EOSC activities in the Photon and Neutron community

Oscar Matilla

ALBA Synchrotron
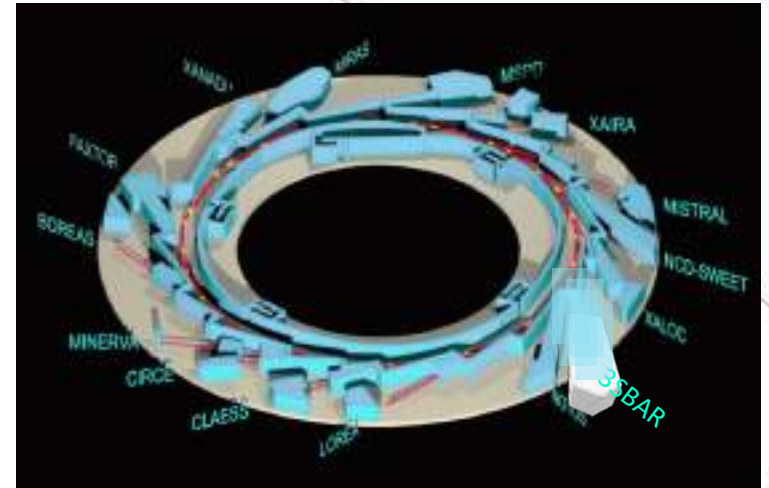
10-10-2022

ALBA Synchrotron

+2.200 users /year

+300 experiments per year

- X-Ray Beamlines
    - 10 beamlines in operation
    - 1 beamline in commissioning
    - 3 beamline under construction
- JEMCA – Electron Microscope Center hosted at ALBA
    - 3 main instruments

Agreement with Portugal signed in 2019 for collaboration in different areas.
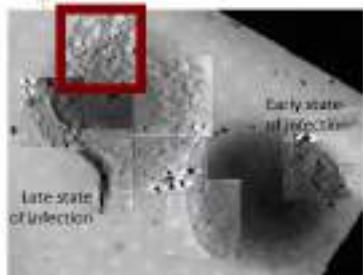
www.albasynchrotron.es/

ALBA Synchrotron

National public institution, funded
50% national Government (Ministerio de Ciencia e Innovación)
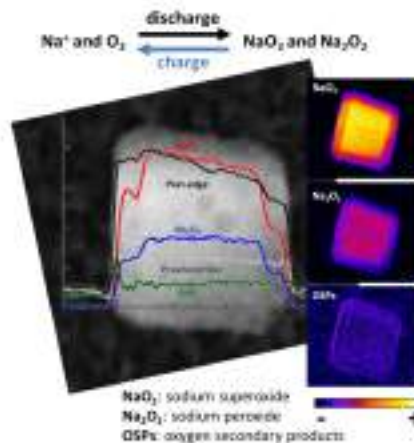50% regional GenCat (Department de Recerca i Universitats)

Life Science
From the protein to the cell

Cell infected by covid-19

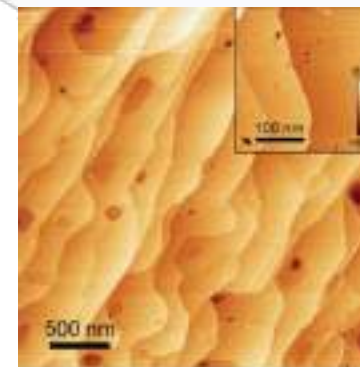### Chemistry and Material Science
Energy material, catalysts, environment

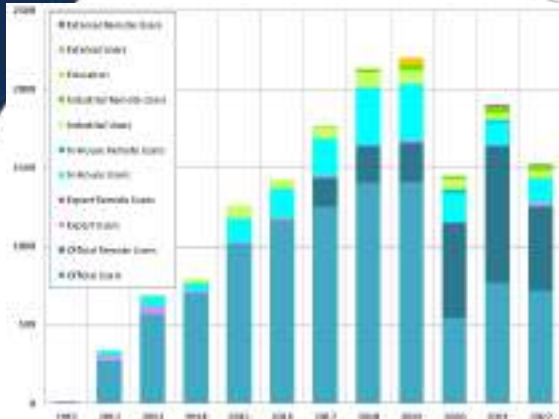
Battery developments

Electronic and Magnetic Structure of Matter
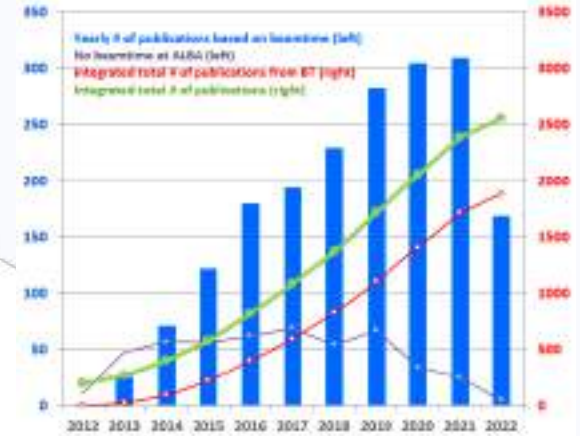Advanced materials
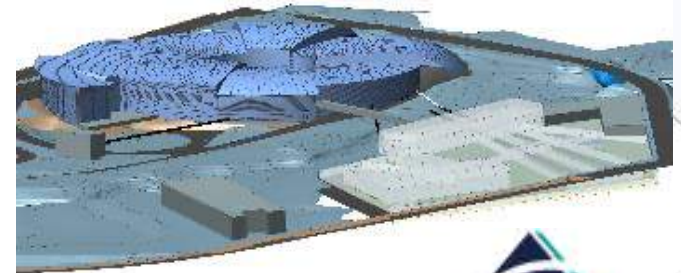

Nanomagnetism for data storage

ALBA Synchrotron

- We are increasing our scientific productivity every year
- 2020-2021: More than 38 publications/BL/Year
- Average IF >8

- And toward a complete upgrade of the facility (**ALBA II**)in 2030 which will multiply by 30 times our brilliance

The Photon and Neutron Community

LEAPS

19 facilities
+800.000 h/year
+35.000 users /year
+300 operating stations
+25.000 publications in the last 5 years

https://leaps-initiative.eu

LENS

9 facilities
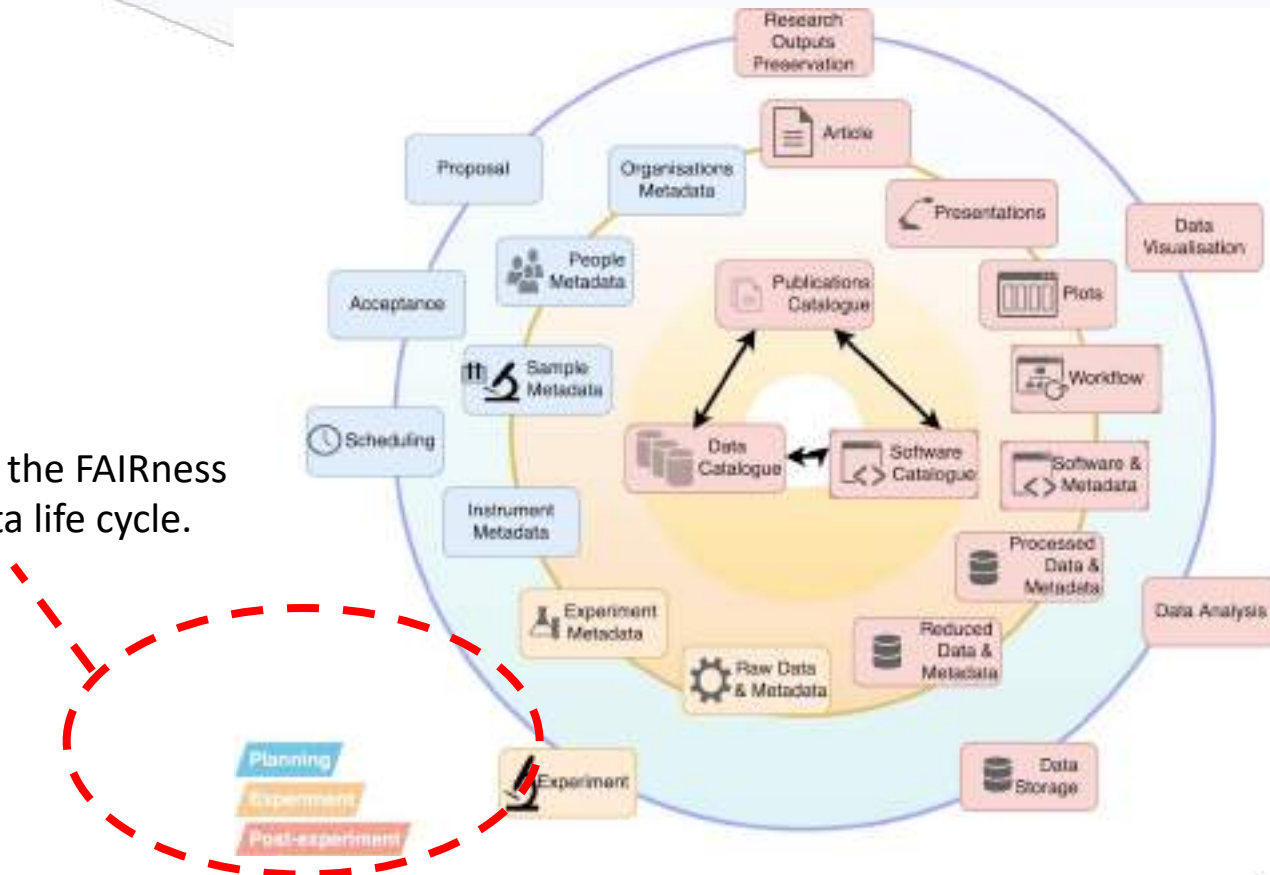
https://lens-initiative.org

- Our community has a long tradition of collaborating in many fields for decades.

Data Management in Photon and Neutron Sources

- RIs are crucial to assure the FAIRness of data in the whole data life cycle.

Data Management in Photon and Neutron Sources

- Our user sends a proposal for an experiment.
- All proposals go through a competitive open call (200% over-subscription)
- If granted, the experiment is scheduled and detailed metadata information is requested.

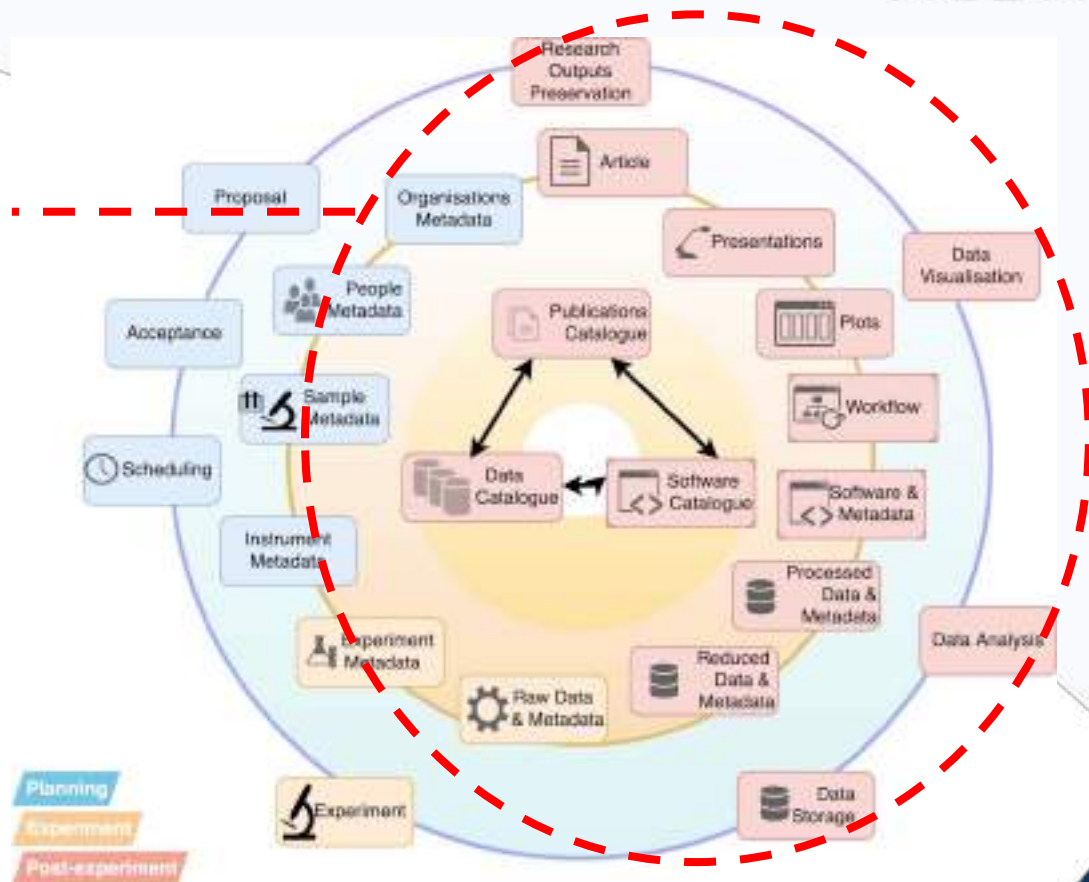# Data Management in Photon and Neutron Sources

- Data acquired is highly heterogeneous. Each beamline can perform multiple types of experiments with great flexibility and each type have hundreds of metadata fields.
- Data Catalogue implementation and guaranteeing **FAIR data when leaving the facility** is a major challenge.
- During experiments Data processing is necessary to validate the acquired data.
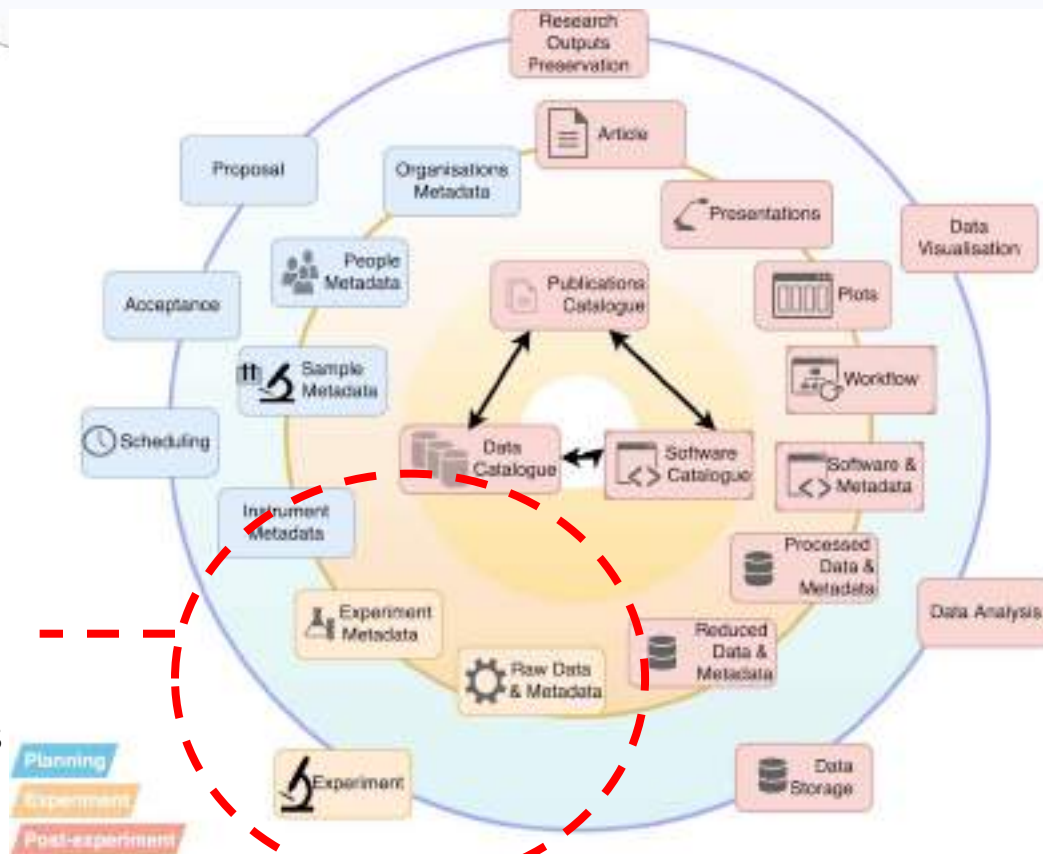
# Data Management in Photon and Neutron Sources

- Data is curated, stored and preserved.
- Following the experiment, complete **data analysis pipelines** are required to extract the required scientific knowledge.
- Open Scientific software in every field and containerized tools will be required to guarantee **Data Provenance** in all steps leading up to a publication.

Data Management in Photon and Neutron Sources

- Data volumes and needs for computation are increasing year after year. We are immersed in a **data deluge**.

    - For example, the ALBA FAXTOR beamline will produce +45GB/s continuous data in two years.
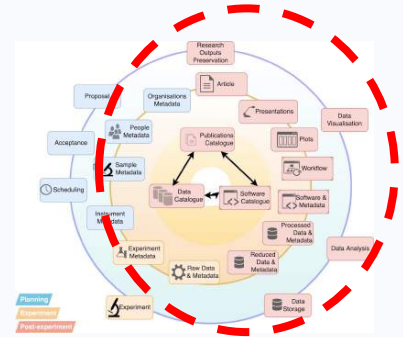
Evolution in the Generation of raw data (no processed data considered) in ALBA.
The ALBA II is expected to produce a big increase in addition.
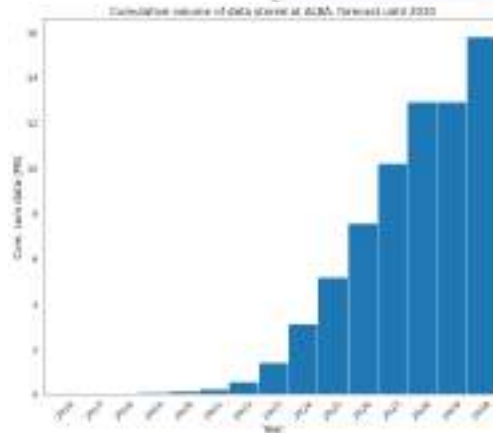
- Our users will not be able to process the raw data without our support, as a powerful HPC cluster will be needed to extract the scientific results.

Data Management in Photon and Neutron Sources

- Our strategy as RIs and service providers for the EOSC is to extend the data services we need to provide to our scientific users to the broader scientific community.
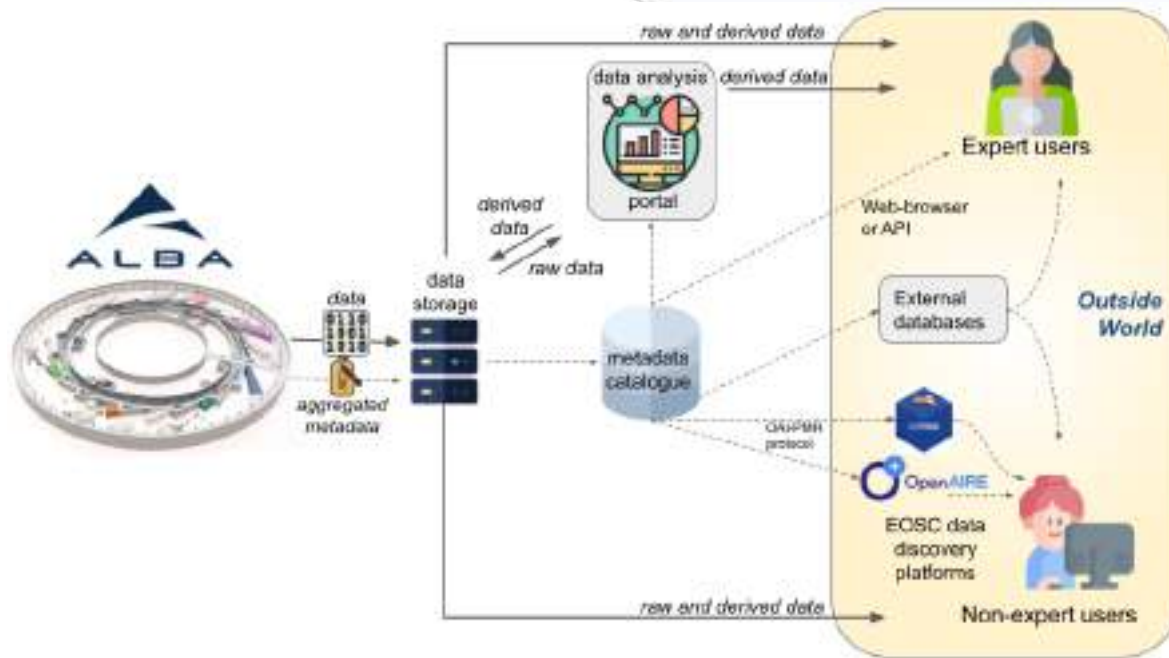
PaNOSC and ExPaNDS projects

- In PaNOSC and ExPaNDS we (18 Research Infrastructures) are laying the foundations on every brick needed for implementing all the required services to the EOSC.

  - Guarantee of FAIRness in the collected data once outside the facility.
  - Generation of DOIs during experiment.
  - Implementation of a common Data Portal.
  - Building the motor for Data Catalogue and search APIs.
  - Establishing common data formats.
  - Start ontologies frameworks.
  - Implementing common containerized DaaS platforms.
  - Developing environments to guarantee data provenance.

* Detailed info next Wednesday @16h - FAIR data in the Photon and Neutron community (Patrick Fuhrmann)

https://expands.eu
https://www.panosc.eu

PaNOSC and ExPaNDS projects

| FACILITY	FAIR data policy	DMPs	DOIs	Nexus HDF5	Search API	Open Data Portal	AAI	Jupyter Lab	VISA	VINYL/ OASYS/ McStas	Pan-learning / training
ALBA	P	P	WIP	WIP	WIP	WIP	P	Y	WIP	N	U
DESY	WIP	WIP	WIP	Y	WIP	P	WIP	Y	U	Y	WIP
CERIC-ERIC	Y	WIP	Y	WIP	Y	Y	Y	Y	Y	Y	Y
ELETTRA	Y	WIP	Y	Y	Y	Y	Y	Y	Y	Y	Y
ESRF	WIP	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
ELI-ERIC	WIP	Y	P	Y	Y	Y	WIP	Y	Y	Y	Y
ESS	WIP	Y	Y	Y	Y	Y	Y	WIP	WIP	Y	Y
EuXFEL	WIP	WIP	Y	WIP	Y	Y	WIP	Y	WIP	Y	Y
FELIX	Y	P	WIP	U	U	WIP	U	U	N	N	U
HZB	Y	P	WIP	Y	P	Y	P	U	U	U	U
HZDR	WIP	WIP	Y	N	N	WIP	WIP	WIP	P	N	Y
ILL	Y	WIP	Y	Y	WIP	Y	Y	Y	Y	Y	WIP
MAX-IV	WIP	U	Y	Y	Y	Y	Y	Y	U	U	U
PSI	WIP	WIP	Y	WIP	Y	Y	WIP	WIP	N	N	N
PTB	Y	WIP	Y	WIP	N	Y	N	N	N	N	N
SOLEIL	Y	WIP	WIP	Y	WIP	WIP	Y	WIP	WIP	U	Y
SESAME	WIP	U	P	Y	P	WIP	P	P	N	Y	N

Yes, already adopted (Y)   Planned to be adopted (P)
Not Planning to be adopted (N)   Under evaluation (U)
In progress of being adopted (WIP)

https://expands.eu
https://www.panosc.eu

Conclusions

- **PaNOSC** and **ExPaNDS** have **laid strong foundations** towards the future implementation of the **EOSC Services** in the **PaN** community.

- There should be a **continuation** of these efforts to guarantee future success and not lose the current inertia. There are multiple open fields:
  - **Assuring the FAIRness** of the data collected is a big challenge for the future.
  - The **reusability of the data** implies comprehensive software developments in multiple fields.

- The **Data Policy of the RIs must be updated**, embracing FAIR principles. But the impact on the required IT resources may limit its application. **A model to guarantee its sustainability must be found.**
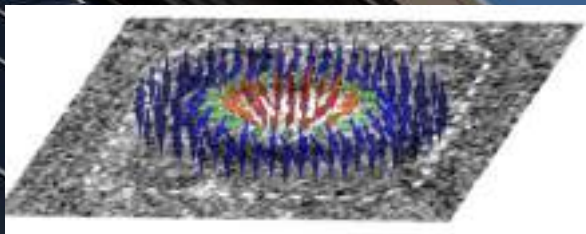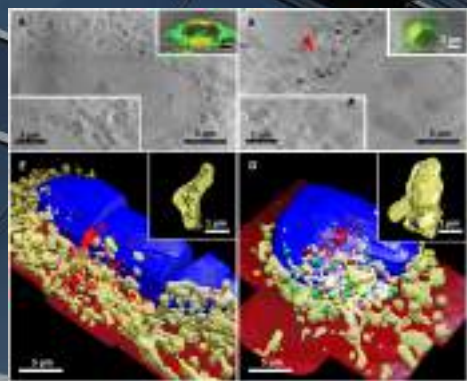
Nature Nanotechnology (2016) O. Bulle ….L. Aballe,
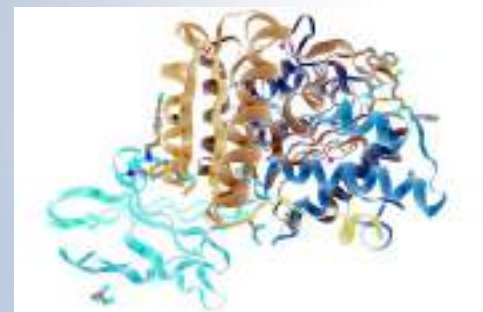M. Foester, …G. Gaudin , ALBA
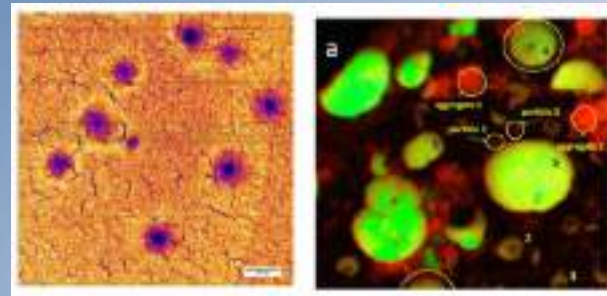
Synchrotron X-ray studies of Li battery
materials (courtesy: ALBA)

Thanks!

A combination of X ray tomography
(ALBA) and fluorescence imaging (ESRF)

DOI: 10.2210/pdb6I9A/pdb
Porphyromonas gingivalis gingipain K (Kgp) in complex
with inhibitor KYT-36