Plan Overview

A Data Management Plan created using DMPonline.be

Title: Establishing a multifunctional catalytic platform for sustainable and divergent C-C coupling reactions via combined C-H activation and C-S, C-O or C-C heterogroup activation

Creator: Niels Van Herck

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Project abstract:
Carbon-carbon bonds form the bedrock of organic chemistry, yet their synthesis is still dominated by the use of harmful and waste-generating coupling partners such as organometallic reagents and organohalides. Consequently, novel methods that forge C-C bonds directly from C-H bonds and biologically innate heterogroups are highly desirable. Unfortunately, the inherently lower reactivity of these molecules causes reaction systems to become encumbered by an inefficient use of catalytic and energetic resources, hampering their wide-spread implementation in the industry. Inspired by Nature's way of leveraging a maximized functional involvement of catalytically active entities to mediate complex reactions, we here propose a multifunctional catalytic framework that enables challenging, divergent C(sp2)-C(sp/sp2/sp3) coupling reactions with sustainable and abundant C-H and C-heterogroup feedstock molecules. The cornerstone of our framework is laid by the use of multifunctional agents that mediate both C-H activation and C-S, C-O or C-C heterogroup activation in cooperation with a central Pd-catalyst. The cooperative catalytic assembly will be meticulously designed, characterized and manipulated to enable highly innovative C-C disconnections, including e.g. C-H formylation with CO_2-derived thioformates, C-H (fluoro)alkylation with carboxyl derivatives and C-H/C-CN metathesis with acetonitrile. As such, our rationale forms the impetus for a more resource-efficient catalytic design.

ID: 211253

Start date: 01-11-2024

End date: 31-10-2028

Last modified: 25-11-2024

Establishing a multifunctional catalytic platform for sustainable and divergent C-C coupling reactions via combined C-H activation and C-S, C-O or C-C heterogroup activation
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.

Dataset Name	Description	New or reused	Digital or Physical	Only for digital data Digital Data Type	Only for digital data Digital Data format	Only for digital data Digital data volume (MB/GB/TB)	Only for physical data Physical volume
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Primary experimental data	Raw datafiles containing material characterization (NMR, XAS, FTIR, etc.) Chromotographic data of reactions (GC, GC-MS, GPC etc.)	Generate new data	Digital	Experimental	• .gcd • NMR • .csv • .txt	< 100 GB	
Physical samples	catalysts, reaction mixtures, purified products	Generate new data	Physical				< 2000 vials
Data processing	Excel files, summaries, manuscripts, presentations	Generate new data	Digital	Experimental	• .xslx • .pptx • .pdf • .docx	< 1 GB	

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:

NA

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).

Reaction procedures will be described in a physical lab notebook and in a spreadsheet (.xslx) stored locally and in a cloud service (onedrive). Presentations and publications will provide sufficient details to allow replication by scientific peers. ReadME-files or recording methods for advanced characterization are stored in parallel with generated data.

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

Although no metadata standard is available for the advanced characterization in this project, most techniques store their own ReadMe files containing recording parameters and file information in parallel with recorded datasets. besides, some datasets contain metadata headers on recording parameters (eg. GC-FID). The data will be stored and named in consistent manner with unambiguous identifiers. Descriptions of experiments and materials are kept in physical lab notes for each identifier. The data will be processed and summarized in a comprehensive excel files with reference to the identifiers. The resulting

summaries will culminate in presentations and publications.

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

Where will the data be stored?

Physical notebooks of procedures and observations. Digital data with unambiguous identifiers for reactions and materials are stored in cloud storage from "Box". KU Leuven offers and Enterprise Box account 3, which offers and easy and secure storage space of up to 100 GB for each user.

How will the data be backed up?

The usage of cloud storage "Box". Digital data will be stored in parallel on the cloud and on the personal computer of the applicant, guaranteeing backup of the data. Most raw data is also stored and frequently backed up off-line on external hard drives.

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 cloud storage provides 100 GB per user, but the project is expected to generate much less than 100 GB of data.

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

The data will be stored in the university's secure environment for private data. An Enterprise Box account 3 for cloud storage ensured a secure environment

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

The costs of the Box data storage are included in an internal service contract with KU Leuven IT support service (SET-IT). No additional costs are expected for the storage of data. In case of additional costs, they will be covered by the project budget or reserve funds.

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 the data will be retained for the expected 5 year period after the end of the project with the exception of physical lab samples. These samples will stored for 2 years or until publication of a peer-reviewed manuscript. These samples have a shelf life and take up valuable space, making them difficult and unreliable to store over long time periods.

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

Upon termination of the contract, the data will be transferred to an external hard drive (Samsung Portable SSD T5 1 TB), which

is managed by Annelies Van Vlasselaer.

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

The high storage capacity of the external hard drive (1 TB) enables the preservation of the data for multiple years for various finished/terminated projects. As of now, 15 GB is used on average for the projects of each user, which allow around 60 users to safely store their data. The price of such a hard drive is around 120 EUR, which makes the cost per user negligible. The ITexpenses are included in the project's consumable expenses or covered by reserve funds.

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.

- Other, please specify:

General data and summaries (see "data processing" in part 1) will be made avaible through means of supplementary information files, which accompany a publication. Extra data such as chromatograms can be made available when asked by email.

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

Only research purposes will be allowed and commercial use is excluded.

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.

Data and summaries are available through supplementary information files which accompany a publication. Extra data such as chromatograms can be made available upon request by email.

When will the data be made available?

Upon publication of the results.

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

CC BY-NC-SA 4.0

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.

- No

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

No additional costs are expected a the data can be shared by online platforms (e.g. WeTransfer) or the already budgeted storage cloud from Box. If additional costs would occur, they will be covered from reserve funds.

6. Responsibilities

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

Niels Van Herck (Applicant)

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

Niels Van Herck (Applicant), Dirk De Vos (Promotor) and Annelies Vanvlasselaer

Who will manage data preservation and sharing?

Dirk De Vos (Promotor) and Annelies Vanvlasselaer

Who will update and implement this DMP?

Niels Van Herck (Applicant)