# KLEOPATRA – Design and implementation of effective cOmbination of Phages and Antibiotics for improved TheRApy protocols against KLEbsiella pneumoniae

*A Data Management Plan created using DMPonline.be*

**Creators:** Jeroen Wagemans, Rob Lavigne

**Affiliation:** KU Leuven (KUL)

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

**Template:** FWO DMP (Flemish Standard DMP)

**Principal Investigator:** Rob Lavigne

**Project Administrator:** Jeroen Wagemans

**Grant number / URL:** S008722N

**ID:** 202180

**Start date:** 01-06-2023

**End date:** 31-05-2026

**Project abstract:**

Klebsiella pneumoniae is a critical pathogen which displays pan-resistance against prevalent antibiotics. Moreover, this bacterium contains hyper virulence traits-associated genes encoding capsular polysaccharides (K-antigens or capsules) that protect it against the human innate immune system as well as other harsh environments. The KLEOPATRA consortium focuses on the use of bacterial viruses (phages) and their components as sustainable combinations with antibiotics. As natural predators of bacteria, strictly lytic phages have been identified that specifically recognize and target the K-antigen as a receptor for phage infection.

In this proposal we introduce the concept of K-sensitization, which utilizes capsule-specific (CS) phages and their capsule-degrading enzymes (depolymerases) to (1) resensitize bacteria towards antibiotics treatment, (2) sensitize the pathogens to the host complement system and susceptible to neutralization by phagocytic cells and (3) make them susceptible to capsule-independent (CI) phages.

The envisioned K-sensitization preclinical proof-of-concept aims to develop a synergistic phage therapy/antibiotic treatment protocol against encapsulated pathogens that helps preserve the commensal microbiome diversity. The approach is supported by an environmental genotype analysis on the prevalent clinical and environmental isolates. Furthermore, the design of phage/depolymerase/AB cocktails will be driven by tailored diagnostics tools, a unique phage production approach of dedicated phage banks, as well as data-driven mathematical models for optimal synergy. The latter will be developed at KU Leuven.

**Last modified:** 10-10-2023

# KLEOPATRA – Design and implementation of effective cOmbination of Phages and Antibiotics for improved TheRApy protocols against KLEbsiella pneumoniae
# FWO DMP (Flemish Standard DMP)

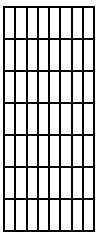
### 1. Research Data Summary

**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 |
| *Klebsiella*strains | Collection of strains from the different partners of the consortium | ☐ Generate new data<br>☒ Reuse existing data | ☐ Digital<br>☒ Physical | ☐ Audiovisual<br>☐ Images<br>☐ Sound<br>☐ Numerical<br>☐ Textual<br>☐ Model<br>☐ Software<br>☐ Other: | | ☐ < 1 GB<br>☐ < 100 GB<br>☐ < 1 TB<br>☐ < 5 TB<br>☐ > 5 TB<br>☐ NA | Maximum 2 boxes stored at -80°C |
| *Klebsiella* phages | Collection of phages from the different partners of the consortium | ☒ Generate new data<br>☒ Reuse existing data | ☐ Digital<br>☒ Physical | | | | Maximum 2 boxes stored at 4°C |
| Phage depolymerases | Collection of proteins | ☒ Generate new data<br>☒ Reuse existing data | ☐ Digital<br>☒ Physical | | | | Maximum 1 box stored at -80°C |
| Phage infectivity and depolymerase activity data | Data on the infectivity of the phage collection and the activity of the depolymerases on the strain collection | ☒ Generate new data | ☒ Digital<br>☐ Physical | ☐ Audiovisual<br>☒ Images<br>☐ Sound<br>☒ Numerical<br>☒ Textual<br>☐ Model<br>☐ Software<br>☐ Other: | .jpg<br>.tiff<br>.png<br>.xls<br>.csv | ☒ < 1 GB<br>☐ < 100 GB<br>☐ < 1 TB<br>☐ < 5 TB<br>☐ > 5 TB<br>☐ NA | |
| Bacterial and phage sequencing data | Raw and processed sequencing data of phages and bacteria | ☒ Generate new data | ☒ Digital<br>☐ Physical | ☐ Audiovisual<br>☐ Images<br>☐ Sound<br>☒ Numerical<br>☐ Textual<br>☐ Model<br>☐ Software<br>☐ Other: | .fastq<br>.fast5<br>.fasta<br>.gb | ☐ < 1 GB<br>☒ < 100 GB<br>☐ < 1 TB<br>☐ < 5 TB<br>☐ > 5 TB<br>☐ NA | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Methods, protocols, lab books | Associated data | ☒ Generate new data<br>☒ Reuse existing data | ☒ Digital<br>☐ Physical | ☐ Audiovisual<br>☐ Images<br>☐ Sound<br>☐ Numerical<br>☒ Textual<br>☐ Model<br>☐ Software<br>☐ Other: | .docx<br>.txt<br>.pdf | ☒ < 1 GB<br>☐ < 100 GB<br>☐ < 1 TB<br>☐ < 5 TB<br>☐ > 5 TB<br>☐ NA | |
| Communicative data | Scientific reports, presentations, papers, theses | ☒ Generate new data<br>☐ Reuse existing data | ☒ Digital<br>☐ Physical | ☐ Audiovisual<br>☐ Images<br>☐ Sound<br>☐ Numerical<br>☒ Textual<br>☐ Model<br>☐ Software<br>☐ Other: | .docx<br>.txt<br>.pdf<br>.pptx | ☒ < 1 GB<br>☐ < 100 GB<br>☐ < 1 TB<br>☐ < 5 TB<br>☐ > 5 TB<br>☐ NA | |
| Models/code | Bioinformatic pipelines to predict infectivity of phage on strains | ☒ Generate new data<br>☐ Reuse existing data | ☒ Digital<br>☐ Physical | ☐ Audiovisual<br>☐ Images<br>☐ Sound<br>☐ Numerical<br>☒ Textual<br>☒ Model<br>☐ Software<br>☐ Other: | .txt | ☒ < 1 GB<br>☐ < 100 GB<br>☐ < 1 TB<br>☐ < 5 TB<br>☐ > 5 TB<br>☐ NA | |

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



**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 will reuse sequencing data from Klebsiella strains and/or phages freely available from NCBI or other public repositories.
We will also reuse strains and phages/depolymerases from existing collections within the different partner labs.

**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

For the KU Leuven part, there are no ethical issues.

**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 direct expected outcomes of KU Leuven's part of this project will be (i) a biobank of both capsule-specific and -independent phages actively targeting Klebsiella strains isolated from one health settings; (ii) the development of new mathematical models for optimization of antimicrobial combinations in treatment protocols (in silico phagogram) and companion diagnostics for K. pneumoniae, as proof of concept for encapsulated pathogens

Economic valorization: (i) The already characterized phages which will be used to optimize treatment protocols within the KLEOPATRA project are partly owned by the KU Leuven partner (IP ownership agreements in place, thereby facilitating a smooth valorisation in Flanders and beyond). These phages can be directly licensed to companies active in phage therapy (e.g. the Belgian Vésale Biosciences) for an estimated 1-10% of the royalties of the net sales. (ii) From a societal perspective, the phage bank can directly be valorized within UZ Leuven which apply phage therapy as standard of care for patients with difficult-to-treat musculoskeletal infections. Currently some patients cannot be treated because there is no Klebsiella phages available with a passport according to the Belgian monograph (see also societal impact). (iii) The development of an in silico phagogram and mathematical model to define the best phage/antibiotic combinations for treatment of a specific infection will also directly lead to IP generation, which can be valorized by patenting.

**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).**

All (meta)data and how it was generated and processed is tracked using a combination of digital lab books, which are kept on university-secured network drives and cloud-based systems (such as Google Drive). Both lab books and data are time-stamped, allowing to easily trace back (experimental) details of the corresponding data. Files will follow a standard naming format (e.g. yyyymmdd_reasearcherinitials_experimentdescriptor) to improve traceability of data and make them easily findable.

Nearly all digital data formats are commonly used in the field of research and can be visualised/analysed with common software suites or online tools, making them interoperable.

**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

Where possible, we will strive to include how data was processed into the file or folder of the respective processed data (as .txt file).

Data published in or as supplement of open-access peer-reviewed publications will be conform to the standards of the publisher.

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

**Where will the data be stored?**

☒ Shared network drive (J-drive)
☐ Personal network drive (I-drive)
☒ OneDrive (KU Leuven)
☐ Sharepoint online
☐ Sharepoint on-premis
☒ Large Volume Storage
☐ Digital Vault
☐ Other:

**How will the data be backed up?**

☒ Standard back-up provided by KU Leuven ICTS for my storage solution
☐ Personal back-ups I make (specify)
☐ Other (specify)

**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

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

Data (both raw and processed) will be stored on the storage facilities of the respective university from each partner and a secure cloud service (only for processed data), to ensure accessibility by the different partners.
In case data is linked to publications, it will be also made available on public databases which have their own storage facilities (e.g. Sequence Read Archive, Genbank of NCBI).
The university facilities have rigorous data back-up plans in place to avoid data loss as much as possible. These include frequent back-ups, redundancy in storage and spatial separation of data storage sites.
Biological data will be stored at 4°C fridges and -80°C freezers with limited access. A copy of the bacterial, phage and protein database will be kept at each consortium partner. Freezers with biological data are located in the labs which have restricted access for unauthorized personnel (e.g. by means of a badge-system).

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

We will stay within the limits of what is currently provided by our university for digital data storage.
For physical data storage, we have the required materials and space available (fridge, freezers)

## 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 data will be preserved for 10 years according to KU Leuven RDM policy
☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans
☐ Certain data cannot be kept for 10 years (explain)

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

☒ KU Leuven RDR
☒ Large Volume Storage (longterm for large volumes)
☒ Shared network drive (J-drive)
☐ Other (specifiy):

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

We will stay within the limits of what is currently provided by our university for digital data storage.
For physical data storage, we have the required materials and space available (fridge, freezers)

## 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.**

- Yes, in an Open Access repository

All relevant data of the project and associated data such as lab books will be archived for a minimum of 10 years. After this time, irrelevant versions of digital DNA sequences, digital lab notebooks and experimental datasets will be scrutinized for prolonged storage or disposal. All relevant data, however, will be preserved to keep them reusable.
After the IP strategy has been sorted out, relevant data will be published in peer-reviewed journals, making them easily discoverable and identifiable through DOI codes. Specifically, the FAIR policy will be also implemented through Open Access publications. Bioinformatic pipelines produced in KLEOPATRA will be made freely available at GitHub.

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

NA

**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

Since some of the data can lead to new IP, the data will be kept confidential within the consortium until the valorisation of the data is clear for everyone. The consortium agreement will describe how protection and valorisation of project results will be achieved. IP and copyrights will be handled according to this agreement.

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

☐ KU Leuven RDR
☒ Other data repository (specify)
☐ Other (specify) NCBI, Github

**When will the data be made available?**

☒ Upon publication of research results
☐ Specific date (specify)
☐ Other (specify)

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

☒ CC-BY 4.0 (data)
☐ Data Transfer Agreement (restricted data)
☐ MIT licence (code)
☐ GNU GPL-3.0 (code)
☐ Other (specify)

**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

☒ Yes, a DOI will be added upon deposit in a data repository
☐ My dataset already has a PID
☐ No

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

Open access publication costs are foreseen within the project budget.

## 6. Responsibilities

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

Cédric Lood, Jeroen Wagemans, Sabrina Green

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

Cédric Lood, Jeroen Wagemans, Sabrina Green

**Who will manage data preservation and sharing?**

Rob Lavigne, Jeroen Wagemans

**Who will update and implement this DMP?**

Rob Lavigne, Jeroen Wagemans

# KLEOPATRA – Design and implementation of effective cOmbination of Phages and Antibiotics for improved TheRApy protocols against KLEbsiella pneumoniae
# GDPR

### GDPR

#### Have you registered personal data processing activities for this project?

- No

# KLEOPATRA – Design and implementation of effective cOmbination of Phages and Antibiotics for improved TheRApy protocols against KLEbsiella pneumoniae
# DPIA

**DPIA**

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

- Not applicable