EEN GEÏNTEGREERDE AANPAK OM STAPHYLOCOCCUS AUREUS BIOFILMS TE BESTRIJDEN DOOR MIDDEL VAN DESIGNER ENZYMES

A Data Management Plan created using DMPonline.be

Creators: Vincent De Maesschalck, Jeroen Wagemans, Rob Lavigne

Affiliation: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Grant number / URL: C3/23/005

ID: 203768

Start date: 01-09-2023

End date: 31-08-2025

Project abstract:
De alarmerende toename van antibioticaresistentie bij de bacteriële ziekteverwekker *Staphylococcus aureus* (Sa) vormt een kritieke uitdaging in de humane geneeskunde. Bovendien komt Sa veel voor in abcessen en fibrine-geassocieerde biofilms. Deze biofilms beperken de toegankelijkheid van antibacteriële componenten en veroorzaken chronische infecties. De aanvragers ontwikkelden en patenteerden een enzyme-gebaseerde benadering om Sa te bestrijden (ZL922003), gebaseerd op een 'proof-of-concept' voor abcessen geassocieerd met prothetische gewrichtsinfecties. Dit voorstel zal de voorlopige indiening van het patent versterken door (1) de designer-eiwitten uit te breiden naar een breder gamma aan enzymen en (2) de POC uit te breiden naar andere Sa-biofilm-geassocieerde klinische omgevingen, waaronder subcutane toepassingen (Hidradenitis suppurativa), chronische rhinosinusitis, evenals katheter-geassocieerde biofilms. Vanuit een technisch perspectief zullen we onze kennis en expertise uitbreiden in (a) recombinante eiwitproductie in nieuwe fermentatiestammen (uitbreiding op een 2e patent-ZL922073) (b) de ontwikkeling van een op bioluminescentie gebaseerd, *in vivo* model in *Galleria mellonella* larven voor implantaat-geassocieerde biofilms en (c) valideren van ons vermogen om deze eiwitten efficiënt te coaten op titaniumoppervlakken voor lokale afgifte en preventie van biofilmvorming.

Last modified: 21-12-2023

EEN GEÏNTEGREERDE AANPAK OM STAPHYLOCOCCUS AUREUS BIOFILMS TE BESTRIJDEN DOOR MIDDEL VAN DESIGNER ENZYMES

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 (N) or reuse (E)	Digital (D) or Physical (P) data	Data Type Indicate: Audiovisual Images Sound	File format	Data volume Indicate: <1GB <100G	Physical volume

				Numerical Textual Model SOftware Other (specify)		B <1TB <5TB >5TB NA	
Staphylococcal isolates	Strain collection (Laboratory of Gene Technology)	**E**	**P**				Maximum 1 box stored at -80°C
Expression strains	Strains for the recombinant expression of proteins	**E**	**P**				Maximum 1 box stored at -80°C
Staphylococcal domains of interest, encoded on an expression plasmid	Cloned expression constructs to make fusion proteins/the fusion proteins itself	**N**	**P**				Maximum 4 boxes stored at -20°C
Protein-coated implant materials	Metallic biomaterials with and without proteins coated to the surface	**N**	**P**				Maximum 1 box stored at 4°C
Data on genome sequencing to mine for novel *Staphylococcal* domains	Data generated by next-generation sequencing technologies, necessary for *de novo* genome assembly and annotation	**N**	**D**	Other	.fastq .fasta .fna .gb	<100GB	
Data on the deposition of proteins on metallic biomaterials	Data generated by the Braem lab	**N**	**D**	**I** **N** **T**	.jpg .tiff .png .xls .csv	<1GB	
Peptide data	Data on the purification of recombinant proteins, what binds to metallics	**N**	**D**	**I** **N** **T**	.jpg .tiff .png .xls .csv	<1GB	
Data on the antibacterial character of engineered proteins	Associated data	**N**	**D**	**N** **T**	.xls .csv	<1GB	
Animal experiment data	Data obtained from the *in vivo* trial: histological data, BLI	**N**	**D**	**I** **N** **T**	.jpg .tiff .png .xls .csv .seq	<100 GB	

	data,...					
Methods, protocols, lab books	Associated data	**N**	**D**	**T**	.docx .txt .pdf	<1GB
Communicative data	Scientific reports, presentations, papers, theses	**N**	**D**	**T**	.docx .txt .pdf .pptx	<1GB

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 *S. aureus* strains previously isolated and characterized at KU Leuven, as well as expression vectors encoding engineered proteins developed at the Laboratory of Gene Technology.

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.

- Yes, animal data (Provide ECD reference number below)

During the second year of this project, an animal trial is scheduled, the ECD reference number will be provided when the application for this trial is filed.

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

Concepts being developed in this project as well as sequences of the engineered proteins and all data from both *in vitro* and *in vivo* experiments are strictly confidential and therefore will be stored at internal network drives of KU Leuven. Only people with the right credentials will have access to this folder.

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

A research collaboration and exploitation agreement for joint IP was established with AO Research Institute by LRD. This projects fits in the valorization path of the ongoing collaboration, as well as in the scope of creating new IP.

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

Reusage of previously generated data is covered in the research collaboration and exploitation agreement, established by LRD.

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. 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.

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.

- Yes

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. Cross-references in the digital lab book to raw data files will improve traceability of the raw data files.

DATA STORAGE & BACK-UP DURING THE RESEARCH PROJECT

Where will the data be stored?

- Shared network drive (J-drive)
- Large Volume Storage

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

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

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 network drives of KU Leuven. The university has 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.
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).
Biological data will be stored at 4°C fridges and -80°C freezers with limited access. 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).

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

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)

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).

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.

- Yes, as embargoed data (temporary restriction)

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.

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

Confidential data will be accessible for partners upon signing an NDA or for internal members continuing this research line.

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 some of the data can lead to new IP, the data will be kept confidential within the consortium until the valorization of the data is clear for everyone.

Where will the data be made available?

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

- KU Leuven RDR (Research Data Repository)

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

- CC-BY 4.0 (data)

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

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

Open access publication costs are foreseen within the project budget.

RESPONSIBILITIES

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

Vincent De Maesschalck

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

Vincent De Maesschalck

Who will manage data preservation and sharing?

Rob Lavigne, Jeroen Wagemans

Who will update and implement this DMP?

Vincent De Maesschalck