Characterization of natural posttranslational modifications of CCL2 and CCR2: role in activity regulation and quantification of proteoforms in inflammatory disorders

A Data Management Plan created using DMPonline.be

Creators: Paul Proost https://orcid.org/0000-0002-0133-5545, n.n. n.n.

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Data Manager: n.n. n.n.

Project Administrator: n.n. n.n.

Grant number / URL: G036423N

ID: 200885

Start date: 01-01-2023

End date: 31-12-2026

Project abstract:

CCL2 is part of the chemokine family and is crucial for recruiting
immune cells by binding to its receptor CCR2. Due to its central role in monocyte recruitment and activation during immune responses and the expression of CCR2 on endothelial cells, the CCL2/CCR2 axis has been implicated in inflammatory disorders and angiogenesis. Although CCL2 was found to be O-glycosylated more than three decades ago, limited information about the functional effects of this modification exists. In addition, the receptor CCR2 carries several predicted glycosylation and tyrosine sulfation sites in its, for chemokine binding, crucial N-terminal domain. We will produce defined O-glycosylated CCL2 proteins and compare these in binding and signaling assays on immune cells and endothelial cells to study the effects of CCR2 ligand and receptor modifications on molecular interactions and biological responses. We will develop methods to quantify expression of individual CCL2 glycoforms in complex samples to evaluate whether O-glycosylation is regulated in a cell-type/tissue specific manner and to detect glycoforms in patient samples such as synovial fluids and bronchoalveolar lavages. We will correlate ligand and receptor modifications with patient pathology and disease intensity. This project will enhance our understanding of the regulation of the activity of the CCL2/CCR2 system in immune responses and cancer.

Last modified: 30-06-2023

Characterization of natural posttranslational modifications of CCL2 and CCR2: role in activity regulation and quantification of proteoforms in inflammatory disorders
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

This application will generate raw data from experiments (numerical) which will be noted in lab books and kept in word and excel files. Some instruments (Mass spectrometry, flow cytometry, microscopy,...) will generate electronic data. These data will be analyzed and used to write manuscripts. Patient data will be kept by our clinical partners at UZ Leuven and we will only receive coded samples without personal information.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

The supervisors (Prof. P. Proost and Prof. Schjoldager) are responsible for long-term data storage. Electronic data (excel and word files, instrument-data files) will be stored locally: on internal computer disks or hard disks connected, to the workstations & on dedicated redundant NAS-devices (1st backup) and centrally (2nd back-up) on the KU Leuven central storage infrastructure or Univ. of Copenhagen data center. For research data KU Leuven Desktop File Storage is the first choice. KU Leuven and Univ. of Copenhagen datacenters provide storage on 2 locations and promise high availability and disaster recovery.

What’s the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

No deviation needed.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

Patient samples will be analyzed in this project. Patient identities and personal data will only be accessible to treating physicians with whom we collaborate at UZ Leuven. All patient samples will receive an anonymous code in the hospital that will be used in our laboratory. This code can only be linked to the individual patient by the treating physicians.

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

none

Characterization of natural posttranslational modifications of CCL2 and CCR2: role in activity regulation and quantification of proteoforms in inflammatory disorders
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
		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	
Proteins	CCL2 and CCR2 proteoforms and glycoforms	N	P	Frozen peptides (at -80) or cells (in liquid nitrogen)	NA	NA	20 tubes
binding studies	interactions with GAGs and (modified) receptors	N	D	Experimental	.xls or .fcs	<1GB	
activity measurements	detection of signaling and chemotactic activity	N	D	Experimental	.xls or .docx or .fcs	<1GB	
CCL2 proteoform quantification	Detection and quantification of CCL2 proteoforms and glycoforms by mass spectrometry	N	D	Experimental	.d	<1TB	
animal experiments	cell counts, tissue sections, ELISAs	N	D, P	Experimental	.xls and .docx	<100MB	100 sections
biological samples patients	synovial fluids, BAL fluids, plasma	N,P	P	synovial fluids, BAL fluids and sera (either already collected or newly collected)	NA	NA	<1L

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:

Synovial fluids and BAL fluid samples from patients with RA, JIA, AOSD (collaboration with Prof. C. Wouters, L. De Somer, S. Vanderschueren, B. Neerinckx, P. Verschueren and E. De Langhe), bacterial or viral pneumonia (collaboration with Prof. J. Yserbyt, E. Wauters and J. Wauters) or patients at enhanced risk for lung transplant rejection (collaboration with Prof. R. Vos, G. Verleden and B. Vanaudenaerde). Plasma and synovial fluids from crystal-induced arthritis patients will be provided by Prof. E. De Langhe as disease control for arthritis patients.

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.

- Yes, human subject data
- Yes, animal data

Animal experiments for evaluation of the activity of chemokine proteoforms (ethics application ongoing).
Use of human samples has been approved by the UZ Leuven ethics committee under the following numbers:
S65508 and S59874 for use of synovial fluids and plasma of arthritis patients
S58418 for blood samples of healthy volunteers
S63881, S51577, S61168, S58418, S63357 for BAL and plasma samples of pneumonia and lung transplant patients

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.

- Yes

Only pseudonymised samples will be obtained. Our clinical collaborators will provide info on sex and age of patients.

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

In case specific proteo- or glycoforms would be linked to specific diseases, these could be used as biomarker and this has potential for commercial exploitation. In addition, if specific proteo- or glycoforms would have specific activities they could be targets for drug development. As such, data obtained during the whole project should be evaluated for
potential commercial exploitation and data will be made available only after this evaluation and protection of potential IP.

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.

- Yes

An MTA has been signed with the Glycocenter of the Univ. of Copenhagen for restricted use of cells transfected or knock-out for enzymes involved in protein glycosylation. These cells (and cells created from these cells) can only be used in our group. Broader distribution of the cells is only possible upon agreement of Prof. H. Clausen from Univ. of Copenhagen.

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.

- Yes

see above

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).

1. Synthetic and recombinant proteins and biological samples will be stored at -80 °C and the location of tubes (exact freezer and position in the freezer) will be stored in the FreezerPro database of our laboratory.
2. Results related to GAG and receptor interactions, signaling and activity experiments (numerical data) will be registered in lab books (with detailed description of the used protocols) and electronically as excel or word (.xls or .docx) documents which automatically imprint the metadata (user, date, time, equipment, parameters) from the experiments. Information on quantification and experimentation parameters will be embedded byt he users on the document folders an in the lab books in order to improve reproducibility and maintenance of data.
3. Microscopy images: Imaging data is created by default with metadata imprinted by the image acquisition software’s automatically. That includes information on user, date and time, duration of experiments, equipment parameters and imaging configurations. The metadata is saved (also in OME format) and transferred with the original imaging file. The created data files will be organized in folders named by the date of the experiment (YYYYMMDD) followed by the researcher who performed it and the title of the experiment. The methodology and protocol of each experiment will be described in detail in a lab book.
4. Flow cytometry data: Flow cytometry templates are saved which automatically stores the parameters (voltages, compensation...) that are used during the acquisition of the 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

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

Where will the data be stored?

The data will be stored in several locations, including on internal computer disks (at KU Leuven and Univ. of Copenhagen), at the shared local virtual drive (Rega drive), in One Drive, in redundant NAS (network adapted storage)-devices, and on the KU Leuven and Univ. of Copenhagen central storage servers. The KU Leuven datacenters provide storage on two locations and promise high availability and disaster recovery to preserve data for a long period. Hard copy notebooks with raw data will be stored physically in our laboratory. The large raw data volumes from analysis equipment are stored redundant on hard disks in or connected to the lab computers and work stations. The backups of the analysis data are stored on dedicated redundant NAS-services. Also, we will use the Lirias platform as data repository for published material.

How will the data be backed up?

We will use the centrail server storage of KU Leuven (Data centre ICTS Luna storage) and Univ. of Copenhagen, which provide a daily automatic backup. Moreover, the data will be backed up on the Rega Institute Virtual Drives (Rega NAS (network adapted storage)) and on external hard-drives by the investigators.

**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

Available storage & backup capacities by far exceed our needs.

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

All research data generated during this project will be secured by the need for login, registration on datacenter/luna and use of u-number and password, which are also restricted. Freezers with stored material are located in our laboratories which are all under electronic access badge control.

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

Long-term data storage and costs will be managed by the principal investigator working in the project, Paul Proost. The costs for data storage is 520€/y/TB.

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

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

The data will be stored redundantly during and after the research in our PCs, in external hard-drives, and in the KU Leuven data centers (ICTS Luna storage and Rega NAS).

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

Long-term data storage and costs will be managed by the principal investigator working in the project, Paul Proost. The expected cost for data storage is 520€/Y/TB.

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 a restricted access repository (after approval, institutional access only, ...)

All data that are not anymore subject to potential IP protection and that are not restricted by the MTA agreement with Prof. Clausen (concerning transfected cells, see above).

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

Access by external users will be evaluated and needs to be authorized by Paul Proost. Access to cells modified in the glycosylation pathways needs to be negotiated with Prof. Henrik Clausen of Univ. of Copenhagen.

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, Privacy aspects
- Yes, Intellectual Property Rights

patient samples are pseudonymised and personal data will only be available from the treating physician and upon approval by the ethical committee
All data obtained will be checked for potential IP protection before they are shared.

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 protection of IP and publication of research results

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

Data transfer agreement (restricted data) or material transfer agreement

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

DOI will be used for publications

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

Costs for transfer of animal or patient material (only after ethics approval) will be covered by the researcher requesting the materials. Costs depend on transport needs. Data transfer to partners will be at the partners cost.

6. Responsibilities

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

The researchers who generate the data.

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

The researchers who generate the data.

Who will manage data preservation and sharing?

The project coordinator Paul Proost and research expert Dr. Mieke Gouwy for KU Leuven and Katrine T. Schjoldager for signaling data obtained at Univ. of Copenhagen

Who will update and implement this DMP?

Paul Proost and Mieke Gouwy