11P4824N_The role of crystal growth in secondary nucleation from nuclei breeding: the nature of the crystal-solution boundary layer

A Data Management Plan created using DMPonline.be

Creator: Lorijn De Vrieze

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 11P4824N

ID: 205521

Start date: 01-11-2023

End date: 31-10-2027

Project abstract:

Crystallization is a crucial but challenging step in the development and production of pharmaceutical products. Preformed crystals of the desired chemical compound (seeds) are commonly added to a crystallizer. This drastically speeds up crystallization and ensures required critical product qualities. In seeded crystallizers, the two processes creating new crystalline material are (1) secondary nucleation and (2) crystal growth. Although both occur simultaneously in real-life, research studies typically focus on only one of both.

This way, the role of crystal growth in the process of secondary nucleation is highly overlooked. In this project, we hypothesize that crystal growth plays a non-negligible role in the process of nuclei breeding, a mechanism significantly affecting the outcome of secondary nucleation. This will lead to fundamental insights into secondary nucleation, aiding in better control of industrial crystallization.

This project focuses on experimentally showing the hypothesized role of crystal growth in secondary nucleation using a novel flow crystallizer. As both processes take place in the crystal-solution boundary layer, we propose to investigate the nature of this boundary layer using Raman microspectroscopy. To provide more theoretical insight into the stability of the boundary layer and a possible nucleation pathway, a 2D-Ising model will be developed.

Last modified: 29-03-2024

11P4824N_The role of crystal growth in secondary nucleation from nuclei breeding: the nature of the crystal-solution boundary layer

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	
Formed crystals	Solid material (paracetamol, glycine, NaClO3) generated from experiments	Generate new data	Physical				Every sample < 10 g
Liquid samples	Solutions of solvent (water) containing amount of solute (paracetamol, glycine, NaClO3)	Generate new data	Physical				Every sample < 50 mL
Raman spectroscopy	Raman spectra for crystals and solutions	Generate new data	Digital	Experimental	.csv, .tvb, .txt	< 1 GB	
Scanning electron microscopy	SEM pictures for crystal surface characterization	Generate new data	Digital	Experimental	.tiff	< 100 GB	
Light microscopy	Pictures of crystals (crystal growth + new crystals)	Generate new data	Digital	Experimental	.tiff	< 100 GB	
Photos	Pictures/photos of setups and crystals	Generate new data	Digital	Experimental	.jpeg, .png, .tiff	< 1 GB	

Temperature profiles	Recorded temperatures from crystallization experiments	Generate new data	Digital	Experimental	.csv, .dat, .txt	< 100 MB	
3D printing	Design of setup parts	Generate new data	Digital	Compiled/aggregated data	.form, .par, .stl	< 100 MB	
Ising model simulations	Computer simulations modelling solute clustering	Generate new data	Digital	Simulation data	.gif, .png, .py, .txt	< 1 GB	
Plots and graphs	Data elaboration and visualization	Generate new data	Digital	Compiled/aggregated data	.gnu, .eps, .pdf, .png, .tiff	< 1 GB	
Images	Graphics	Generate new data	Digital	Compiled/aggregated data	.eps, .pdf, .pdn, .png	< 100 GB	
Notes	(Electronic) Lab notebook	Generate new data	Digital	Compiled/aggregated data	.docx, .xlsx	< 100 MB	
Manuscripts	Written texts/scientific publications	Generate new data	Digital	Compiled/aggregated data	.docx, .pdf, .tex	< 1 GB	
Presentations	Slides and presentation notes	Generate new data	Digital	Compiled/aggregated data	.pdf, .ppt	< 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:

Not applicable

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).

- **Formed crystals:** Created crystalline material will be stored in the lab according to safety standards and protocols. Every sample will be labelled unambiguously ("name_time_conditions") and will be recorded in a lab notebook. Subsequently, the sample name (combined with the corresponding conditions and observations) is added and saved into an electronic list (.docx, .xlsx), which also contains a more detailed description of the followed experimental procedure.
- **Liquid samples:** Vials containing liquid solution are stored in the lab according to safety standards. Every sample will be labelled unambiguously ("name_time_conditions") and will be recorded in a lab notebook. Subsequently, the sample name (combined with the corresponding conditions and observations) is added and saved into an electronic list (.docx, .xlsx), which also contains a more detailed description of the followed experimental procedure.
- **Raman spectroscopy:** Spectra are received in .csv or .txt format and named "name_time_conditions". Standard Operating Procedures (SOP) are followed to analyze the Raman spectra before plotting. The followed procedures and observations are written in a lab notebook and subsequently saved digitally (electronic list in .docx, .xlsx).
- **Scanning electron microscopy:** Pictures are captured according SOPs and saved as .tiff files with the name "name_time_conditions". The device settings are also automatically displayed in the saved pictures. Afterwards, samples cannot be stored.
- **Light microscopy:** Pictures are taken and saved according to SOPs and saved as .tiff files with the name "name_time_conditions". The pictures are saved together with their metadata following OME standard.
- **Photos:** Relevant photos of the setup, samples, observations etc. taken in the lab are saved in .jpeg, .png or .tiff format.
- **Temperature profiles:** Temperatures will be recorded using SE374 software. Data will be saved in .dat or .txt format. The name of the experiment is added and saved into an electronic list (.docx, .xlsx), which also contains a more detailed description of the followed experimental procedure.
- **3D printing:** 3D models for setup construction are design in Solid Edge (.par) and saved digitally. After exporting the 3D model into an .stl file, the physical part is made by 3D printing. Both the .stl files as the physical part will be stored (either digitally or inside a setup).
- **Ising model simulations:** Computer simulations are conducted in Python. README.txt files will be made to accompany each source code file and will contain the simulations conditions, the followed procedures and observations. The source code files are saved as .py files while the output is provided as .gif or .txt files. All files will be stored on a private GitHub profile.
- **Plots and graphs:** Data from experiments or simulations are provided as .txt files that can be plotted using Gnuplot software. Both the code for plotting as the output will be saved digitally. The code for plotting is in .gnu format, while the output can be .eps, .png, .pdf, .tikz depending on the final application of the plot/graph.
- **Images:** Images will be created using Paint.net and stored as .pdn files. The exported images will have the .eps, .png, .pdf, .tikz format depending on the final application.
- **Notes:** The physical lab notebook is used for setup design (sketches) and writing down of observations. When not used in the lab, the notebook will be stored in a locked drawer in the office. Important notes (samples names, conditions etc.: see above) are compiled into an electronic list (.docx or .xlsx).
- **Manuscripts:** Communication of scientific results is done via manuscripts and saved in .docx, .pdf or .tex.
- **Presentations:** Presentations containing scientific results are produced using PowerPoint and stored accordingly (.ppt or .pdf).

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?

All data will be stored on the work laptop of the researcher, on an external hard drive and in a shared cloud (OneDrive). After completion of (sub)WPs, the data will be additionally stored on the KUL service servers.

How will the data be backed up?

Data are backed up on the cloud (OneDrive) immediately. The software indicates the update status (green, blue or red) and, in case of a non-sync, action can be taken using the online version of the tool. Data are further back up regularly on an external hard drive. After completion of (sub)WPs, data will be additionally backed up on the KUL service servers.

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

Data storage should not exceed 1.5 TB.

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

Only authorized researchers have access to the OneDrive account: all users need to use a two-factor authentication to log in (MFA app from KUL). On lab computers, a log-out is always performed to prevent modification by unauthorized people. Hard drives and physical lab notebooks are stored in a closed drawer in the office.

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

The expected costs for data storage and backup during the research project are around €5000. These costs will be covered by Prof. Simon Kuhn.

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 years period after the end of the project.

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

All the data will be stored for the long-term on the KUL service servers (with automatic back-up procedures), conform with the KUL RDM policy.

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

The expected costs of data storage and preservation are around €5000. These costs will be covered by Prof. Simon Kuhn.

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
- Yes, in a restricted access repository (after approval, institutional access only, …)

Data relevant for publication will be made available in an Open Access repository. Full datasets will only be made available upon request.

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

Not applicable

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 relevant for publication will be made available in an Open Access repository. Further data can be made available upon valid request via email to the researcher and/or the responsible of the data after project ending.

When will the data be made available?

Data will be made available upon publication of scientific findings.

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

There are no data that need a usage license.

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?

The expected costs for data sharing is €0. Free tools like Belnet FileSender (KUL account) will be used for data sharing. In the unlikely event that there would be costs, these costs will be covered by Prof. Simon Kuhn.

6. Responsibilities

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

Lorijn De Vrieze

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

Lorijn De Vrieze

Who will manage data preservation and sharing?

Prof. Simon Kuhn

Who will update and implement this DMP?

Lorijn De Vrieze; Prof. Simon Kuhn bears the end responsibility of updating and implementing this DMP.