
Plan Overview

A Data Management Plan created using DMPonline.be

Title: Theoretical Modelling of Photoinduced Processes in Fluorescent Proteins

Creator: Janko Civic

Principal Investigator: Janko Civic

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Janko Civic

Project abstract:

The characterization and development of fluorescent proteins marked a milestone in molecular imaging, offering a versatile toolset for tracking cellular processes and protein dynamics. Despite significant progress in enhancing their properties, further optimization remains a challenge, particularly in understanding complex photoinduced processes such as photoswitching and photoactivation. This project aims to advance our understanding of these phenomena through computational methods. By employing quantum-chemical methodologies and molecular dynamics simulations, this project seeks to explain the mechanism of photoinduced decarboxylation underpinning photoconversion and photoactivation of fluorescent proteins. Furthermore, the project will investigate the structure-function relationship of photoswitching, and explore implementation of novel methodologies for predicting protonation states of fluorescent proteins. Successful completion of this project will not only deepen our understanding of photo-transformations of fluorescent proteins but also pave the way for the design of improved variants with tailored photochemical properties. Moreover, insights gained from this research may extend to broader applications in understanding photoinduced processes in other biological systems.

ID: 213614

Start date: 02-10-2023

End date: 31-10-2028

Last modified: 25-03-2025

Theoretical Modelling of Photoinduced Processes in Fluorescent Proteins

DPIA

DPIA

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

- Not applicable

Theoretical Modelling of Photoinduced Processes in Fluorescent Proteins

GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Theoretical Modelling of Photoinduced Processes in Fluorescent Proteins

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
Raw	Collection of all input and output files of computational chemistry software.	New	Digital	<ul style="list-style-type: none"> Software input files Force-field parameter files Output files of quantum chemical calculations Processed trajectories and individual snapshots from molecular dynamics simulations 	<ul style="list-style-type: none"> .txt binary 	< 500GB	
Analysis	CSV files and jupyter notebooks used to generate plots and figures for analysis.	New	Digital	<ul style="list-style-type: none"> Results in CSV format Tables and graphs analysing raw data Jupyter notebooks containing the analysis procedure 	<ul style="list-style-type: none"> .txt .pdf .png .ipynb 	< 1GB	
Documentation	Research diary and other documents such as manuscripts and thesis.	New	Digital	<ul style="list-style-type: none"> Research diary Manuscript drafts/final versions Thesis drafts/final versions 	<ul style="list-style-type: none"> .pdf 	< 100MB	

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:

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.

- No

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).

A clear folder structure will be maintained for each work package with a descriptions in the form of Readme files.

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

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

Where will the data be stored?

Throughout the research, bulk of the data will be stored on the division's computer cluster, "Dirac." Other smaller documents like the research diary or manuscript drafts will be stored on a personal OneDrive account.

How will the data be backed up?

All data on the division's computer cluster, "Dirac", is backed up daily. The backups are managed by the ICTS department of KU Leuven.

**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

Each user on the division's computer cluster, "Dirac", has 400GB of available storage which is sufficient for the purposes of this project.

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

The access to the division's computer cluster, "Dirac", is protected through a secure authentication system and access is only possible when connected to a KU Leuven network. Off-campus access is possible only using a VPN.

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

No direct costs related to data storage during the time of the research are expected.

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...).

As dictated by the KU Leuven RDR policy all research data will be preserved for a minimum of 10 years.

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

All data will be stored on the division's computer cluster, "Dirac". If additional disk space is required, the university's K-drives (specialized for archiving) will be considered. Relevant data for reproducing the results will be deposited in the Research Data Repository (RDR) of KU Leuven. Published manuscripts along with their corresponding supporting information files will be deposited in the KU Leuven's repository system, "Lirias".

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

No direct costs related to data storage after the end of the research are expected. If K-drives are used, the cost of storing data for 10 years is expected to be less than 500 EUR. This cost will be covered by funds from my supervisor Jeremy Harvey.

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

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

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.

- No

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

Relevant data for reproducing the results will be deposited in the Research Data Repository (RDR) of KU Leuven. Published manuscripts along with their corresponding supporting information files will be deposited in the KU Leuven's repository system, "Lirias". Preprints of manuscripts will be deposited in ChemRxiv.

When will the data be made available?

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

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?

There are no expected costs related to data sharing.

6. Responsibilities

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

Janko Civic

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

Janko Civic, Hans Vansweevelt (ICTS)

Who will manage data preservation and sharing?

Janko Civic, Jeremy Harvey

Who will update and implement this DMP?

Janko Civic