Plan Overview

A Data Management Plan created using DMPonline

Title: Exploring glycoproteins in Alzheimer's disease: Mass Spectrometry insights into Disease Related Abnormalities and their ramifications

Creator: Amalia Kontochristou

Principal Investigator: Amalia Kontochristou

Data Manager: Amalia Kontochristou

Project Administrator: Dr. Melissa J. Bärenfänger, Prof. Dr. Anouk M. Rijs

Affiliation: Vrije Universiteit Amsterdam

Template: 1 - VU DMP template 2021 (NWO & ZonMW certified) v1.4

ORCID iD: 0009-0006-5988-0024

Project abstract:

Neurodegenerative diseases are rising in our ageing society, but effective treatments are missing, and the disease pathway is not fully understood. Several studies have connected the altered protein glycosylation to the progression of neurodegeneration; however, a causal relationship between the pathological disease progression and abnormal protein glycosylation still needs to be established. This project aims to elucidate the role of glycoproteins that interfere with disease-causing protein aggregation in the brain by establishing a structure-function relationship between protein glycosylation and glycoprotein activity.

ID: 136010

Start date: 01-08-2023

End date: 31-07-2027

Last modified: 09-01-2024

Grant number / URL:
https://www.rijksoverheid.nl/documenten/kamerstukken/2022/07/14/bestuursakkoord-2022-hoger-onderwijs-en-wetenschap

Exploring glycoproteins in Alzheimer's disease: Mass Spectrometry insights into Disease Related Abnormalities and their ramifications

0. General information

0.1 Document version & date

Version 3.0
Date: 09 / 01 / 2024

0.2 Project title

Exploring glycoproteins in Alzheimer's disease: Mass Spectrometry insights into Disease Related Abnormalities and their ramifications

0.3 Project summary

Alzheimer's disease (AD) is a neurodegenerative disease accounting for 60–70% of cases of dementia. The number of Alzheimer's patients is increasing significantly in our ageing society, resulting in an enormous financial and social burden that affects patients, family members and caregivers alike. Without any effective treatment for Alzheimer's disease (AD), we need to re-evaluate our strategies as the conventional approaches are failing. A central focus is the amyloid hypothesis that describes the accumulation of amyloid β-peptides (Aβ) in the brain as the leading cause of pathological neurodegeneration. Tremendous efforts have been made to develop drugs that target these Aβ-peptides or formed amyloid fibrils. However, they still lack efficacy in slowing down or reversing the progression of fibril formation. To open novel routes for therapeutic approaches, this project will not focus on Aβ itself but on glycoproteins that directly interact with Aβ and can promote or inhibit aggregation. These glycoproteins hold great potential to direct Aβ aggregation by controlling folding or acting as disaggregase (reversing aggregation). This project will explore this potential by investigating the effects of selected glycoproteins to correct Aβ aggregation by utilizing advanced mass spectrometry (MS) approaches. Understanding the pathological processes in AD is complex because of the intricate interplay between biophysical and biochemical parameters. Studies focusing on isolated aspects, such as abnormalities in proteomics, often fail to uncover causal relationships but provide valuable correlational data. These studies show that protein glycosylation, the attachment of sugar moieties to proteins, is abnormally regulated in the brains of Alzheimer's patients. For example, N-glycosylation shows increased fucosylation, bisecting N-acetylglucosamine and reduced sialylation. These abnormalities can influence protein function, which also holds for glycoproteins in this project.

A striking example is the glycoprotein clusterin, where different glycoproteoforms promote or inhibit Aβ aggregation. However, the relationship between protein glycosylation and its influence on pathological peptide aggregation is still poorly understood, and a structure-function relationship is missing. This project will unravel how glycoproteins modulate Aβ aggregation and decipher the impact of the altered protein glycosylation observed in AD patients. Therefore, this project will focus on the glycoproteins: prostaglandin-H2 D-isomerase, clusterin, transferrin, alpha-1-antichymotrypsin and haptoglobin. These proteins are highly N-glycosylated, directly interact with Aβ and are secreted into the circulation.

Aim of the project

This project aims to control Aβ aggregation via glycoproteins that stop or reverse Aβ aggregation. To do so, this project will develop tools for glycoprotein characterization and develop MS-based structural and functional assays for selected glycoproteoforms, which are currently still lacking in our field. Therefore, the following objectives are addressed:

• Objective 1: Identifying disease-specific glycoproteoforms in AD patients.
• Objective 2: Evaluating the structural & conformational changes of disease-specific glycoproteoforms.
• Objective 3: Unravelling the glycoproteoform-specific impact on the Aβ aggregation.

The combined objectives will decipher the role of glycoproteoform structure and function in the aggregation of pathogenic Aβ. This understanding is crucial to potentially prevent the negative consequences of amyloid aggregation on human health.

Strategy

This project is organized into three consecutive work packages (WPs). WP1 will describe differences in protein N-glycosylation observed in AD patients by mass spectrometry. The glycoproteoforms that are overexpressed in AD patients and those corresponding to healthy phenotypes will be enriched in WP2. Subsequently, MS-based structural and functional assays to describe their role in the aggregation of pathogenic Aβ will be performed in WP3. All CSF and plasma samples used in this study will be provided from Amsterdam UMC, VUmc.

WP1: Characterization of glycoproteins

To characterize changes in glycosylation by MS (Objective 1), glycoproteins from cerebrospinal fluid (CSF) and plasma will be isolated from AD patients and control subjects. CSF and plasma are valuable biological sources due to their clinical applications and can represent the glycosylation status of the brain. Immunoprecipitation with commercial antibodies offers a straightforward approach to isolating glycoproteins, and standard protocols can be utilized. A more significant challenge is the accurate assessment of the glycosylation profile. Strategies, such as isoelectric focusing or lectin staining, have already identified AD-specific changes in

glycosylation for several glycoproteins, but these techniques cannot define glycan structures precisely. Hyphenated MS techniques will be employed for an unambiguous characterization. Since the selected glycoproteins differ in size and number of glycosylation sites, a glycoprotein-specific strategy will be developed for each glycoprotein. This will identify disease-specific glycoproteoforms (Objective 1). Additionally, results from this WP will lead to new insights into disease pathology and have the potential to be developed into glycoprotein-based biomarkers.

WP2: Enrichment of relevant glycoproteoforms

Numerous studies revealed that aberrant glycosylation correlates with the development of Alzheimer's disease. However, a causal relationship between function and distinct glycoproteoforms is not explored due to the complexity of isolating glycoproteoforms. This challenge will be addressed by enriching glycoproteoforms carrying the pathogenic and non-pathogenic glycan motifs observed in WP1 from pooled CSF and plasma samples. For this, one can use commercially available lectin beads which recognize distinct glycan motives. Furthermore, this approach can be combined with native separation techniques, such as ion exchange chromatography, to separate glycoproteoforms based on their number of sialic acid moieties. Although this approach is labour-intensive and needs to be optimized for reasonable sample amounts, it offers the unique advantage of having enriched glycoproteoforms that are not recreated in vitro but represent the in vivo condition. While complete isolation of defined and unique glycoproteoforms is most likely not feasible, this strategy allows the enrichment of glycoproteoforms with specific glycan motifs.

WP3: Structure and function of selected glycoproteoforms

While WP3a and WP3b can be performed independently and provide novel insights on their own, the combination of structural and functional information will better understand why certain glycoproteoforms differ in their activity.

- WP3a: Structural analysis of isolated glycoproteoforms
  For proteins, function follows form. Native ion mobility mass spectrometry (IM-MS) can be used to assess the structural stability of different glycoproteoforms observed in WP1 (Objective 2). Collision-Induced Unfolding (CIU), an IM-MS application, allows an estimation of the overall protein structure and stability of mass-isolated glycoproteoforms. It was already demonstrated for transferrin that glycoproteoforms with reduced sialylation show decreased binding to Aβ due to conformational differences. The greatest challenge here is to preserve the native state of glycoproteins while obtaining resolved MS spectra. This is especially difficult for glycoproteins that present a heterogeneous glycosylation profile with overlapping MS signals. I will overcome this as-yet-unsolved obstacle by working with isolated glycoproteoforms obtained in WP2, which result in less complex and deconvoluted MS spectra. This allows me to utilize MS to provide rapid and glycoproteoform-specific experimental setups, which are rarely available.
- WP3b: Functional analysis of isolated glycoproteoforms on Aβ aggregation
  The influence of selected glycoproteoforms on Aβ aggregation will be studied by mass spectrometry. The goal is to study the glycoproteoform-specific impact on Aβ aggregation (Objective 3) by incubating isolated glycoproteoforms from WP2 with Aβ peptides. It is expected to detect the complex formation of Aβ with different glycoproteoforms, resulting in stochiometric information. Furthermore, using IM-MS, the aggregation kinetics can be determined by detecting the ratio of Aβ monomers and oligomers over time. Aβ aggregation studies are commonly performed with fluorescent assays. However, mass spectrometry-based methods will provide additional insights into the exact molecular mechanism. Moreover, these experiments will be performed with glycoproteins that have not yet been considered potential drug candidates. Conducting these studies on isolated glycoproteoforms will reveal their potential to slow down, stop or reverse Aβ aggregation and define the influence of deviating glycosylation observed in AD patients.

0.4 At which VU Faculty is this project situated?

- Faculty of Science (BETA)

0.5 Your contact details

Full name: Amalia Kontochristou
Your role in the project (please refer to the CRediT contributor roles): Conceptualization (1), Data Curation (2), Formal Analysis (3), Investigation (5), Methodology (6), Software (9), Validation (11), Visualisation (12), Writing- Original Draft (13)
Email: a.kontochristou@vu.nl
ORCID (LibGuide): https://orcid.org/0009-0006-5988-0024
University: Vrije Universiteit Amsterdam
Faculty/Institute: Faculty of Science
Department/Research Group: Bioanalytical Chemistry

0.6 List other people involved, including those at partner organisations in the project (if applicable)

Full name of person(s) involved:

Prof. Dr. Anouk M. Rijs
Their role(s) in the project (please refer to the CRediT contributor roles): Conceptualization (1), Funding Acquisition (4), Resources (8), Supervision (10), Writing- Review & Editing (14)
Email: a.m.rijs@vu.nl
ORCID (LibGuide): https://orcid.org/0000-0002-7446-9907
University: Vrije Universiteit Amsterdam
Faculty/Institute: Faculty of Science
Department/Research Group: Bioanalytical Chemistry

Dr. Melissa J. Bärenfänger
Their role(s) in the project (please refer to the CRediT contributor roles): Conceptualization (1), Funding Acquisition (4), Investigation (5), Methodology (6), Resources (8), Software (9), Supervision (10), Writing- Review+Editing (14)
Email: j.m.baerenfaenger@vu.nl
ORCID (LibGuide): https://orcid.org/0000-0002-2855-924X
University: Vrije Universiteit Amsterdam
Faculty/Institute: Faculty of Science
Department/Research Group: Bioanalytical Chemistry

0.7 Funding organisation & grant number (if applicable)

Funding organisation: *National Government, Ministry of Education, Culture and Science (OCW)*

0.8 Project code (if applicable)

Not applicable

0.9 Consulted data management expert(s)

Name: Dr. Brett G. Olivier
Email: b.g.olivier@vu.nl
University: Vrije Universiteit Amsterdam,

Date of consultation: DD / MM / YYYY
(has yet to be planned)

1. Data description

1.1 Will you collect and/or process personal data in this project?

- Yes

Specialised hospital staff will utilise cerebrospinal fluid and plasma samples from humans. It's important to note that only limited information related to the gender and age of the patient will be collected for the purposes of this project. All data involving human subjects will be provided to us, analysed, and stored by us according to the Declaration of Helsinki developed by the World Medical Association (WMA).

1.2 Will you use existing data? If yes, what is their source?

No

1.3 Will you collect or produce new data? If yes, please describe how.

New data will be produced in this project. Due to the use of multiple analytical techniques, such as liquid chromatography and mass spectrometry, using different instrument-specific software, various data formats will be obtained. The techniques mentioned will be used to analyse glycosylated proteins contained in cerebrospinal fluid (CSF) and plasma samples (from humans). CSF and plasma samples will be provided to us by the Amsterdam UMC, VUmc. The samples will be stored at -20oC (locked freezer; only people involved in the project will have access to the samples), and they will be analysed using first separation techniques (such as affinity chromatography and reverse phase chromatography) and then detection techniques (such as mass spectrometry).

Data asset: Liquid Chromatography (LC)
Description: Chromatograms will be generated using either HPLC (Agilent) or nanoLC (Thermo Fischer)
Formats: .CSV, .txt, .png, .jpeg

Data asset: Mass Spectrometry (MS)
Description: Mass spectra will be generated using different types of mass spectrometers.
Formats: .d (Bruker DataAnalysis), .wiff (SCIEX), .CSV

1.4 Describe the population/participants/subjects that will be studied

Cerebrospinal fluid and plasma samples will be collected from Alzheimer's patients and healthy individuals, female and male, from 40 to 90 years old. The sample will be collected by specialised hospital staff and provided to us from the Amsterdam UMC, VUmc. The samples, after collection, will be stored at -20oC. The number of samples that will be used has still to be determined.

1.5 Do you process any of the following (personal) data?

- Other, please specify below

In this project, only limited personal data will be processed. Specifically, the patient's name will not be mentioned. The data collected and processed will be restricted to the age and gender of the patient and whether the patient has Alzheimer's disease.

1.6 Do you process the personal data based on informed consent?

- Yes, through a physical form

I will not personally collect the consent form. Specialised hospital staff will ensure that CSF and plasma samples are collected and that the patients sign the consent forms before sample collection. Then, they will provide us with the collected samples and a list mentioning the age and gender of the patients and whether they have Alzheimer's disease.

1.7 On what legal ground will the data processing take place if it is not based on informed consent?

- Not applicable, I use informed consent

1.8 Does the data collection include any of the following types of personal data?

- Data concerning health

The data collection for this project includes information related to the patient's health. The focus will be solely on the presence or absence of Alzheimer's disease; no other detailed health information will be collected.

1.9 If your research involves special categories of personal data (previous question) and you will not use explicit informed consent, what is the legal ground for the exemption?

Not applicable

1.10 What kinds of outputs will you produce in this project? Please describe these data assets.

Amsterdam UMC, VUmc will provide the CSF and plasma samples and patient data, including the age, gender and whether the patient has Alzheimer's disease. The samples will be stored at -20oC and analysed using separation and detection techniques. The output data will include the following:

Raw data:
Data asset: Instrumental data: Liquid Chromatography (LC)
Description: Measured data using an HPLC (Agilent) or a nanoLC (Thermo Fischer)
Formats: .CSV, .txt, .png, .jpeg
Data asset: Instrumental data: Mass Spectrometry (MS)
Description: Measured data generated using mass spectrometers.
Formats: .d (Bruker DataAnalysis), .wiff (SCIEX), .CSV
Data asset: Parient's information
Description: Information describing the age, gender and whether patients have Alzheimer's disease.
Formats:.txt

Processed data:
Data asset: Mass Spectrometry (MS) data used for proteomics and glycoproteomics analysis Description: Mass spectrometry data raw data has to be converted to other data formats (such as mgf, mzML, and mzXML) to be comparable to the analysis software used.
Format: .mgf, .mzML, .mzXML

Analysed data:
Data asset: Pictures, graphs, tables, powerpoints, documents
Description: Visualisation of produced and processed results
Format: .xls, .pdf, .docx, .ppt, .png, .jpeg

Other Analysis software:

- Bruker Compass DataAnalysis version 5.0, https://www.bruker.com/en/products-and-solutions/mass-spectrometry/ms-software.html;
- MacCoss Lab Software: Skyline version 23.1, https://skyline.ms/project/home/software/Skyline/begin.view;
- MSfragger version 20.0, https://www.nesvilab.org/software.html;
- MSConvertGUI (64-bit) version 3.0.23272-0f4c78d, https://proteowizard.sourceforge.io/download.html;
- Matrix Science Mascot Deamon version 2.8.0 (64-bit), https://www.matrixscience.com/daemon.html

1.11 How much digital data storage will your project require?

- 1 - 2 TB

The exact amount of digital data storage needed depends on the amount of raw data files. The estimated amount needed is between 500GB and 2TB. Data is stored on the SURFsara and Yoda cloud drive provided by Vrije Universiteit Amsterdam.

1.12 Will you collect physical data? If yes, please describe these.

No physical data will be collected. The biological samples, the consent forms from patients and the patient information are not collected by us but by specialised hospital staff.

1.13 Will you take measures to ensure data quality? Please describe these, if applicable.

1. Implement validation checks during data collection to identify and correct errors.
2. Develop cleaning protocols to address inconsistencies, outliers, and missing values.
3. Document data collection methods, including instruments, protocols, and any changes made using eLabFTW.
4. Conduct regular audits of the dataset to identify and rectify errors or inconsistencies.
5. Implement access controls to ensure that only authorized personnel can modify the dataset.
6. Regularly monitor and log data access to identify and address any unauthorized changes.
7. Regularly back up the dataset to prevent data loss.

2. Legal and ethical requirements, codes of conduct

2.1 What legislation applies to your research project? Please tick the relevant boxes for your project.

- General Data Protection Regulation (GDPR)/ Algemene Verordening Gegevensbescherming (AVG)

2.3 Do you require approval of an ethical committee for this project? If yes, please indicate which ethical committee and whether you have obtained approval for this project.

- Yes

Yes, ethical approval is potentially required for this project. First, I have to contact the Faculty's Privacy Champion to confirm it.
Ethics Committee: BETCHIE
Approval status: Not yet
Review code:
Approval date: DD/MM/YYYY (has to be planned)
(Needs to be updated)

2.4 Will you work with data for which intellectual property and/ or confidentiality are an issue? If yes, please describe.

- Yes

Since I am working with personal data, I must comply with the General Data Protection Regulation. As a result, the CSF and plasma samples will be stored in a locked freezer (-20oC), and all documents related to personal data will be stored in Yoda drive (restricted access). Only people involved in the project will have access to these files.

2.5 Do you plan on generating a marketable product from your research project? if yes, please describe

- No

3. Storage and back-up during the research process

3.1 What measures will you take to secure and protect data during the research process? Please describe, for each separate data asset you described for question 1.10, how you will ensure data security, where the data assets are stored & backed up, and who has authorization to access the asset.

All data will be stored in SURFDrive and Yoda and archived in Yoda for long-term storage. All data (stored and backed up) will be available only to people involved in the project.

In details:

Raw data:
Data asset: Produced liquid chromatography (LC) and mass spectrometry (MS) data and patient information
Storage: SURFDrive
Backup: Yoda
Access: restricted access (only people involved in the project)
Security measures: All personal data will be anonymised, and access to all digital data assets (protected by a password) will be restricted, according to General Data Protection Regulation (GDPR) regulations.

Processed data:
Data asset: Mass spectrometry data used for glycoproteomics analysis
Storage: SURFDrive
Backup: Yoda
Access: restricted access (only people involved in the project)
Security measures: There will be restricted access to all digital data assets (protected by a password), according to General Data Protection Regulation (GDPR) regulations.

Analyzed data:
Data asset: Liquid Chromatography and Mass Spectrometry data
Storage: SURFDrive
Backup: Yoda
Access: restricted access (only people involved in the project)
Security measures: There will be restricted access to all digital data assets (protected by a password), according to General Data Protection Regulation (GDPR) regulations.

CSF and plasma samples provided by the hospital will be stored in a locked freezer, and only people involved in the project will have access to the samples.

3.3 Which tools are used in the collection, processing or storage of data during research?

- Yoda
- SURFDrive

3.4 What other tools or software do you intend to use during your research?

Name: Bruker Compass DataAnalysis version 5.0 (https://www.bruker.com/en/products-and-solutions/mass-spectrometry/ms-software.html)
Role: Analysis of mass spectrometry data
Country: The Netherlands

Name: MacCoss Lab Software: Skyline version 23.1 (https://skyline.ms/project/home/software/Skyline/begin.view)
Role: Analysis of mass spectrometry data
Country: The Netherlands

Name: MSfragger version 20.0 (https://www.nesvilab.org/software.html)
Role: Analysis of processed mass spectrometry data
Country: The Netherlands

Name: MSConvertGUI (64-bit) version 3.0.23272-0f4c78d (https://proteowizard.sourceforge.io/download.html)
Role: Convert mass spectrometry data to other types of data
Country: The Netherlands

Name: Matrix Science Mascot Deamon version 2.8.0 (64-bit) (https://www.matrixscience.com/daemon.html)
Role: Process proteomics data
Country: The Netherlands

Name: Matrix Science Mascot Database search (https://www.matrixscience.com/search_form_select.html)
Role: Process proteomics data
Country: The Netherlands

3.5 Is it necessary to transfer the (physical or digital) data assets to other locations or research partners? If yes, please describe how you secure the file transfer.

- No

No, transferring the digital data assets to other locations or research partners is unnecessary. All individuals involved in the project can access the SURFDrive and Yoda, where the data is securely stored.

3.7 Do you transfer personal data outside of the European Economic Area (EEA)? If Yes, please provide additional information

- No

4. Data archiving and publishing

4.1 Which data assets will be archived and which will be published?

After completion, all data assets generated in each subproject will be archived in Yoda Drive. Personal data will be archived in Yoda with restricted access, and a mandatory anonymization process will precede any data publication, ensuring adherence to GDPR regulations and protecting individuals' privacy. Assets related to publications will be designated for public access (apart from personal data) and published through Yoda Drive.
All CSF and plasma samples will be used during the research, so no archive is required.

4.2 Where will you archive your data assets?

- Yoda

4.3 What other archive(s) do you intend to use to archive data assets?

Not applicable for now.

4.4 For how long will the data be available in the archive?

The data generated in the course of this project will be available in Yoda for a minimum period of 10 years after its completion.

4.6 Where will you publish your data assets?

The data assets from this project will be published through Yoda, ensuring the protection of personal data, appropriately documented and accompanied by metadata to enhance its visibility and utility for other researchers. The documents containing personal data will not be published.

4.8 How will you ensure your data assets get a persistent identifier (e.g. a DOI-code)?

The data assets will be assigned a persistent identifier, specifically a Digital Object Identifier (DOI), through Yoda. Yoda automatically generates a DOI upon publishing and storing the data, providing a unique and persistent link directly connected to the dataset. Metadata, including comprehensive information about the dataset, will be meticulously curated and associated with the DOI to enhance the visibility, accessibility, and proper attribution of the data assets within the scholarly community.

4.9 Will you register your datasets in an online registry other than PURE? If yes, where?

No

4.10 Are there restrictions to data publishing? If yes, please specify the reasons and list the data assets you do not wish to share publicly.

Yes, with the exception of data associated with personal data, all other data from this project can be published. Due to privacy and confidentiality concerns, information linking samples to specific patients will not be made publicly available. Access to this specific data will be restricted to authorised personnel to ensure compliance with ethical standards and privacy regulations.

4.12 When will you share the data? If not immediately after completion of the project, please specify the reasons.

The data from this project will be shared immediately after completion. However, due to privacy and confidentiality concerns, personal data linking samples to specific patients will only be accessible to individuals directly involved in the project.

4.13 Please indicate the license and/ or terms of use under which you share your data.

Attribution-NonCommercial 4.0 International

5. Documentation

5.1 What documentation and metadata will accompany the project?

Dataset registrations in PURE and CERIF metadata standards will be followed.
Metadata will be generated by Yoda after each publication and also at the end of my project. Additional documentation will include the lab journal (eLabFTW) linked to each experiment performed, describing methodologies, sample info and preliminary results. Moreover, a *Readme* file in .txt format will describe how the data is stored.

5.2 What metadata and documentation will accompany the data assets?

Readme files in .txt format will accompany data folders of all subprojects, and they will be stored in both Yoda and SURFsara.
In addition, PowerPoint and Word documents will be provided to explain and elaborate on the data. All raw and processed/analysed data will be mentioned within these files.
Metadata document is automatically generated from Yoda after data publication. DataCite Metadata Schema v4,4 properties, indicated in Tables 1 and 2, will be used for data documentation.

Table 1: DataCite Mandatory (M) Properties

ID	Property	Obligation
1	Identifier (with mandatory type sub-property)	M
2	Creator (with optional name identifier and affiliation sub-properties)	M
3	Title (with optional type sub-properties)	M
4	Publisher	M
5	PublicationYear	M
10	ResourceType (with mandatory general type description sub-property)	M

Table 2: DataCite Recommended (R) and Optional (O) Properties

ID	Property	Obligation
6	Subject (with scheme sub-property)	R
7	Contributor (with type, name identifier, and affiliation sub-properties)	R
8	Date (with type sub-property)	R
9	Language	O
11	AlternateIdentifier (with type sub-property)	O
12	RelatedIdentifier (with type and relation type sub-properties)	R
13	Size	O
14	Format	O
15	Version	O
16	Rights	O
17	Description (with type sub-property)	R
18	GeoLocation (with point, box and polygon sub-properties)	R
19	FundingReference (with name, identifier, and award related sub-properties)	O
20	RelatedItem (with identifier, creator, title, publication year, volume, issue, number, page, publisher, edition, and contributor sub-properties)	O

5.3 What methods, software or hardware are needed to access and use your data?

Multiple software will be needed to access and use the raw data, including:

- Bruker Compass DataAnalysis version 5.0
- MacCoss Lab Software: Skyline version 23.1
- MSfragger version 20.0
- MSConvertGUI (64-bit) version 3.0
- Matrix Science Mascot Deamon version 2.8.0 (64-bit)

An additional Readme file will be provided to explain in detail how the mentioned software can be used to access and use the data.

6. Data management responsibilities and resources

6.1 Who will be responsible for management of the data assets during the project? Please specify their name, position, role in the project, and faculty/ institution/ group.

Full name: Amalia Kontochristou
Your role in the project (please refer to the CRediT contributor roles): Investigator
Email: a.kontochristou@vu.nl
ORCID (LibGuide): https://orcid.org/0009-0006-5988-0024
University: Vrije Universiteit Amsterdam
Faculty/Institute: Faculty of Science
Department/Research Group: Bioanalytical chemistry

6.2 Who will be responsible for management of the data assets after completion of the project (e.g. the project lead/ dedicated data manager/ department head)? Please specify their name, position, role in the project, and faculty/ institution/ group.

The department head will be responsible for the management of the data assets after the completion of the project.
Full name: Matthias F. Bickelhaupt
Your role in the project (please refer to the CRediT contributor roles): Investigator
Email: f.m.bickelhaupt@vu.nl
ORCID (LibGuide): 0000-0003-4655-7747
University: Vrije Universiteit Amsterdam

Faculty/Institute: Faculty of Science
Department/Research Group: Chemistry and pharmaceutical sciences (CPS)

6.3 For data that are only available upon request, what methods will be used to handle requests for access and how will data be made available to those requesting access?

Access will be handled in compliance with the European Data Protection Regulation (GDPR) regulations for data that is only available upon request. The following methods will be employed:

1. All requests for access to restricted data should be submitted through a formalised process, which may involve a designated contact person or a secure online portal.
2. Requests will be subject to a thorough verification process to ensure the identity and legitimacy of the requester following GDPR requirements.
3. Each request will undergo a legal and ethical review to assess compliance with GDPR regulations and other relevant legal frameworks.
4. For approved requests, data access will be granted under a restricted access protocol, limiting access to the specific data requested while maintaining compliance with data protection regulations.
5. Data provided in response to requests will undergo anonymisation or pseudonymisation processes as necessary to protect the privacy of individuals, ensuring compliance with GDPR principles.
6. Data will be transferred securely to the requester using encryption and other secure methods to prevent unauthorised access during transmission.
7. All requests, approvals, and data transfers will be thoroughly documented to ensure a transparent and auditable process consistent with GDPR record-keeping requirements.

6.4 What resources (for example financial and time) will be dedicated to research data management? Please estimate their cost.

Active data management will be done every week to ensure proper data organisation and storage. The data will be stored in both SURFsara and Yoda cloud drives. The online storage drives used are provided by the VU (cost unknown).