# Type III intermediate filaments: molecular structure to understand muscular disease

*A Data Management Plan created using DMPonline.be*

**Creator:** Sergei Strelkov

**Affiliation:** KU Leuven (KUL)

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

**Template:** FWO DMP (Flemish Standard DMP)

**Grant number / URL:** G035223N

**ID:** 201068

**Start date:** 01-01-2023

**End date:** 31-12-2026

**Project abstract:**

Intermediate filaments (IFs) are core elements of metazoan cytoskeleton. The diverse and vast family of IF-forming proteins plays a key role in cell mechanics. A growing number of inherited IF mutations cause currently incurable skin, muscular and neuronal diseases. Moreover, while contributing to cell migration, IFs are important in the context of cancer progression. At present, understanding of IF structure and assembly, which is driven by the interactions of elongated, partially disordered protein dimers, remains elusive. Here we will address the molecular architecture of vimentin IFs, as well as the impact of myofibrillar myopathy mutations R350P and R406W on assembly and properties of desmin filaments, via an integrative approach. To this end we will, first, create atomic models of IF dimers and tetramers based on crystallographic data. Second, we will perform cryoelectron microscopy on the ‘unit-length’ vimentin filaments as well as on the filaments formed by mutated desmins in immortalized myoblasts. Third, chemical cross-linking coupled to mass spectrometry will be used to decipher the interactions between dimers in the wild-type and mutated filaments. Finally, integrative modelling should marry the atomic resolution data with the spatial constraints provided by other methods, yielding a trustworthy molecular model of cytoplasmic IFs. This should enable novel insights into the molecular mechanism of currently incurable desminopathies.

**Last modified:** 13-07-2023

# Type III intermediate filaments: molecular structure to understand muscular disease
# 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)**

Protein purification and other biochemical data (lab notes and graphs in digital format)
X-ray diffraction data (raw and reflection lists)
CryoEM data (raw images and processed data)
Atomic models of protein/ligand complexes (PDB format)
Crosslinking data (Excel spreadsheets)
In silico modelling data

**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)**

Steven Beelen (permanent technician/lab manager in Biocrystallography lab) will be responsible for data archivation and storage. This is based on the already existing practice in the lab (for instance, Xray diffraction data have been stored in entirety over the last 10 years).
At present, the lab owns a Network Attached Storage (Synology) with 30TB capacity (high-quality 10TB HDDs). We plan to gradually extend this to at least 100TB, depending on demand. The storage is organized as a RAID array. These data files will be preserved for at least 5 years following the completion of the project.

**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)**

We do not wish to deviate from this principle

**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)**

N/a

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

N/a

# Type III intermediate filaments: molecular structure to understand muscular disease
# 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 |
| | | | | | | | |
| Proteins | Biochemical data | New data | Digital | Experimental | .pdf | <100GB | |
| EM | CryoEM data | New data | Digital | Experimental | .img and others | <50TB | |
| Diffraction | X-ray data | New data | Digital | Experimental | .img, .mtz | <5TB | |
| Structures | Atomic coordinates | New data | Digital | Experimental | .pdb | <1GB | |
| Models | In silico results | New data | Digital | Experimental | various | <1TB | |
| Cross-links | Cross-linking data | New data | Digital | Experimental | .xls and others | <5TB | |

**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:**

No reuse

**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).**

Data sheets will be kept for all large datasets (X-ray, cryoEM, cross-linking)

**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?**

Lab network-attached storage (NAS) of 100TB

**How will the data be backed up?**

Lab NAS

**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

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

The NAS is only for internal use, under a firewall

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

Covered by the FWO funding, several thousand euros

## 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

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

Lab NAS

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

No extra costs beyond initial storage and backup

## 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.**

- No (closed access)

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

The data will be for internal use only

**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

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

N/a

**When will the data be made available?**

N/a

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

N/a

**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?**

N/a

## 6. Responsibilities

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

Steven Beelen

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

Steven Beelen

**Who will manage data preservation and sharing?**

Steven Beelen

**Who will update and implement this DMP?**

Sergei Strelkov

# Type III intermediate filaments: molecular structure to understand muscular disease

# GDPR

## GDPR

**Have you registered personal data processing activities for this project?**

- Not applicable

# Type III intermediate filaments: molecular structure to understand muscular disease
# DPIA

## DPIA

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

- Not applicable