DMP_12X2722N

Project Name FWO - DMP_12X2722N
Project Identifier u0072505
Grant Title 12X2722N
Principal Investigator / Researcher Chris Ulens/Mieke Nys
Project Data Contact mieke.nys@kuleuven.be
Institution KU Leuven

1. General Information

Name applicant

Mieke Nys

FWO Project Number & Title

12X2722N

A broad-spectrum antidote against venomous snakebites based on engineering of nicotinic acetylcholine receptors and derivatives.

Affiliation

- KU Leuven

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Type of data	Format	Volume	How created
Cryo-EM images	.jpeg	400 GB	Cryogenic electron microscopy
X-ray diffraction patterns	.cbf	400 GB	X-ray crystallography at synchrotron radiation sources
Chromatograms	.tif	10 GB	Fast protein liquid chromatography
Scanned images	.jpeg, .tif	10 GB	SDS-gels, agarose gels
Binding data	.xls, .jpeg .nta, .pdf	10 GB	Microscale thermophoresis, Biolayer interferometry

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

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, add the reference to the formal

approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- Yes

The data are part of a patentable invention and viable for commercial exploitation. The ultimate goal of this project is to deliver lead decoy receptors or combinations of lead decoy receptors with venom-lethality neutralizing efficacy in vivo against the most medically important snakes of distinct geographical distributions. Prior to dissemination, P-I Ulens will contact LRD so that a confidential invention disclosure document can be prepared and the result research can obtain protection. For future reuse of the data, we suggest a Creative Commons Attribution-NonCommercial-NoDerivs (CC-BY-NC-ND) licence.

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

All data generated in our wet labs will be documented in traditional written lab journals. Additionally, all conducted procedures are stored in seperate word files per experiment, containing protocols, specific remarks and description of the data generated. These files are stored on the server together with the resulting data. Each file is named according to date (year-month-day) and technique, for example 20220425_MST (MST for microscale thermophoresis).

Once procedures for experiments are established these are written down in protocols on the KULeuven wiki site (wiki.kuleuven.be/xtal), which are accesbile to lab members only. The wiki format guarantees the protocols are accessible from anywhere in the lab (and world) and in a universal format (not dependent on PC versus Mac for example).

Once certain aspects of the project are completed these results will be published in scientific journals. We always pay careful attention to describing experimental procedures in detail in the according section of the publication. This guarantees that the data can be reproduced later, both by others in our lab as well as other labs throughout the world.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- Yes

Our work concerns the 3-dimensional structure determination of proteins. A prerequisite for publication is that the atomic coordinates of these structures, as well as the experimental data are made publically available to the community via the Protein Data Bank (PDB). The metadata in this databank is in the CIF standard (Crystallographic Information File). At present our laboratory has 46 entries in the Protein Data Bank resulting from previous projects.

5. Data storage and backup during the FWO project

Where will the data be stored?

1. Data generated at synchrotron radiation sources (1) are copied to USB drives and then transported back to our lab, where we then copy them (2) on a local disk in our crystallography workstation (3) on the large volume storage capacity provided by the IT department of the biomedical sciences campus. In other words, 3 physically separate copies of the data exist and they are also stored in 3 physically different locations, one of which is a fire-proof safe in the office of Prof.Chris Ulens.

2. A 2nd set of data is generated at the cryo-electron microscopy facilities, these data will be backed up according to the same procedure mentioned in 1. This includes 3 separate copies in 3 separate locations.

3. A 3rd set of data concerns various equipment in the lab connected to PC's. Some of these computers are not connected to the university IT network, they are physically backed up by connecting USB disks. This procedure is carried out by Prof. Chris Ulens on a monthly basis after which the data are copied to network drives provided (at a cost) by the biomedical sciences IT department.

4. Notes generated in traditional lab journals (not electronic!). We still find this an efficient way to establish a record of everyday experiments in the wet lab as well as tracking down these data efficiently over many years of experiments. All of our lab journals are currently stored in a fireproof file cabinet. For this particular FWO fellowship, all experimental work conducted by Mieke Nys are also documented in seperate word files for each experiment. These files are stored on our central server and personal computer of Mieke Nys.

5. Since we collaborate with researchers from orther research units and groups, we will use OneDrive for active use of the data during the project.

How is backup of the data provided?

Most data will be stored on the university's central servers with automatic daily back-up procedures.

Backups of the large volume storage capacity are provided by IT department of the biomedical sciences campus.

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

Yes, we currently have several TB of data capacity for backup. This will be expanded depending on future needs.

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

Our current yearly expenses on IT storage and back-ups is 3000-4000 euro. We expect this amount will only increase in the future, as our volumes keep increasing and initiative like the data management plans only further increase IT needs (and time). The costs are divided proportionally among different project budgets (depending on volume of data generated in each project).

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

1. Our backups are stored on the university IT network and central servers. It is the university's responsability to guarantee security, for which we pay the according costs.

2. The lab books are stored in a fireproof file cabinet.

3. The local disks for storage of Cryo-EM data and X-ray crystallographic data are stored in a fireproof file cabinet.

4. Data on Mieke Nys personal computer is only accessible via Touch ID.

5. Data generated by various equipments connected to PC's are only accessible by Mieke Nys (staff number - password).

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

We will back up **all data mentioned in section 2**, no matter how old they are. This is a matter of scientific philosophy as even raw data from a long time ago could become important again at some point in the future. This matter has been widely discussed in the structural biology community. Inevitably, this means our data volumes will continue to increase in a cumulative manner and so will our associated IT costs.

Where will the data be archived (= stored for the longer term)?

1. All our data are long-term storaged on the large volume data storage bought from our IT department and central servers.

2. Three-dimensional structures will be deposited and long-termed stored in the Protein Data Bank (PDB).

3. All lab journals and local large volume disks are long-term stored in fireproof file cabinets.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

Our current yearly expenses are 4000-5000 euro. This amount will increase during the years to come, as IT continues to increase charges and our volumes continue to increase cumulatively. The costs are divided proportionally among different project budgets (depending on volume of data generated in each project).

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- Yes. Specify:

Prior to dissemination, P-I Ulens will contact LRD so that a confidential invention disclosure document can be prepared and the result research can obtain protection. Where patentable IP arises, and does not jeopardise the delivery, and time of delivery of health benefits, publications will be reviewed by LRD and publications will be paused to allow for a patent application to be filed.

Which data will be made available after the end of the project?

Our 3-dimensional structures of proteins are deposited into the Protein Data Bank. Additional data including biophysical characterization of decoy receptors designed during this fellowship will be published under an open access format, as requested by FWO.

Where/how will the data be made available for reuse?

- In an Open Access repository

See the Protein Data Bank and a recent deposition on Dryad, example see https://datadryad.org/stash/dataset/doi:10.5061/dryad.pv4097s

When will the data be made available?

- Upon publication of the research results

All scientific journals request that our entries into the Protein Data Bank are released upon formal acceptance of our manuscripts. Additionally, raw data will be submitted to journals upon request (source data file).

Who will be able to access the data and under what conditions?

The Protein Data Bank is publicly accesible. All of our publications are open access.

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

So far the Protein Data Bank does not charge users money, but this may change in the future. Publication in open access comes at a significant cost, certainly in high impact journals. We foresee a budget of around 3000 euro or more per publication.

8. Responsibilities

Who will be responsible for data documentation & metadata?

Prof. dr. Chris Ulens and dr. Mieke Nys.

Who will be responsible for data storage & back up during the project?

Prof. dr. Chris Ulens and dr. Mieke Nys.

Who will be responsible for ensuring data preservation and reuse ?

Prof. dr. Chris Ulens and dr. Mieke Nys.

Who bears the end responsibility for updating & implementing this DMP?
Prof. dr. Chris Ulens bears the end responsibility of updating & implementing this DMP.