Other (specify below)

None: the software to access and use PDB files is regulated by the RCSB protein data bank provider.

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

Team of Abhi Singh: Cryo-EM maps and coordinates of protein structures deposited in public databases (EMDB and RCSB-PDB) will be assigned unique DOI.

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

No costs apply for pdb files deposited in RSCB.

Responsibilities

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

Team of L. Naesens and A. Stevaert: data manager = L. Naesens
Team of A. Singh: A. Singh
Team of E. Groaz: Elisabetta Groaz and Nie Peng (post-doc)

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

Team of L. Naesens and A. Stevaert: data manager = L. Naesens
Team of A. Singh: A. Singh
Team of E. Groaz: Elisabetta Groaz and Nie Peng (post-doc)

Who will manage data preservation and sharing?

Team of L. Naesens and A. Stevaert: data manager = L. Naesens
Team of A. Singh: A. Singh
Team of E. Groaz: Elisabetta Groaz and Nie Peng (post-doc)

Who will update and implement this DMP?

The promoter (L. Naesens) and three co-promoters (A. Stevaert, E. Groaz and A. Singh)