DMP title

Project Name PHGDH-ASM, Phosphoglycerate dehydrogenase activators as novel antiseizure medications with dual mode of action

Grant Title research project C3/22/016

Principal Investigator / Researcher Karin Thevissen

Description The nature of the research is pre-clinical research, toward finality of patent application(s). Our preliminary data strongly support the idea that activating PHGDH results in anti-seizure and anti-inflammatory activity and represents a novel therapeutic approach against drug-resistant epilepsy. In this project, we will advance the hit NCEs that activate PHGDH further as novel anti-seizure drugs using zebrafish and mouse epilepsy models.

Institution KU Leuven

1. Data Description

What data will you collect or create? Fill out the table below and/or describe.

Type of data	Format	Volume	How created?
e.g observational, experimental, reference data,...	e.g. textual, numerical, multimedia	e.g 200MB, 1GB	Computer task, observations, blood sample, ...
(WP1) PHGDH enzyme activity	numerical	100MB	computer task
(WP1 and WP3) Observational toxicity - MTC (zebrafish larva)	numerical .xls	500MB	microscopy (visual observation)
(WP1 and WP3) Experimental locomotion data (zebrafish larvae)	numerical .xls	10GB	Daniovision / Ethovision (Noldus) with the program Ethovision,
(WP1 and WP3) Local field potential measurements - LFP equivallent to electroencephalograms (zebrafish larvae)	abf-files	500MB	abf files generated by WindEDR setup (local field potential setup)
(WP4) In vitro Absorption, distribution, metabolism, and excretion - toxicity and solubility data of compounds • Permeability • AMES test • hERG • Plasma protein binding • Microsomes testing CYP450 inhibition	Numerical – Word/Excel files	500MB	computer task
(WP2) Enzymatic profiling over other dehydrogenases	Numerical – excel files	500MB	computer task
(WP2) Compound structure – when shared with Pr. Voet's Lab	Sdf files	100MB	computer task
(WP2) Enzyme kinetics over truncated variants and orthologues	numerical - excel files	1GB	Tecan Spark , computer task
(WP2) BLI affinity data	numerical - excel files	1GB	Sartorius Octet R8, computer task
(WP2) Crystal diffraction and structures	tiff, h5, mtz and pdb files	1TB	computer task

Do you intend to reuse existing data?

-Not applicable for PHGDH enzyme activity data as such data do not exist yet (WP1)

-Not applicable for molecular biodiscovery (WP1 and WP3)

-Not applicable for biophysics and structural biology (WP2)
-Not applicable for CD3/Cistim (WP4)

Do you use personal data (i.e. all data possibly identifying an individual)?

- No

2. Documentation and Metadata

Describe the documentation that will be created for the data. This section deals with the way in which you will document how the dataset was created and subsequently processed.

1. PHGDH enzyme activity data consisting of time-response OD measurements. The following information will be noted: cell type; protocol; date; compound concentrations and DMSO backgrounds. The methodology and protocol will be described in detail in the lab book. An excel file per experiment will be saved. The name of the file will contain the date and cell type/compound tested. An overview txt file will contain date, experiment type and main outcome of all experiments.
2. Maximum tolerated dose (MTC) experiment: observations of toxicity are written down in the lab book and (if needed) in excel file. The following information is noted: date of experiment, compound and concentrations used, signs of precipitation, toxicity parameters.
3. Locomotor data is generated in the Daniovision system from Noldus with the program Ethovision, multiple parameters are being tracked at the same time and exported in excel file along with date of experiment, experimental conditions and results over time. Observations of the larvae tested in the locomotor experiments (post experiment) are written down in a lab book and the conclusions are written down in the excel file.
4. LFP data is generated with WindEDR (John Dempster, University of Strathclyde, UK) and exported to abf-files. Information about the positioning, signal strength and planning of the experiment is written down in the lab book. In the Dravet model, abf files are visually interpreted in Clampfit and the amount of epileptiform events and their duration are noted down in an excel file. In the EKP- model abf files are transformed to atf files and then to text files, a power spectral density analyses is preformed in MATLAB R2020b and the results are exported in excel files. All statistical analyses are preformed in GraphPad Prism 9.3.1.
5. Internal toxicity and solubility testing of new compounds are recorded in Excel data sheets and subsequently uploaded in the dotmatics platform (see question 4).

   ADME tests consists of a panel of different types of cell-based and biochemical assays performed routinely at different CROs. The collaboration with the CROs is described in question 4. The experimental procedures and the internal validation of the assays with reference compounds is shared in Excel sheets and reports. Validated data are subsequently uploaded in the dotmatics platform.
6. Protein purification: the following information wil be noted, protocol, data of purification, purity (SDS page image and SEC profile) concentration and buffer in word file.
7. Enzyme kinetics consists of time-dependent spectroscopical analysis performed on Tecan Spark. The following information will be noted : protocol, data, reagent concetrations. An excel file per experiment will be saved. The name of the file will contain the date and cell type/compound tested, and batch of protein purification. An overview txt file will contain date, experiment type and main outcome of all experiments.
8. Direct binding assays: timedependent spectroscopical analysis. The following information will be noted : protocol, data, reagent concetrations. A raw output file (CSV) and excel file per experiment will be saved. The name of the file will contain the date and cell type/compound tested, and batch of protein purification. An overview txt file will contain date, experiment type and main outcome of all experiments.
9. Structural biology: the procedure of crystal setup will be analyzed, all conditiosn yielding crystals will be documented using images (tiff). They will be picked up and shipped for crystal diffraction. The diffraction images (h5) format will be analyzed using phenix or ccp4 software and tunred into MTZ files and corresponding PDB files. Per crystal a foldering containg all data, and an txt file including crystal conditions, protein batch, cryoprotectant conditions will be saved together with a log detailing all diffraction analysis and solving steps.

Describe the metadata for the data. This section deals with metadata: information contained in your dataset about the research data.

1. PHGDH enzyme activity data consisting of time-response OD measurements. The following information is contained in the dataset (1 excel file/experiment): cell type; protocol; date; compound concentrations and DMSO backgrounds.
2. MTC data are observations of zebrafish larvae containing information about lethality, loss of posture, touch response, necosis, edema, realtive heart rate.
3. Locomotor data generates data about Distance moved, Velocity, Acceleration, Activity, Mobility and Activity state of the centerpoint of the larvae as it swims around in his well over time. Post locomotor toxicity observations are also noted in the excell file.
4. Local field potential data generates an EEG that is interpreted visually in the Dravet zebrafish model and processed in a power spectral density analyses in case of the EKP zebrafish model.
5. All ADME data received from CROs in excel format with unique references to the study number, the date of the experiment, the operator, the reagents used and the date of validation of the experiment.

   All Tox and solubility performed at CD3 are characterized by their own experiment identifier, the name of the operator, the date of the experiment, the protocol number and the time at which it has been countersigned.
6. PHGDH enzyme kinetics. The following information is contained in the dataset (1 excel file/experiment): cell type; protocol; date; protein batch, reagent concentrations, compound concentrations and DMSO backgrounds.
7. BLI affinity data. The following information is contained in the dataset (1 CSV and 1 excel file/experiment): cell type; protocol; date; protein batch, compound concentrations and sensor backgrounds.
8. Crystal structure: The following information is contained in the dataset (1 folder of h5 diffraction images, 1 image of the crystal before diffraction, 1 MTZ file with diffraction density data, 1 pdb file with the solved 3D structure, 1 txt file with refinment statistics, 1 txt logfile/experiment). log file contains: data and batch protein, reagent concentrations, diffraction data, synchrotron facility information and BAG access number, refinement protocols.

3. Ethical, Legal and Privacy Issues

Are there any ethical issues concerning the creation and/or use of the data?

We have ethical permission from KU Leuven to breed zebrafish: 000/(GS1/GS2) and to conduct experiment for neuroactive drug discovery on zebrafish larvae: 027/2019

Did you consider all issues about copyrights and IPR?

Yes - we have FTO

Are the collected data considered to be â€œdata containing personal informationâ€ and are all the requirements about the collection of these data met?

The data do not contain personal information

4. Data storage and Backup during Research

How and where will the data be stored during research?

- Centrally on storage facilities of the research unit

Which back-up procedures are in place?

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

Describe the data security procedures and who has access to the data.

The researchers involved have access to the data.

Data generated internally by CD3/Cistim are stored in a dedicated project folder on an internal server, maintained and secured by KU Leuven ICTS. Access to the data is limited to Cistim employees. Next to this, chemical structures and activity data of compounds are uploaded into the Dotmatics repository hosted on an internal KU Leuven maintained server following standardized procedures. Access to Dotmatics is also limited to internal users. Scientists at Cistim are trained on data security processes. Data can only be stored on the central server or dotmatics and not on PC’s or non-secured data carriers. Compound structures are never shared via email or other none secured data carriers. If confidential data need to be shared with partners a secured Sharepoint is established.

Some activities are outsourced to CRO’s. In most cases a master service agreement is or will be established and if not, terms and conditions will be negotiated with support of KU Leuven LRD to protect data and IP. In any case all shared data, analysis results, analytical reports and products will remain our sole property. Through the contract the CRO is also obliged to prepare adequate, accurate and secured records of their work done in the performance of the services. They are also required to report all data, information, discoveries, observations and the like. If needed, a representative is allowed to inspect the CRO’s facility to audit correct execution of the contract. The contracts will also contain agreements regarding data retention and destruction by the CRO when all services are finished.

Data generated by LBDM (Voet lab) is stored in a double backed-up NAS of 64TB with 4TB reserved for this project.

Data generated by Laboratory for Molecular Biodiscovery is stored in lab books and on local computers which are kept in the laboratory and only accessible by approved personnel with a badge. Backup files are kept on University servers. Access to these files can only be granted by the PI of the research group (Prof. Peter de Witte) via the iMANAGE4.0 system.

5. Data selection and Preservation after Research

What is the long-term preservation plan for these dataset(s)?

1. 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.

2. Data generated by LBDM (Voet lab) will be transferred to the project leader for centralized backup upon conclusion of the project.

Data Selection: Which data will have long time value for the research and will be preserved?

PHGDH enzyme activity data

Maximum tolerated concentration of compounds, locomotor activity and local field potential assessments on zebrafish will be preserved at long term as these provide in vivo efficacy of compounds in disease models.

All chemical structures and activity data of chemical compounds might have value for future projects. Standardized upload and storage of these data in the Dotmatics repository ascertains easy accessibility and re-usability of data

X-ray structures of PHGDH agonists crystallized in PHGDH have great values to understand the binding mode of compounds and will be preserved at long term.

6. Data Sharing

Are there any restrictions for sharing the data?

Within the consortium data will be shared. However, some limitations are implemented. For instance all data on compound activity will be shared with all members of the consortium but chemical structures will only be shared when this is required.

Data sharing with CROs is also limited and tightly monitored. When approaching a CRO an NDA will be established with support of KU Leuven LRD before confidential data are shared. Subsequently, data sharing during the services by the CRO will be described in a contract as described above.

There will be restrictions for data disclosure outside of the consortium as it may contain IP-sensitive information. IP management will be conducted as follows throughout the entire project:

The internal IP management will be the responsibility of the PI. Patent applications will be made for relevant innovative procedures/products and the costs will be shared according to ownership percentage that shall be subject to negotiation on a pro rata basis. The consortium policy will be to patent all innovative results which could be exploited without preventing publication and rapid circulation of information between the partners. Nevertheless, to avoid publication of IPR sensitive results, the PI in collaboration with LRD will screen all anticipated publications including abstracts for presentations at scientific meetings prior to their submission. All partners will be responsible for transfer of anticipated publications of whatever nature to the PI for IP screening and approval of release at least one month before intended publication. The PI will discuss within a three week period and decide on whether the publication can be released or whether postponement of publication with two months to allow for patent protection is necessary.

If there are no restrictions, which mechanisms will be in place to assure that the data

are discoverable, accessible and intelligible?
not applicable - there are restrictions as described above

How will you share the data?

Data will be shared via a Sharepoint put in place by the KULeuven ICTS.

With whom will the data be shared?

- Within the university only

ADME, Tox, solubility, dehydrogenase enzymatic activity, locomotor and local field potential data will be shared with all teams members. Chemical structures required for X-ray structures generation will be shared with Pr Voet's lab, using a password protected document.

7. Responsabilities and Resources

Who is responsible for Data Management during the project? This will be the person who might receive questions on the data management aspects of the research project.
The PI, Karin Thevissen, bears the end responsibility of updating & implementing this DMP.

Which additional resources are needed for the execution of the Data Management Plan?
none

Did you read the KU Leuven Data Management Policy? (find the link to the policy in the guidance).

- Yes