Plan Overview

A Data Management Plan created using DMPonline.be

Title: Development of broad spectrum antibiotics with a new mode of action through establishing new tools for optimal sequence selection in vitro and in silico design of antisense oligonucleotides.

Creator: Sten Reynders

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Project abstract:

Rapid development of antibiotic-resistance urges the search for new antibiotics. For many decades, the bacterial ribosome was successfully targeted by small molecule antibiotics, but many of them lost potential due to the development of antibiotic resistance. This project aims to target the ribosomal RNA (rRNA) of bacteria using oligonucleotides (ON). To date, synthetic ON are used to modulate the expression of specific genes by interaction with mRNA in biotechnology and in the clinic. I will focus on the highly conserved 3'-end of the 16S rRNA, essential in translation initiation within the prokaryotic world. A biotechnological approach will be used to perform a wet lab selection for ribose-based ONs to inhibit protein synthesis through interaction with rRNA in the bacterial ribosome. In parallel, an in silico approach will be established to design morpholino-based ONs with specific stereochemistry and sequence to target the selected RNA fragment. Phosphorodiamidate morpholino ONs (PMOs) are used in a mixture with millions of stereo-isomers as no stereospecific synthesis is available. Recently my host lab prepare synthons for isopure synthesis of methyl-phosphorodiamidate morpholino ONs (Me-PMOs). Stereopure Me-PMOs allow to determine the effect of backbone stereochemistry on biophysical characterisation and biological activity. Selected ON constructs will be screened for inhibition of ribosomal translation and antibacterial activity on clinical relevant pathogens.

ID: 210752

Start date: 01-11-2024

End date: 31-10-2028

Last modified: 06-04-2025

Development of broad spectrum antibiotics with a new mode of action through establishing new tools for optimal sequence selection in vitro and in silico design of antisense oligonucleotides.

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

WP1: Develop a software tool for *in silico* design of stereopure morpholino-based oligonculeotides	**WP1.1: Generate PMO structures *in silico* (all Rp, all Sp and mixed Rp/Sp sequences)**
	1. Quantum Mechanical (QM) data
	  ○ conformational sampling
	2. Molecular Dynamics (MD) data
	  ○ built models
	  ○ pdbs and coordinates of performed simulation and associated analyses
		• Generate new data *(i.e. different linkers)*
		• Reuse existing data *(i.e. morpholino-based nucleosides)*
			• Digital	• Experimental
				• Simulation data
					1. QM:
					  ○ .dat, .txt, .sh, .densities, .engrad, .gbw, .inp, .opt, .xyz, .pdb, .out, .py, .pdf, .prep
					2. MD:
					  ○ .binp, .tinp, .pdb, fcrmod, leaprc, .prmtop, .inpcrd, .log, .rst, .wc, .mdcrd, .mdinfo, .ncrst, .inp, .out, .in, .dat, .sh, .py, .agr, .csv, .png, .pdf
						• <10TB	NA

WP1: Develop a software tool for *in silico* design of stereopure morpholino-based oligonculeotides	**WP1.2: calculate thermodynamic parameters of PMO:RNA and PMO:PMO duplexes**
	1. MD-data:
	  ○ average trimer structures
	2. QM-data:
	  ○ estimated free energies
		• Generate new data	• Digital	• Experimental	• .pdb, .xyz, .inp, .out, .kb
					(to be updated when these calculations are actually performed)
						• <10TB	NA

WP1: Develop a software tool for *in silico* design of stereopure morpholino-based oligonculeotides	**WP1.3: extend PFizer RNAi Enumeration and Design tool (PFRED) for morpholino oligonucleotides** • PFRED software extended with the morpholino data	• Generate new data (regarding the software for implementing the PMO data into the software tool) • Reuse existing data (using the initial software code)	• Digital	• Software	• .py	• <10TB	NA

WP2: Blocking ribosomal translation in a cell-free system	**WP2.1: miRNA 'sponge' to select the sequence of ribose-based ASO** • DNA-templates • Dumbbell-shaped genetic minimal vectors • RNA • 'sponge' RNA • plasmids • IVTT-reactionmixtures • Urea/SDS-page gels + associated analyses • absorbance/fluorescence spectra • Excell spread sheets • R and/or Python scripts (for performing statistical analyses/ making graphs/ visualisation of the data)	• Generate new data	• Digital • Physical	• Experimental	• .xlsx • .png • .tif • .R • .py • .gel	• < 10 GB	Stored at -20 °C during the project and at -80 °C for longterm storage: • DNA-templates • Dumbbell-shaped genetic minimal vectors • RNA • 'sponge' RNA • plasmids • IVTT-reactionmixtures • E. coli
WP2: Blocking ribosomal translation in a cell-free system	**WP2.2: evaluation of selected ASO** • Sequences of ribose-based nucleotides • absorbance/fluorescence spectra • Excell spread sheets • R and/or Python scripts (for performing statistical analyses/ making graphs/ visualisation of the data)	• Generate new data	• Digital • Physical	• Experimental	• .xlsx • .png • .tif • .R • .py • .gel	• < 10 GB	Stored at -20 °C during the project and at -80 °C for longterm storage: • Ribose-based and morpholino-based oligonucleotides • E. coli
WP3: Screening of peptide conjugated ASO *(in collaboration with Prof. E. André, Laboratory of Clinical Microbiology)*	• Synthesized CPP-conjugated PMO-based and ribose based oligonucleotides • plates • data regarding toxicity and antibacterial activity of screened sequences	• Generate new data	• Digital • Physical	• Experimental	• .xlsx • .png • .R • .py	• < 1GB	Stored at -20 °C during the project and at -80 °C for longterm storage: • Synthesized CPP-conjugated PMO-based and ribose based oligonucleotides and conjugates

WP4: Biophyscial characterization of a Me-PMO	**WP4.1: Solution NMR structure of a Me-PMO in chimeric RNA hairpins**
	• NMR spectra (2D-NOESY, 2D-COSY, 2D-TOCSY, 2D-[1H,13C]HSQC, 2D-[1H,13C]HMBC, 2D-[1H,31P] INVCOSY
	• Me-PMO oligonucleotides
		• Generate new data	• Digital
			• Physical
				• Experimental
				  ◦ Raw data (Free Induction Decay, Acquisition Parameters and Processing Parameters)
				  ◦ Export formats for analysis and data sharing
					• ser
					• acqu, acqus
					• proc, procs
					• pdata folder
					• procpar
					• .jdx
					• .txt
					• .csv
					• .xml
					• .cara
						• <1GB	At least one aliquot of the following biological/chemical material need to be stored during/after the project in order to be able to do a retake of the experiment when requested:
							• ON1 (i.e. chimeric Me-PMO*::RNA hairpin, containing one Me-PMO residue)
							• ON2 (i.e. chimeric Me-PMO*::RNA hairpin, containing one Me-PMO basepair)
							*only (R)Me-PMO and only (S)Me-PMO
							Printed spectra + assignment list.

WP4: Biophyscial characterization of a Me-PMO	**WP4.2: Solution NMR structure of Me-PMO in a homoduplex and heteroduplex with RNA**
	• NMR spectra (2D-NOESY, 2D-COSY, 2D-TOCSY, 2D-[1H,13C]HSQC, 2D-[1H,13C]HMBC, 2D-[1H,31P] INVCOSY
	• Me-PMO oligonucleotides
		• Generate new data	• Digital
			• Physical
				• Experimental
				  ◦ Raw data (Free Induction Decay, Acquisition Parameters and Processing Parameters)
				  ◦ Export formats for analysis and data sharing
					• ser
					• acqu, acqus
					• proc, procs
					• .jdx
					• .txt
					• .csv
					• .xml
					• .cara
						• <1GB	At least one aliquot of the following biological/chemical material need to be stored during/after the project in order to be able to do a retake of the experiment when requested:
							• ON3::ON4 (i.e. chimeric Me-PMO*::RNA heteroduplex)
							• ON5 (i.e. chimeric Me-PMO* homoduplex)
							*only (R)Me-PMO and only (S)Me-PMO
							
							Printed spectra + assignment list.

WP4: Biophyscial characterization of a Me-PMO	**WP4.3: Thermal stability** • Melting temperatures	• Generate new data	• Digital • Physical	• Experimental	• .xlsx	• <100MB	At least one aliquot of the above already mentioned oligonucleotides need to be stored during/after the project in order to be able to do a retake of the melting temperature experiment.

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:

WP1.1: Generate PMO structures *in silico* (all Rp, all Sp and mixed Rp/Sp sequences)
Data of the morpholino-based nucleosides will be reused from the following paper: [DOI: https://doi.org/10.1093/nar/gkae135]:
Rihon, J., Mattelaer, C. A., Montalvão, R. W., Froeyen, M., Pinheiro, V. B., & Lescrinier, E. (2024). Structural insights into the morpholino nucleic acid/RNA duplex using the new XNA builder Ducque in a molecular modeling pipeline. Nucleic Acids Research, 52(6), 2836-2847.

WP1.3: extend PFizer RNAi Enumeration and Design tool (PFRED) for morpholino oligonucleotides
PFRED software
GitHub: https://github.com/pfred
Paper [DOI: https://doi.org/10.1371/journal.pone.0238753]:
Sciabola, S., Xi, H., Cruz, D., Cao, Q., Lawrence, C., Zhang, T., ... & Stanton, R. V. (2021). PFRED: A computational platform for siRNA and antisense oligonucleotides design. *PLoS One*, *16*(1), e0238753.

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

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 project has the potential to provide first-in-class products that tackle bacterial protein synthesis

with a mode-of action that remains unexplored to date. Once we have generated proof of antibacterial
activity with a selected sequence of nucleobases, we will file for IP protection (in collaboration of LRD, the tech transfer office of KU Leuven) on the oligonucleotide sequence for targeting ribosomal RNA. This is possible for existing ASO chemistries (IP protection on secondary medical use) as well as for the new chemistry proposed in the project (primary IP protection). This will mainly concern data obtained from WP1-3.

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).

WP1: Develop a software tool for *in silico* design of stereopure morpholino-based oligonculeotides

README.txt files will be added by the different generated datafiles of the QM calculations and MD simulations. This can be accompanied by a flowchart, which describes the different steps for both QM and MD on how to go from calculations to results and can also serve as clarification for the organization of the folder structures. For the various QM calculations, MD simulations and analyses performed, a cloud based e-notebook (Benchling) is kept. A lab notebook (on paper) is also been kept, which will be scanned regulary so that it can be added to the data as a digital copy. The generated python, shell and R scripts and the software be developed will be commented. README.txt will also be added were necessary. All the generated code (scripts for analysis and the software been developed) will be hosted on GitHub.

WP2: Blocking ribosomal translation in a cell-free system

An electronic lab notebook is maintained on Benchling, although a lab notebook on paper will still be used during the experiments, the latter will be scanned regulary so that it can be stored as a digital copy together with the experimental data. This experimental data will include README.txt files. Generated scripts (code) for analyzing the data will be commented appropriately and hosted on GitHub.

WP3: Screening of peptide conjugated ASO *(in collaboration with Prof. E. André, Laboratory of Clinical Microbiology)*

README.txt files will be included by the different generated datafiles. Generated scripts (code) for analyzing the data will be commented appropriately and be hosted on GitHub.

WP4: Biophyscial characterization of a Me-PMO

README.txt files will be included where necessary to give the reader a better understanding of the structure of the folder that is generated after each NMR experiment. The software (TopSpin, Bruker) automatically generate files containing information of the experimental settings as well as the applied transformation and processing steps to the raw data. A cloud based e-notebook (Benchling) is kept, as well a lab notebook on paper. This last one will be scanned regularly so that it can be added to the data as a digital copy.

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.

- Yes

WP4: Biophyscial characterization of a Me-PMO

For the NMR experiments, the software TopSpin (Bruker) generate a set of metadata containing the experimental details of the performed experiments.

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

Where will the data be stored?

Data will be kept as hard copies in notebooks and electronically through diverse clouds (Dropbox, Benchling ...). Obtained oligonucleotides and biological resources will be stored in the provided biological storage units, e.g. stored at -80°C. The obtained spectra of the NMR experiments will be printed and stored in binders. Results will be bundled in a PhD thesis and will be published in peer reviewed manuscripts, and, if possible, raw data will be deposited in open

platforms (Figshare).
Data will be stored carefully on external hard drivers (HDD). The central IT services of the KU Leuven will provide storage of data during at least 5 years after conclusion research.

How will the data be backed up?

The modelling laboratory of Rega Institute's Medicinal Chemistry is already well equipped to handle storage of large amount of data.

- The modelling data obtained from WP1 will be saved to the hard drive of the local modelling computer itself. This computer is mirrored to a drive/cloud (central rega drive), where a backup is generated nightly (i.e. external drive off site). On a regular basis, at least once a quarter and certainly when completing certain partial work packages, a additional backup will be made to an external hard drive (i.e. external hard drive on site). As for the data described in hard copy notebooks, these will be scanned on a regular basis with a copy of this scanned document being kept on the modelling computer/ central rega drive and another copy on the external hard drive.

- Data obtained from WP2 and WP3 will be stored electronically, using a electronic lab notebook (Benchling) (i.e. data storage off site). On a regular basis, at least once a quarter and certainly when completing partial work packages, a backup will be made to an external hard drive (i.e. data storage on site). As for the data described in hard copy notebooks, these will be scanned on a regular basis with a copy of this scanned document being kept on the external hard drive.

- Data obtained from WP4 is automatically saved to the local network (central rega drive) (i.e. data storage off site). On a regular basis, at least once a quarter and certainly when completing partial work packages, a backup will be made to an external hard drive (i.e. data storage on site). As for the data described in hard copy notebooks, these will be scanned on a regular basis with a copy of this scanned document being kept on the external hard drive. Obtained spectra and assignment lists will also be printed and stored in binders.

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

The modelling laboratory of Rega Institute's Medicinal Chemistry is already well equipped to handle storage of large amounts of data.

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

The Lab for Medicinal Chemistry is located in the Rega Institute, which is only accessible to people with a badge (i.e. people that work their). In addition, the modelling computers and external hard drives are located in a locked room. Computers are also protected by the KU Leuven firewall to avoid access from outside.

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

Costs for data storage on rega central data drive is shared by the Rega Institute. Estimate cost for purchase of the external hard drive will be €215,00 and will be covered by my bench fee.

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...).

The data will be preserved for 10 years according KU Leuven RDM policy.

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

The central IT services of the KU Leuven will provide storage of data during at least 5 years after conclusion research.

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

Depending the type of storage we will choose, the expected costs will vary:

- Large Volume storage: €95.14/TB/year
- Shared network drive: €450.76/TB/year
- File storage: €225.38 - €1,913.58/TB/year depinding on the type chosen

A part of my bench fee will be used for this.

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
- Yes, in a restricted access repository (after approval, institutional access only, …)

Data will be bundled to a PhD thesis and will be published in peer reviewed manucripts and, if possible, raw data will be deposited in open platforms (e.g. Figshare).

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

Some of the data obtained during this PhD project will be IP protected in collaboration of LRD, the tech transfer office of KU Leuven. This concerns nucleic acid sequences with proven antibacterial effect obtained from WP1 and WP2 of this research project. Agreements between the lab/KU Leuven (owner of these IP) and third parties will be made considering access to these IP protected 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 in the comment section per dataset or data type where appropriate.

- Yes, Intellectual Property Rights

Some of the data obtained during this PhD project will be IP protected in collaboration of LRD, the tech transfer office of KU Leuven. This concerns nucleic acid sequences with proven antibacterial effect obtained from WP1 and WP2 of this research object. Agreements between the lab/KU Leuven (owner of these IP) and third parties will be made considering access to these IP protected data.

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

Data will be bundled to a PhD thesis and will be published in peer reviewed manucripts and, if possible, raw data will be deposited in open platforms (e.g. Figshare). The software obtained by WP1 will be made available via GitHub.

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.

- Attribution (BY)
- Software license: GLP version 3 licence for the software generated by WP1

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

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

Difficult to say at this moment, part of these possible costs will be covered by my bench fee.

6. Responsibilities

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

Sten Reynders

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

Sten Reynders

Who will manage data preservation and sharing?

Prof. Dr. E. Lescrinier and Prof. Dr. V. B. Pinheiro

Who will update and implement this DMP?

Sten Reynders, Prof. Dr. E. Lescrinier and Prof. Dr. V. B. Pinheiro