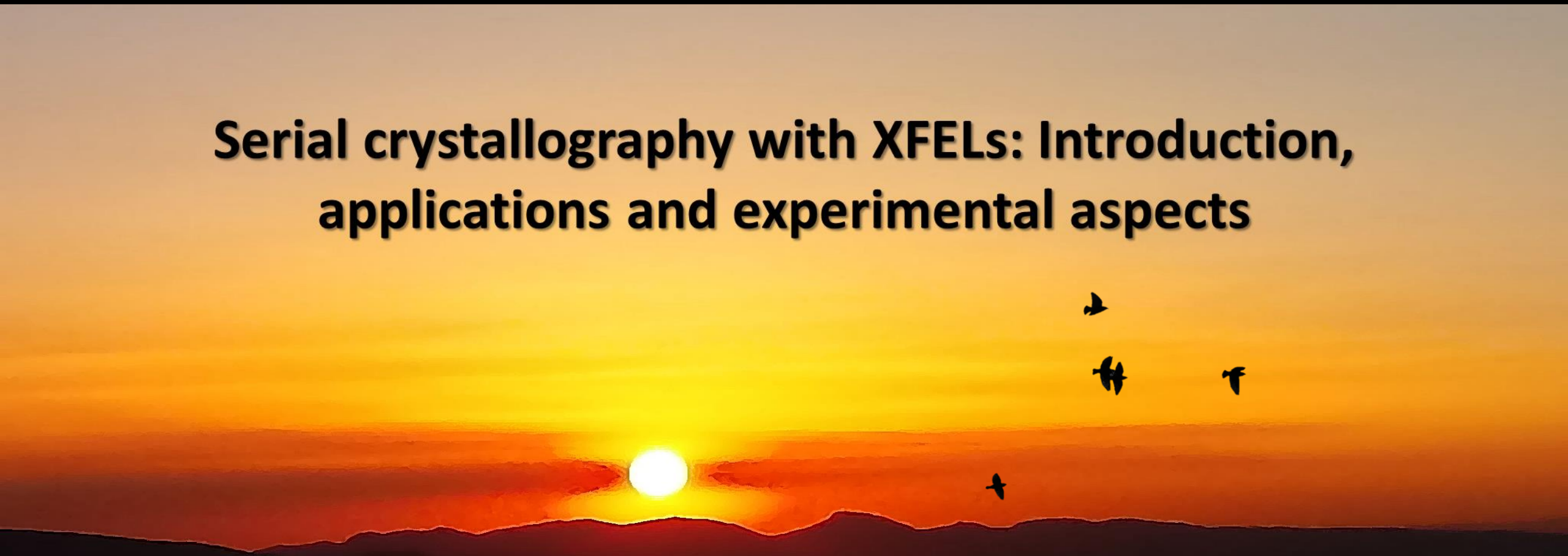
Serial crystallography with XFELs: Introduction, applications and experimental aspects

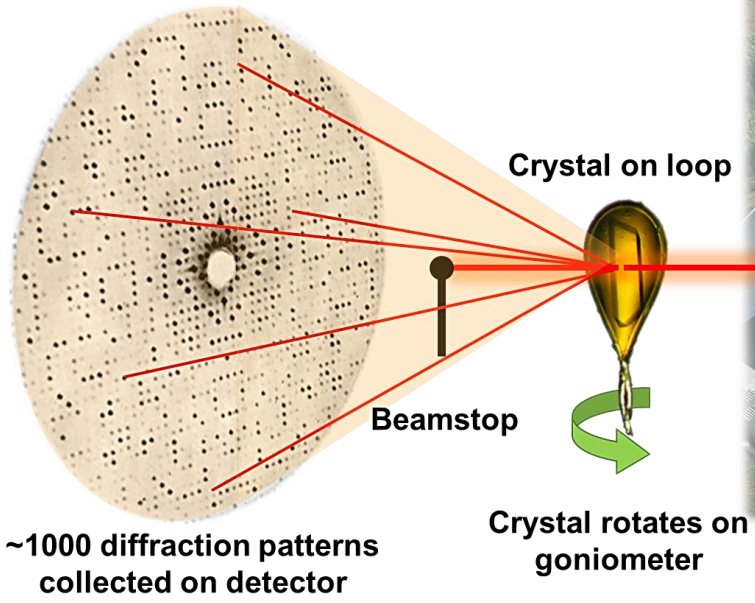


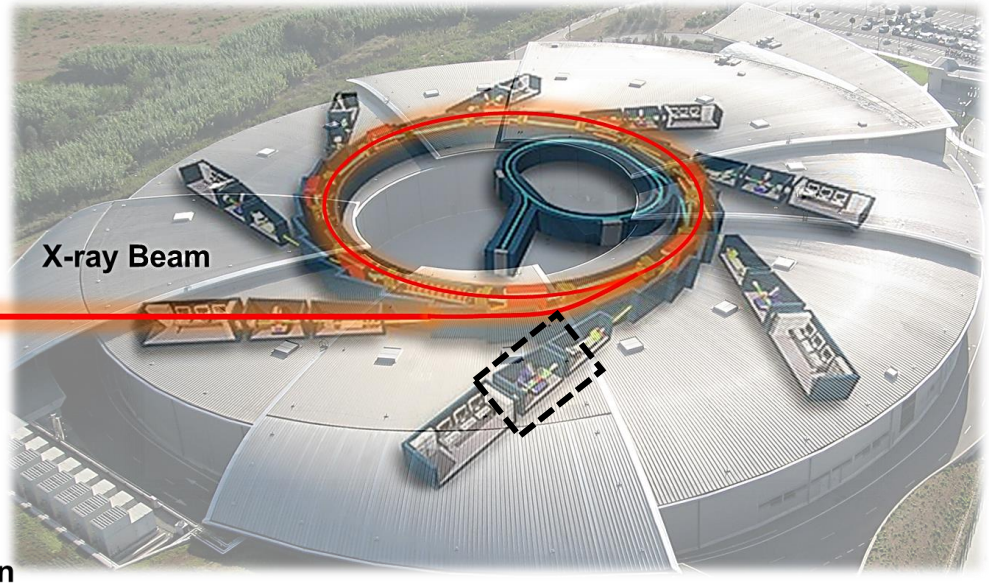
José M. Martín García
Institute of Physical Chemistry Blas Cabrera
CSIC, Madrid
jmmartin@iqfr.csic.es

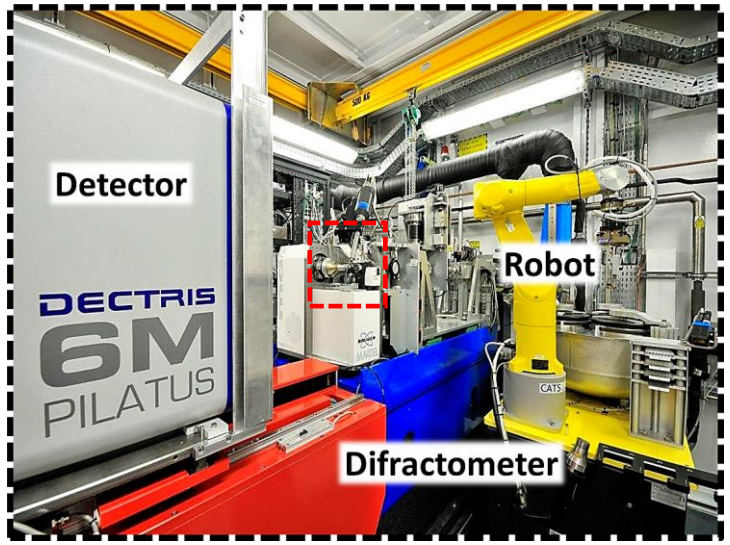
MCCS-2023
Madrid, May 11th, 2023

Classical macromolecular crystallography

Classical macromolecular crystallography





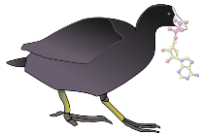


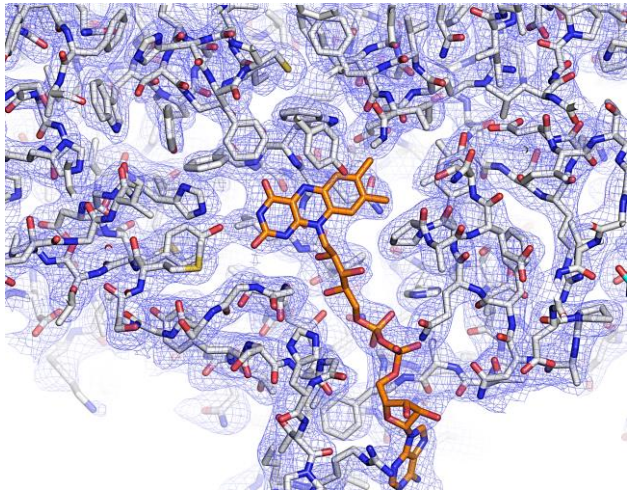
Figures adapted from **Gisriel C, Fromme P, Martin-Garcia JM*. Methods Mol Biol. 2021**

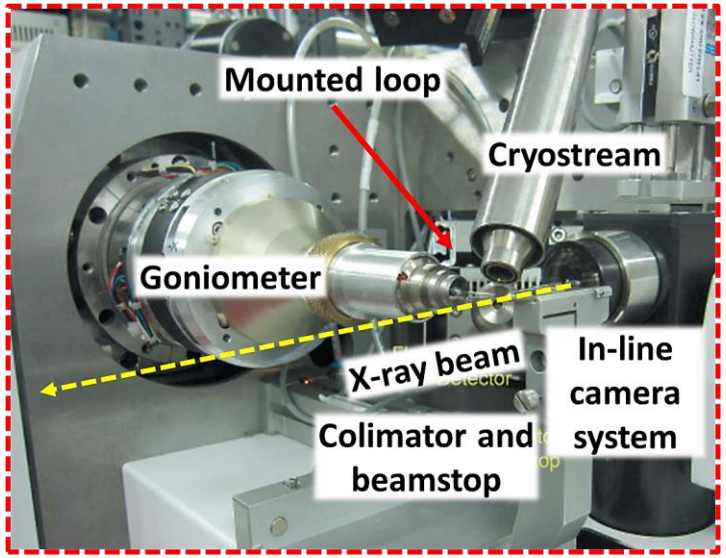
XDS/XDSGUI

Phenix

CCP4

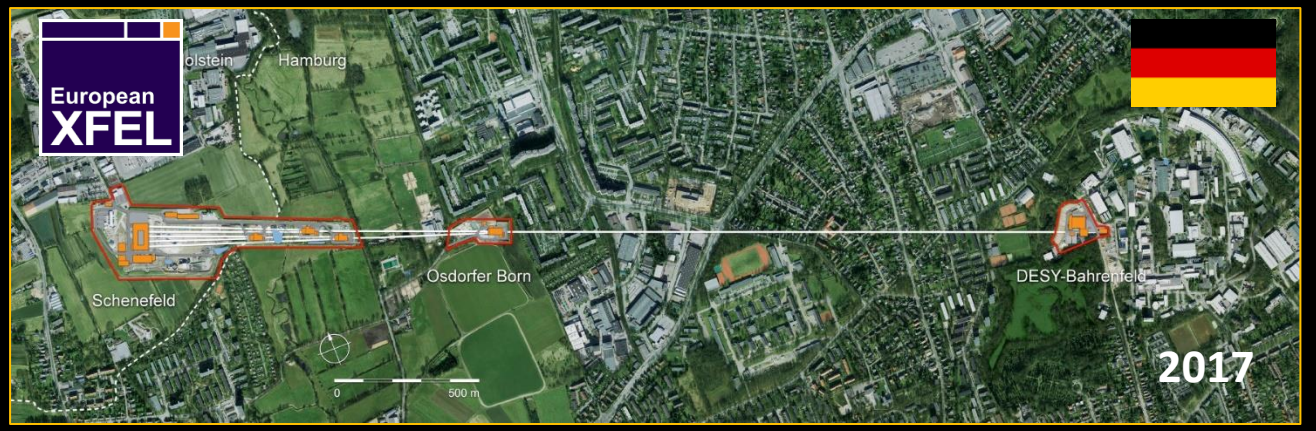


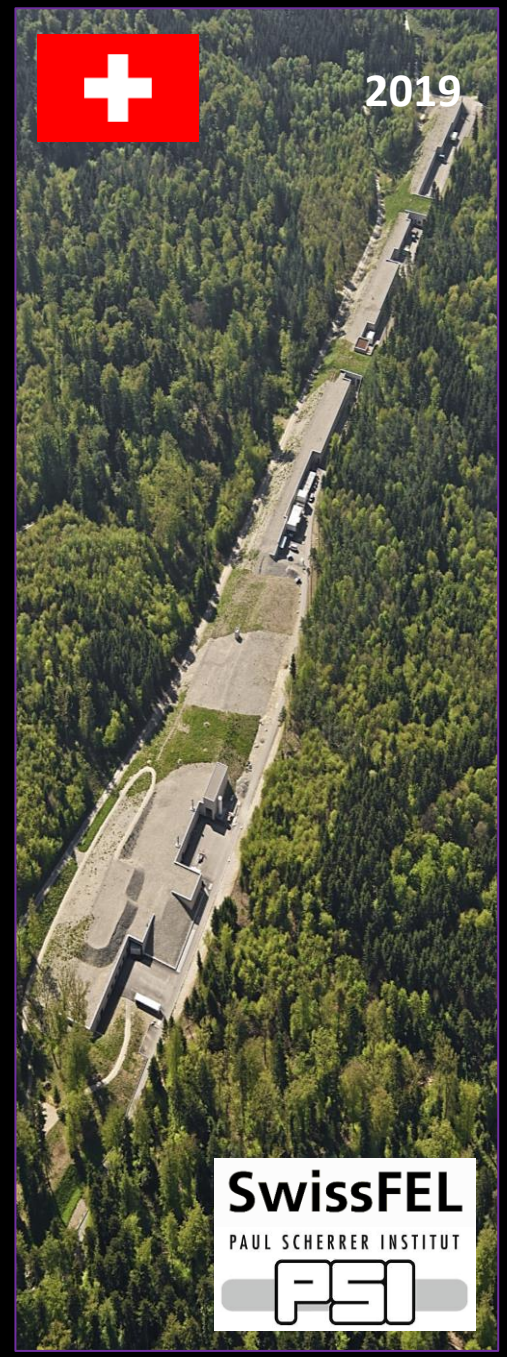




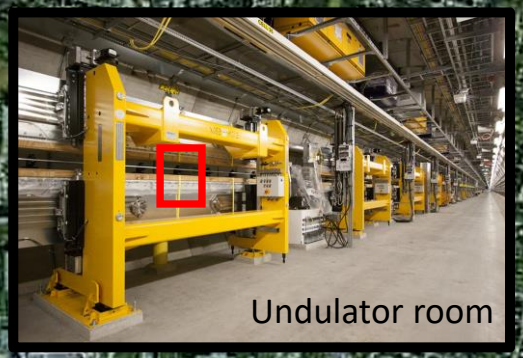
- Large, cryo-cooled crystals
- Measured at synchrotrons
- Remote data collection
- Crystal rotation

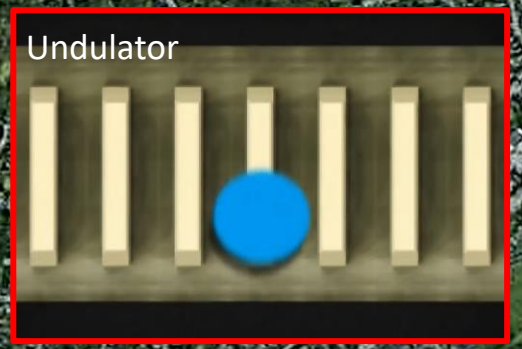
XFELs Worldwide

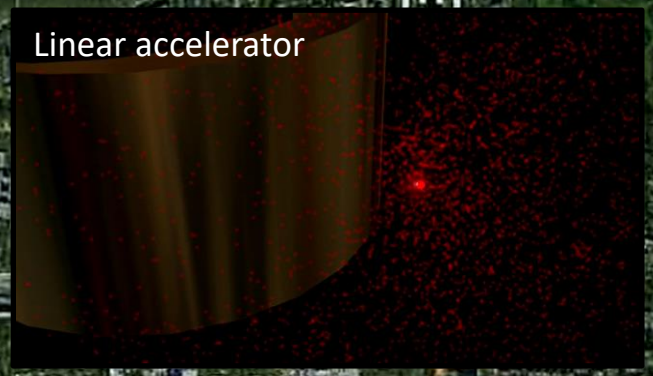




What´s an XFEL?







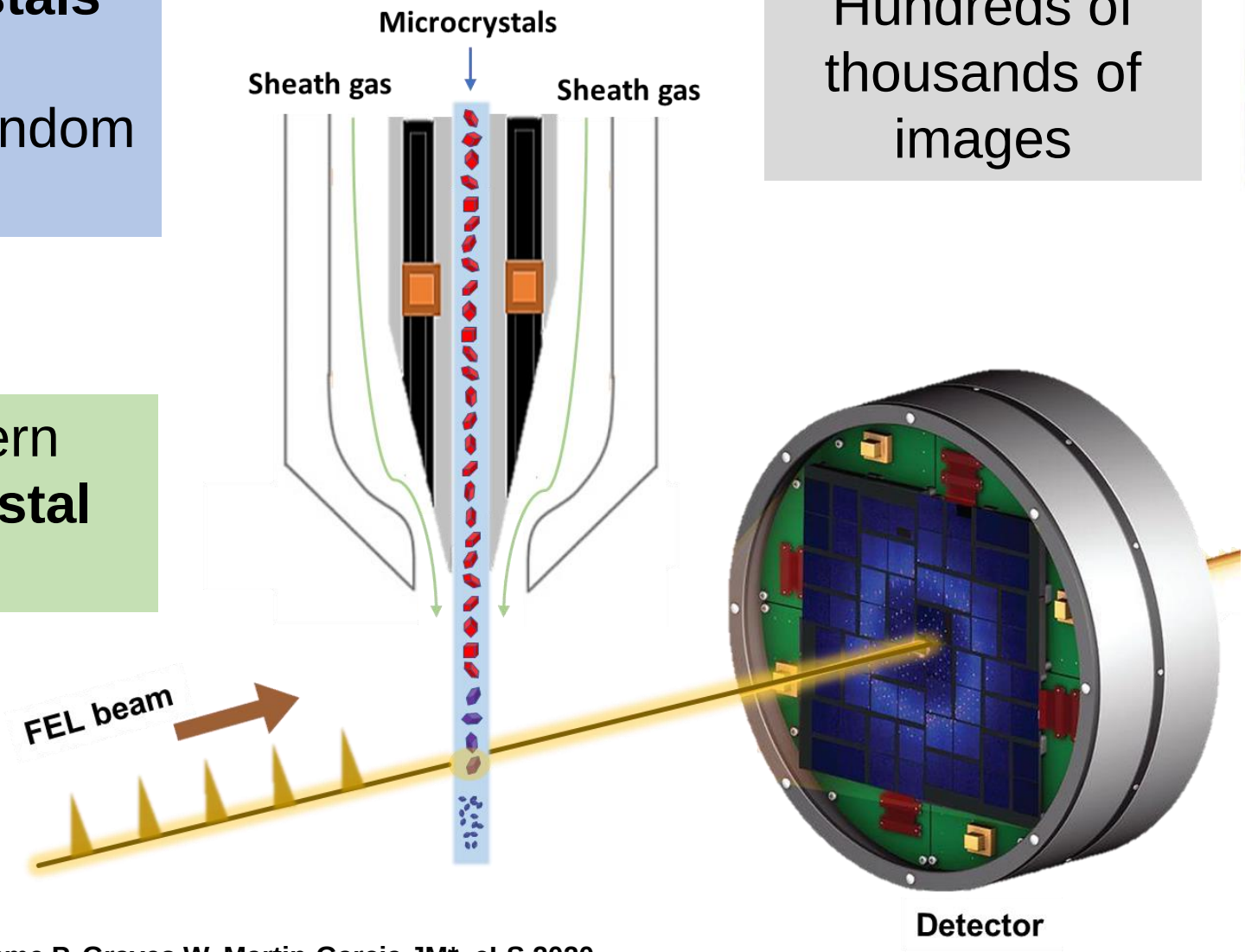
- 5th generation light sources (~ 3Km)
- Extremely bright X-ray pulses
- Ultra-coherent X-rays
- Ultra-short duration (few fs to 100 fs)

Click here to go on a quick virtual tour of the European XFEL!!!

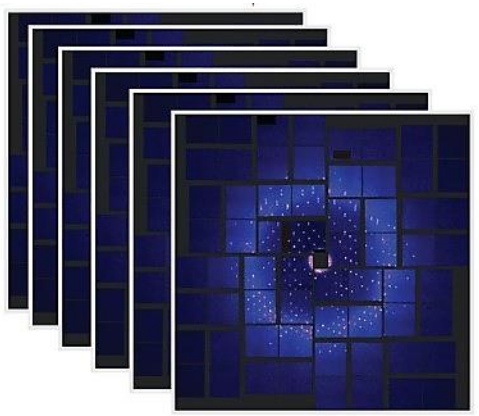
Serial Femtosecond Crystallography (SFX)

Nano/microcrystals injected in serial fashion and in random orientations

A diffraction pattern collected per **crystal** and **X-ray pulse**



Hundreds of thousands of images



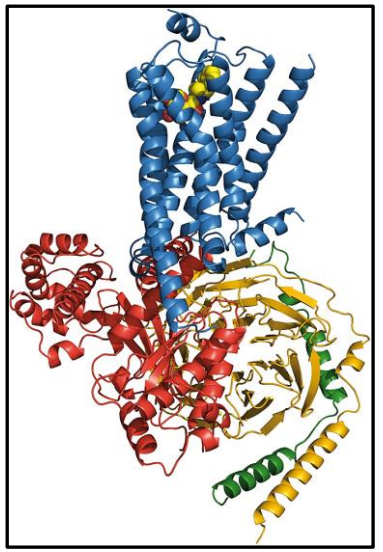


Figure adapted from **Fromme P, Graves W, Martin-Garcia JM*. eLS 2020**

XFELs and instruments available for serial MX experiments

XFEL machine	Beamline	Current beam parameters					Detector / Frame rate	Sample delivery	Sample environment
		keV	Pulses/second	fs	ph/pulse	µm			
LCLS (USA), 2009	**MFX**	5 – 25	120 (1 million soon)	30 – 100	$1{\cdot}10^{12}$ or $1{\cdot}10^{11}$	3-200	ePix10K-2.1M / 120 Hz Rayonix 340 / 30 Hz ePix100 (for XES)	SF-ROX, MESH, GDVN, HVE, DoD, fixed targets	90–278 K Ambient pressure Helium available, humidity controlled
	CXI	6 – 25		30 – 100	$1{\cdot}10^{12}$ or $1{\cdot}10^{11}$	1	Jungfrau-4M / 120 Hz CSPAD 2.3 / 120 Hz	MESH, GDVN, HVE, fixed targets	Vacuum (10^{-5} Torr) Ambient pressure
SACLA (Japan), 2011	**BL3**	4 – 20	30 (60)	2 - 10	10^{11}	~1	MPCCD / 60 Hz	SF-ROX, fixed targets, GDVN, HVE and DoD (in DAPHNIS chamber)	100K - RT Helium at ambient pressure
	BL2	4 – 15		2 - 10	10^{11}	~1	MPCCD / 60 Hz	SF-ROX, GDVN, HVE and DoD (in DAPHNIS chamber)	100K - RT, Helium at ambient pressure
EuXFEL (Germany), 2017	**SPB/SFX**	3 – 16	27000	5 - 300	$1.5{\cdot}10^{12}$	~1 or ~0.1	AGIPD 1M / 4.5 MHz AGIPD 4M / 4.5 MHz (not yet available)	GDVN, aerosol injection HVE, fixted target at lower repetition rate (10 Hz)	Vacuum, $1{\cdot}10^{-6}$ mbar typical, 1.10^{-4} mbar maximum Ambient pressure at low rep rate (second interaction region)
PAL-XFEL, South Korea, 2017	**NCI-SFX**	2.2 – 15	60	25	$3{\cdot}10^{10}$ (mono.) or $1{\cdot}10^{12}$ (pink)	5 × 5	Rayonix MX225-HS / 60 Hz Jungfrau / 60 Hz	GDVN and HVE, fixed targets	RT He at ambient pressure
SwissFEL, Switzerland, 2019	**ALVRA**	2 − 12.4	100	-	$4{\cdot}10^{11}$	2-20	Jungfrau 4M / 100 Hz Jungfrau 16M / 100 Hz	HVE, GDVN (user supplied)	RT Helium at $5{\cdot}10^{-4}$ - 800 mbar

SFX´s brought new challenges to MX:

1- Sample preparation

2- Sample delivery

3- Data collection and evaluation

SFX´s brought new challenges to MX:

1- Sample preparation

2- Sample delivery

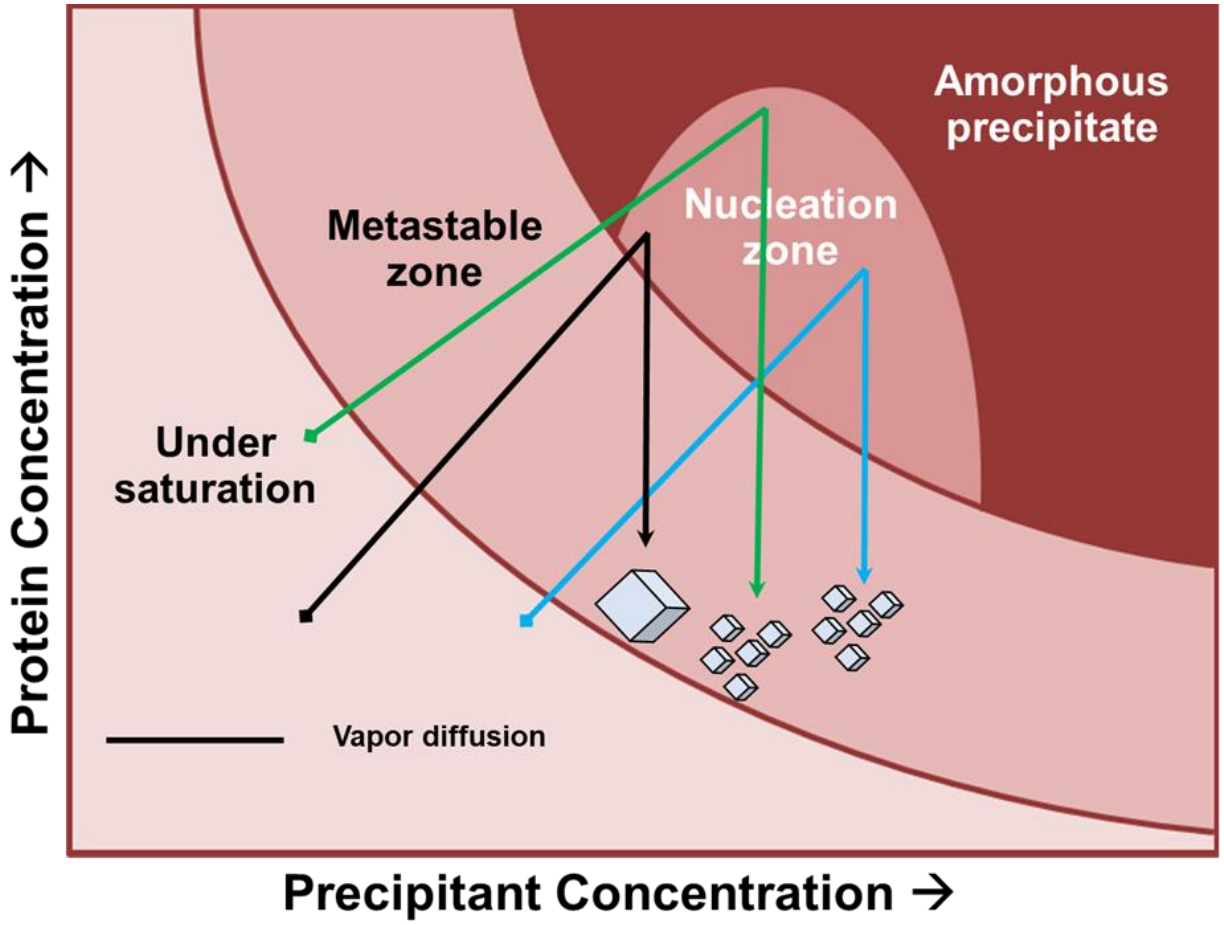
3- Data collection and evaluation

Sample requirements for serial crystallography

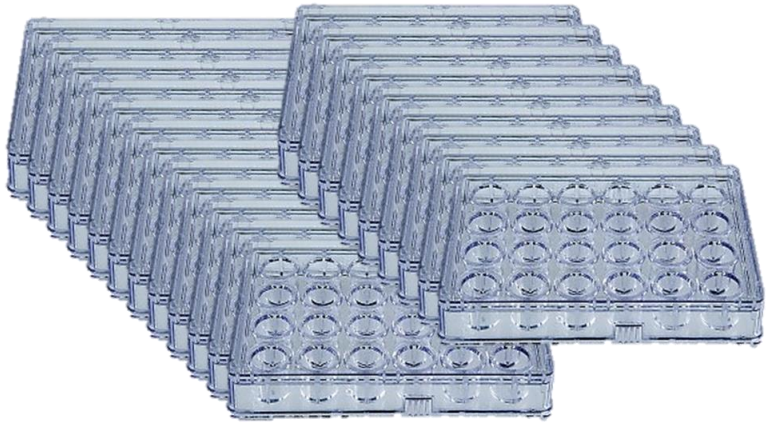
- Large sample volume (from 5-10 μL to 5 mL per experiment)*
- High crystal density (10^9 - 10^{12} crystals/mL)
- Narrow size distribution
  - ❑ Data scaling issues
  - ❑ Injector clogging issues
  - ❑ Ideal for time-resolved studies

** This depends on what kind of data are to be acquired and how (slides ahead).*

Is vapor diffusion suitable?

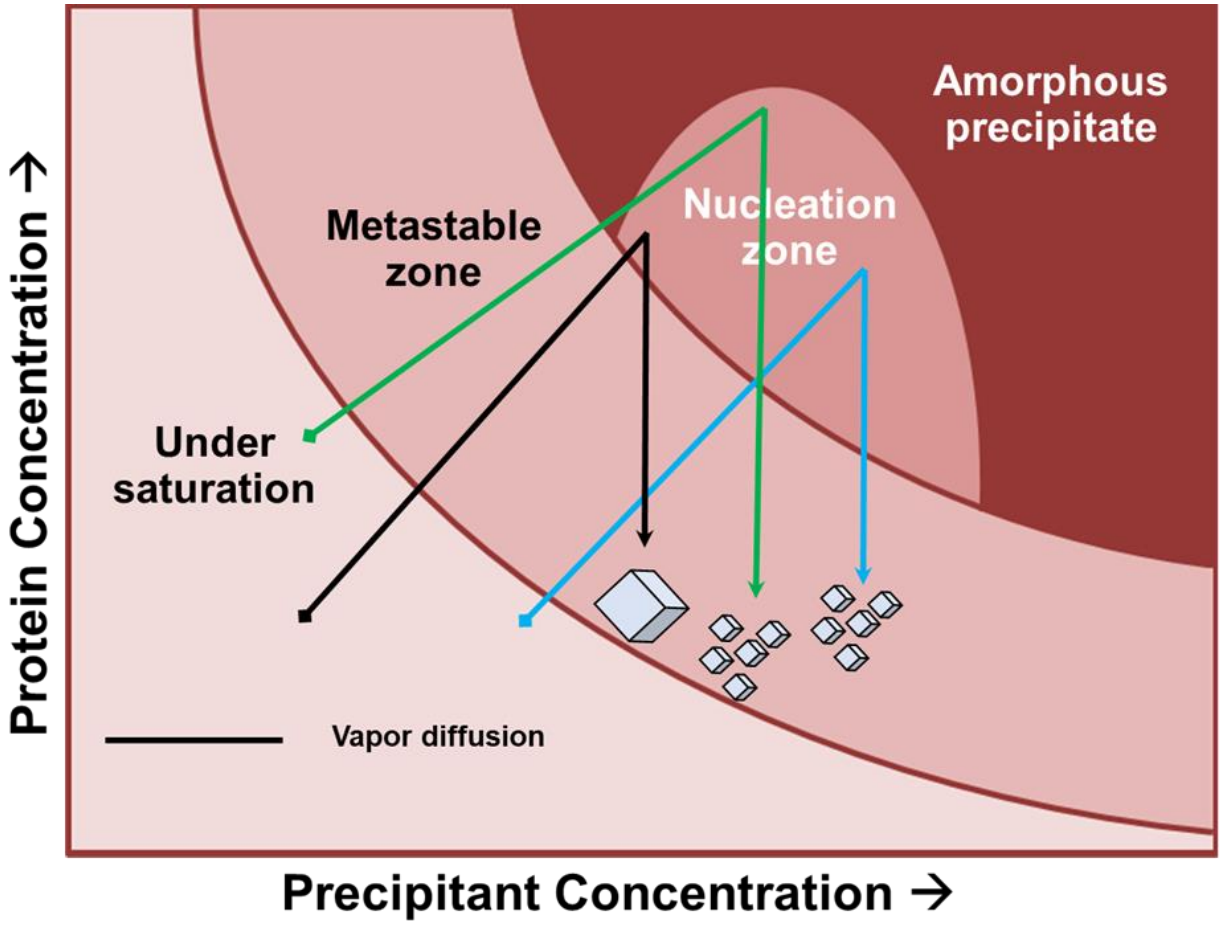


- Vapor diffusion does not require any specialist hardware beyond that found in a normal crystallography lab.
- It is extremely tedious => hundreds of plates

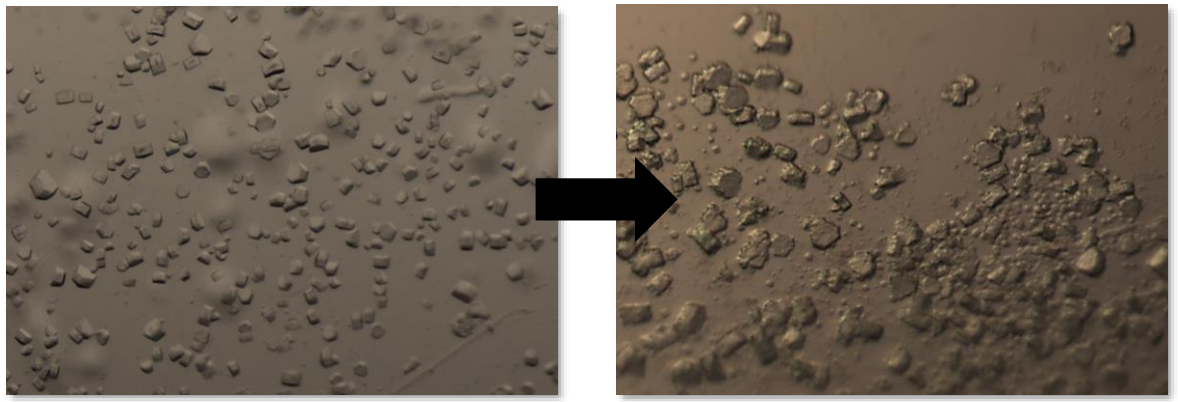


Is vapor diffusion suitable?

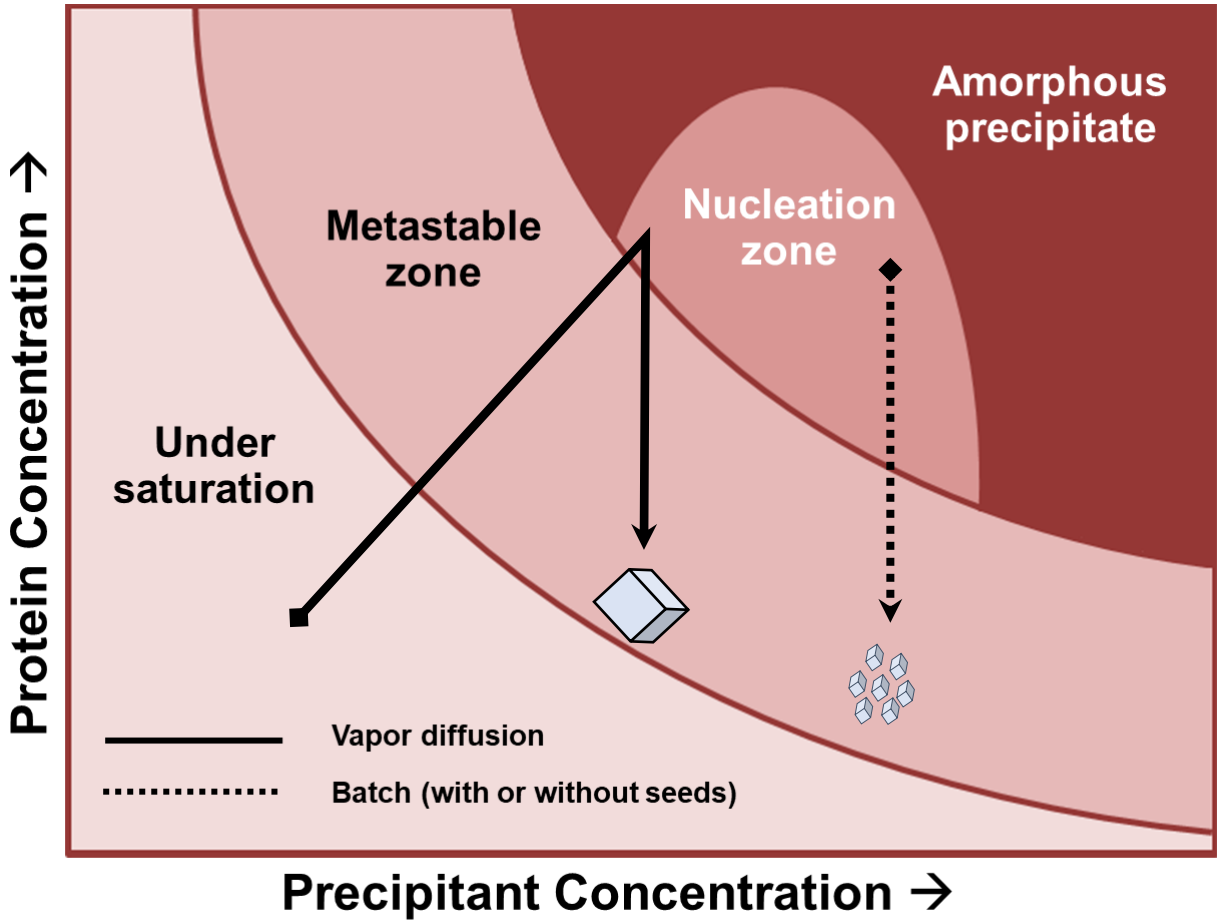
It´s not!!



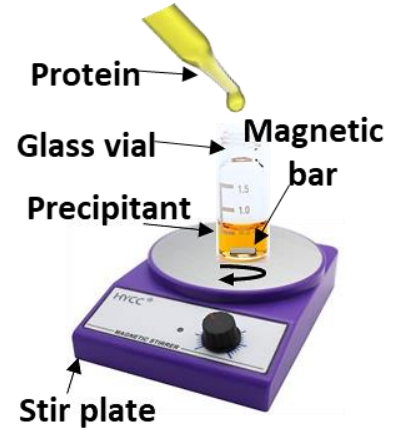
- Vapor diffusion does not require any specialist hardware beyond that found in a normal crystallography lab.
- It is extremely tedious => hundreds of plates
- Harvesting and pooling of drops from 24-well plates can be detrimental to crystals which are sensitive to handling.

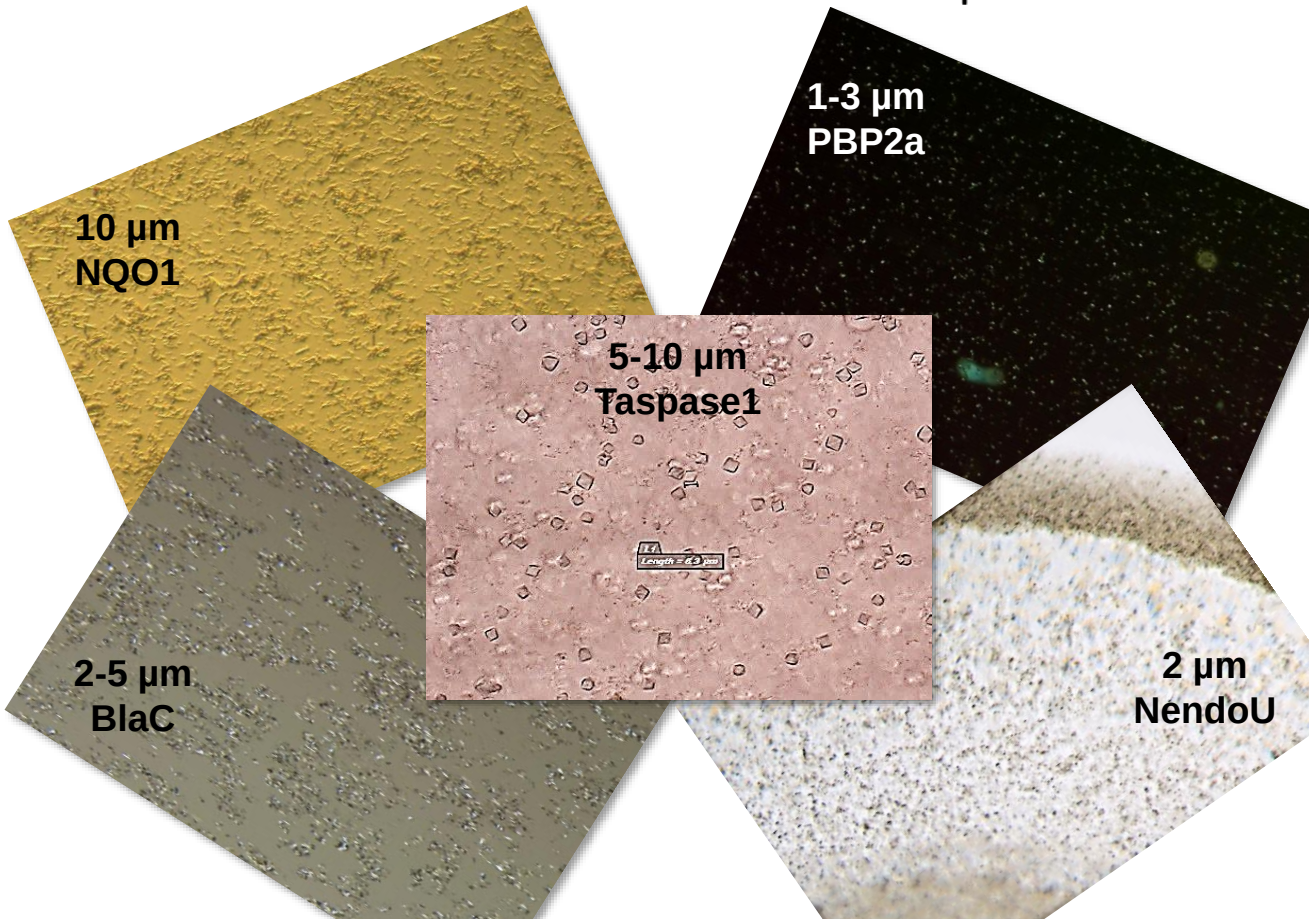


The batch with agitation method

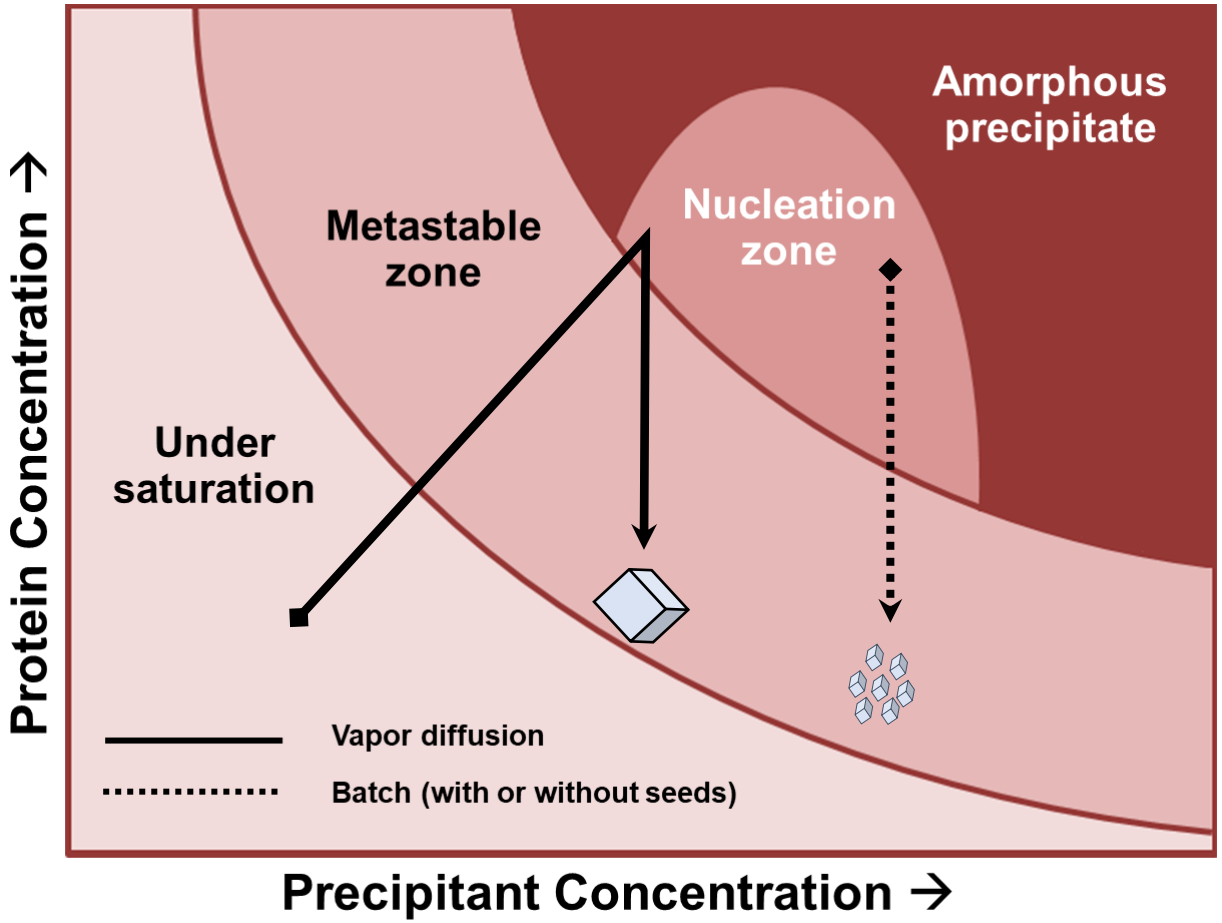


- Best method for scaling up to larger sample volumes (500 µL – 5 mL)
- Protein and precipitant solutions are quickly mixed/homogenized

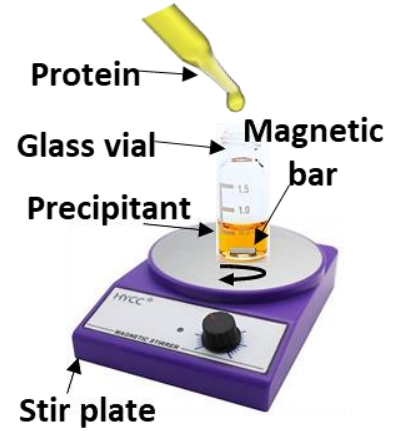


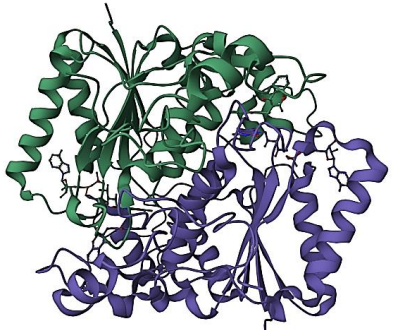


The batch with agitation method

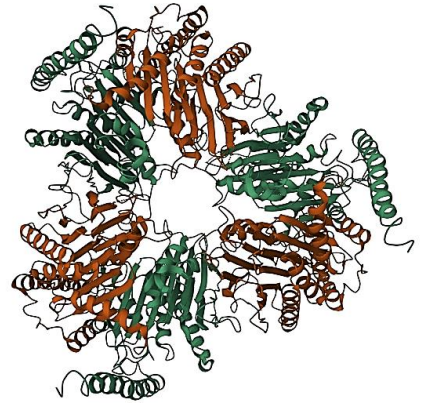


- Best method for scaling up to larger sample volumes (500 µL – 5 mL)
- Protein and precipitant solutions are quickly mixed/homogenized

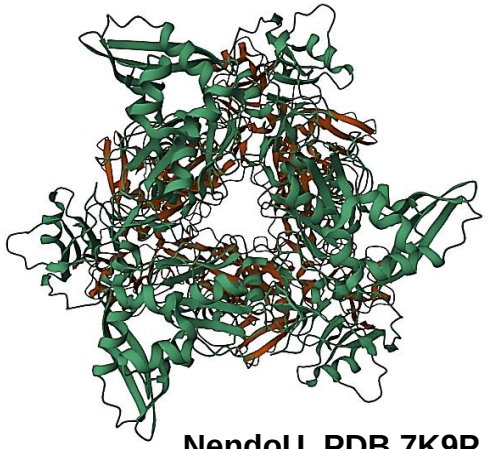




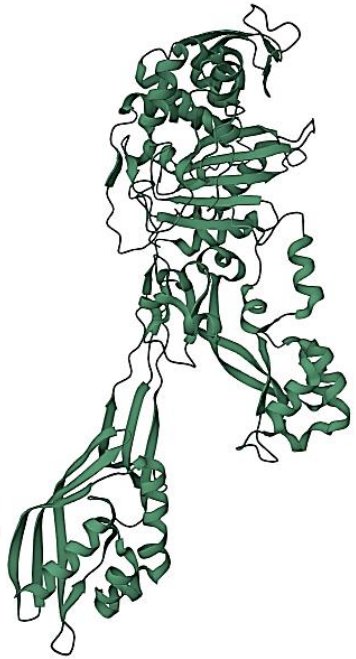
NQO1, PDB 8C9J



Taspase1, PDB 6VIN



NendoU, PDB 7K9P



PBP2a, Unpublished

BlaC, PDB 6B6E

SFX´s brought new challenges to MX:

1- Sample preparation

2- Sample delivery

3- Data collection and evaluation

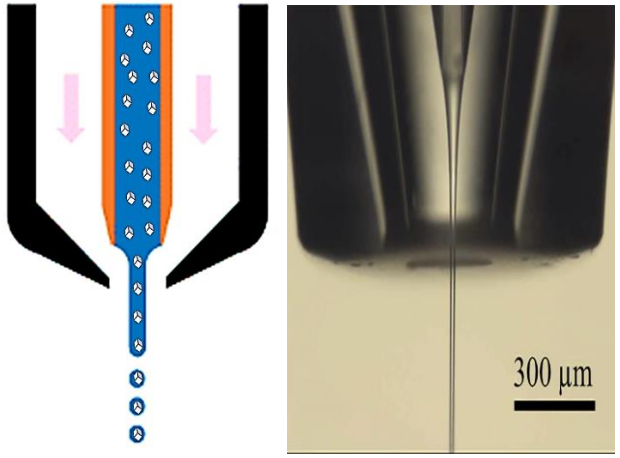
An ideal sample delivery method MUST…

- Produce minimal background scattering on detector
- Replenish crystals quickly
- Reduce sample consumption
- Be reliable without human intervention for many hours

Type of sample delivery methods

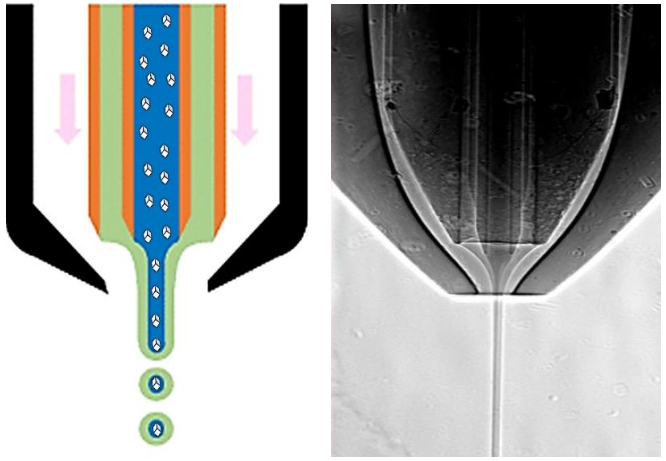
Liquid injectors

GDVN



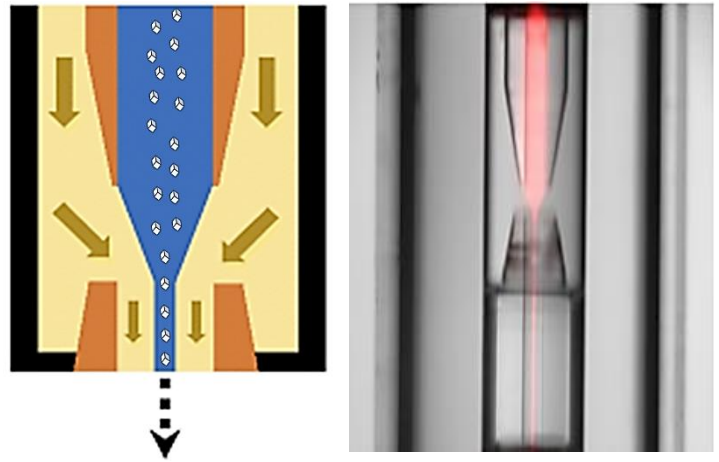
DePonte, et al., J. Phys. D Appl. Phys. 2008

DFFN



Oberthuer D, et al., Sci Rep. 2017

Mixing



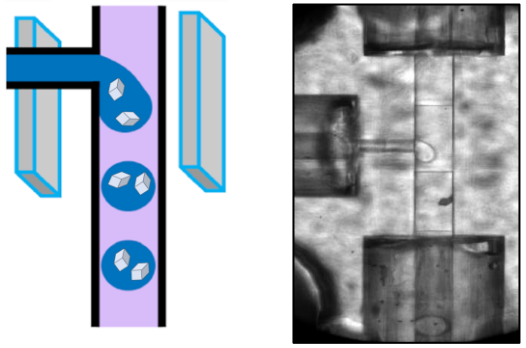
Calvey GD, et al., Struct Dyn. 2016

- Capillary-in-capillary devices that use a high-pressure gas sheath from the outer capillary to focus the sample into a jet of a few micrometres in diameter.
- They are rapid flow injectors that are ideal for TR experiments and ideal to be used at high-rep rate XFELs such as EuXFEL and the upcoming LCLS-II.
- Drawbacks: due to the high flow rate at which they operate: 1) consume a lot of sample; 2) low hit rate; 3) high risk of clogging.

Type of sample delivery methods

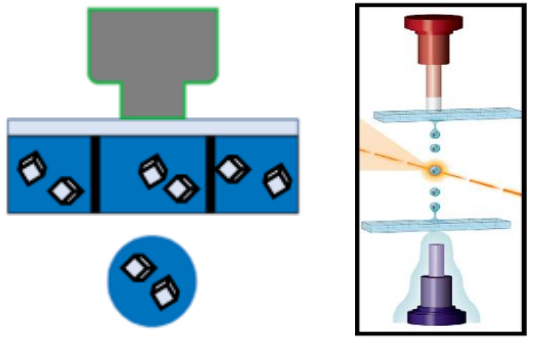
Pulsed injectors

Segmented droplet generator



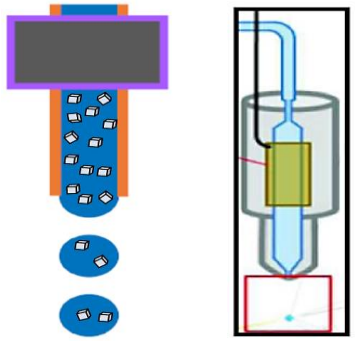
Kim D, et al., Anal Chem. 2019

Acoustic droplet ejection



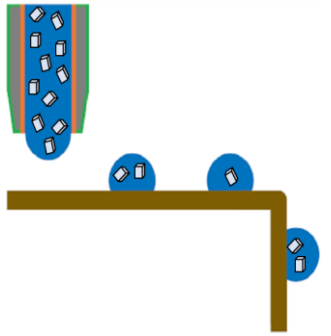
Roessler CG, et al., Structure. 2016

Piezoelectric droplet ejection



Mafune F, et al., Acta Cryst. D. 2016

Drop-on-tape



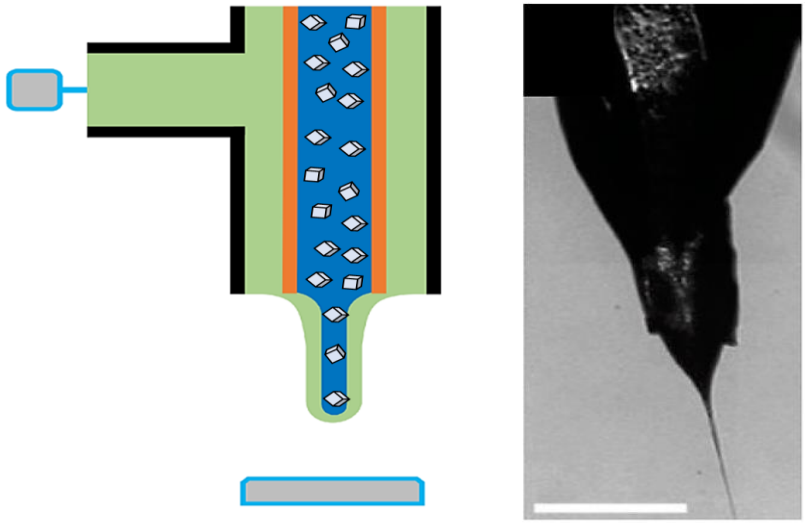
Beyerlein KR, et al., IUCrJ. 2017

- They reduce sample consumption by encapsulating crystals into droplets and then synchronizing them with the X-ray pulses.
- Droplets can be generated by applying electrical fields, acoustic fields and oil.
- They are suitable for TR experiments both at XFELs and synchrotrons.

Type of sample delivery methods

Electrospinning injector

Microfluidic electrokinetic sample holder (MESH)



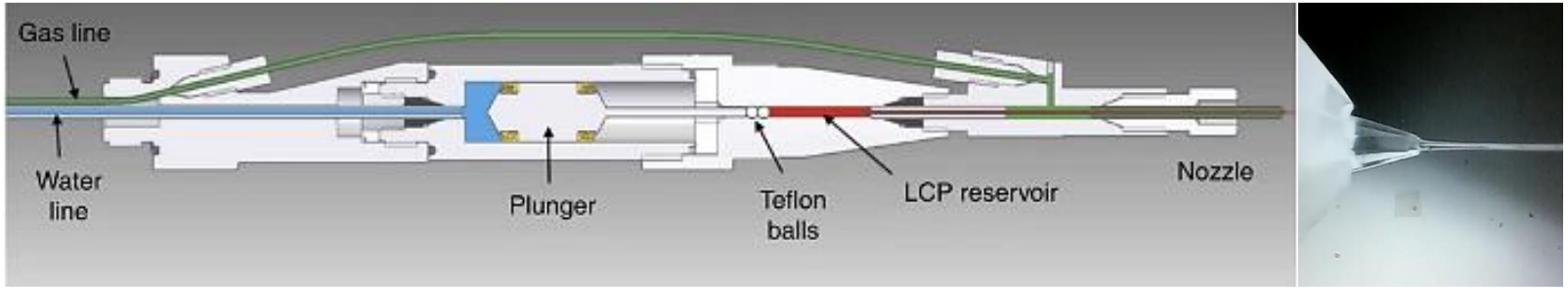
Sierra RG, et al., Acta Crystallogr D Biol Crystallogr. 2012

- Low-flow injector that overcomes sample clogging and high sample consumption by applying a high-voltage to the sample.
- It is compatible with both liquid and viscous samples.

Type of sample delivery methods

Viscous-media injectors

High-viscosity injector (LCP injector)

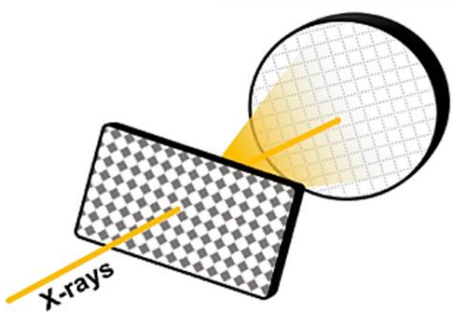


Weierstall U., et al., Nature Communications. 2014

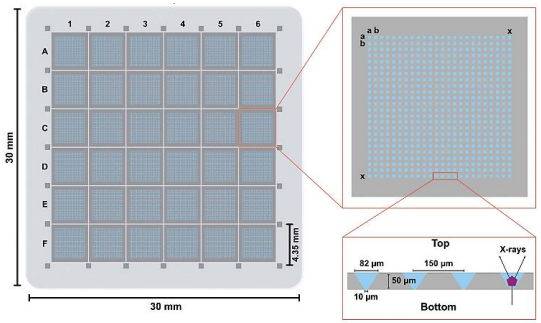
- Reduces sample consumption and clogging issues by mixing crystals with a viscous matrix.
- Originally developed to deliver MP samples in LCP to XFELs it is also suitable for soluble proteins for which novel viscous matrices have been developed: agarose, PEG 8M, and grease among some others.
- They are run at much slower flow rates, which makes it ideal for serial crystallography experiments at synchrotrons.

Type of sample delivery methods

Fixed targets

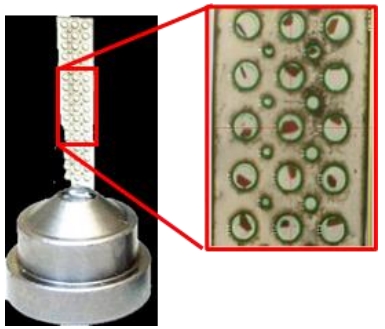


Chips



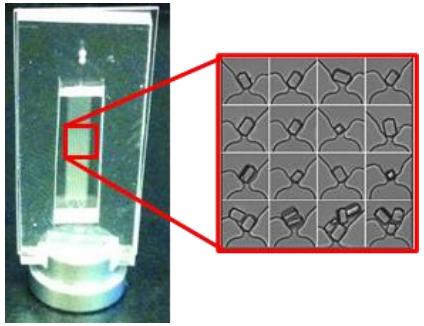
Mehrabi P, et al., J. synchr. radiat. 2020

Grids



Cohen AE, et al., PNAS. 2014

Microfluidic devices



Mueller C, et al., Struct Dyn. 2015

- Crystals are deposited on a solid matrix.
- This is the method that consumes less sample.
- Suitable for TR experiments specially in pump-probe experiments
- Ideal for determining the structures of proteins in steady state.
- Fixed targets can be also used to grown crystals in situ.
- This is the method of choice at synchrotrons.
- They can be used both at RT and in cryo-conditions.

Which sample delivery method is best?

- Sample characteristics (fluidity, crystal size/distribution)
- Sample volume/concentration
- Objectives of the experiment (structure of apo protein, complex, reaction mechanism….)

PCS beamtimes are offered by all XFELs and ~~strongly recommended~~ “mandatory” for assessing crystal and injection quality.

Recommended: Kupitz C, Sierra RG. Preventing Bio-Bloopers and XFEL Follies: Best Practices from your Friendly Instrument Staff. *Crystals*. 2020

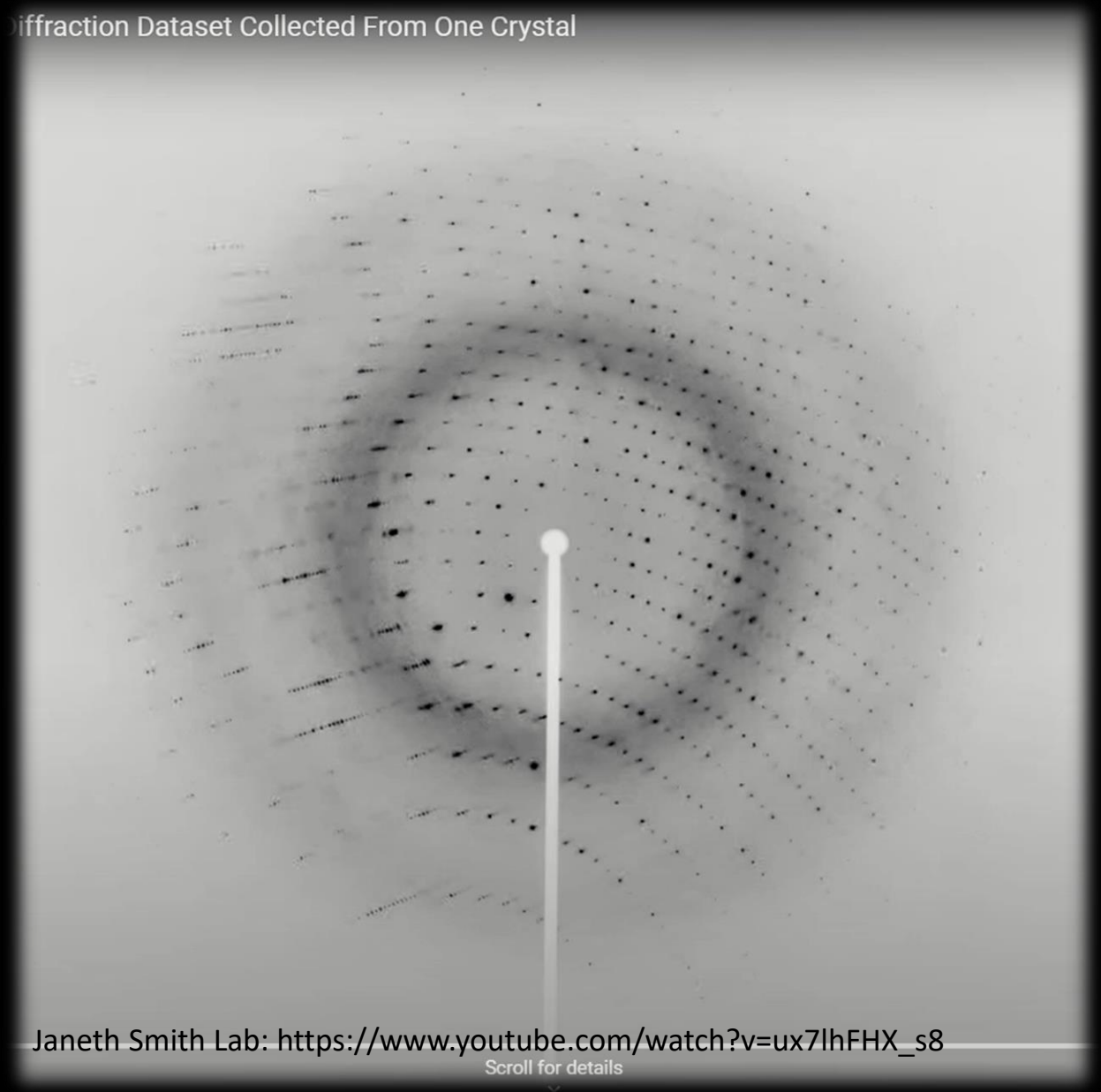
SFX´s brought new challenges to MX:

1- Sample preparation

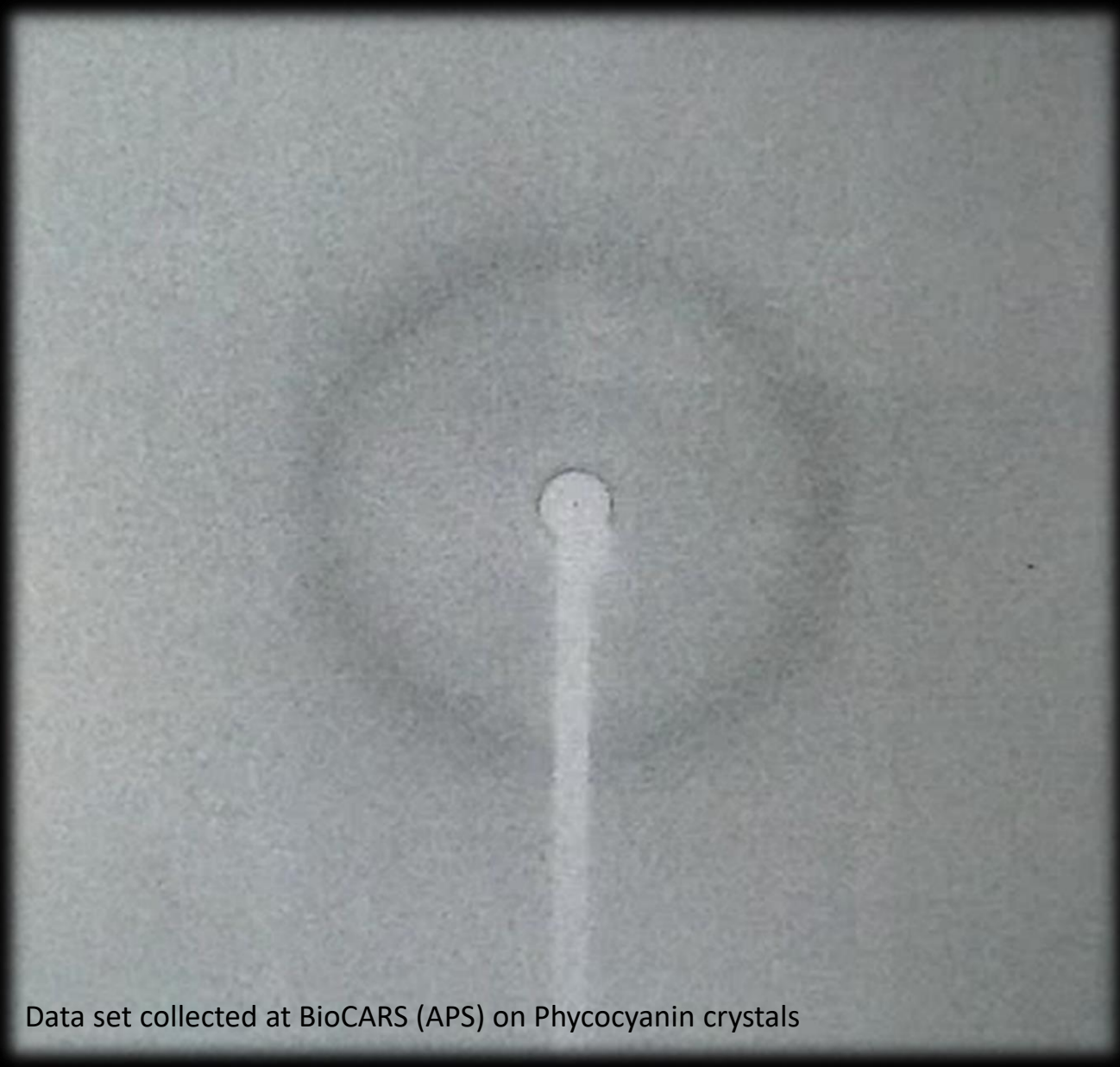
2- Sample delivery

3- Data collection and evaluation

Classical rotation vs serial data collection



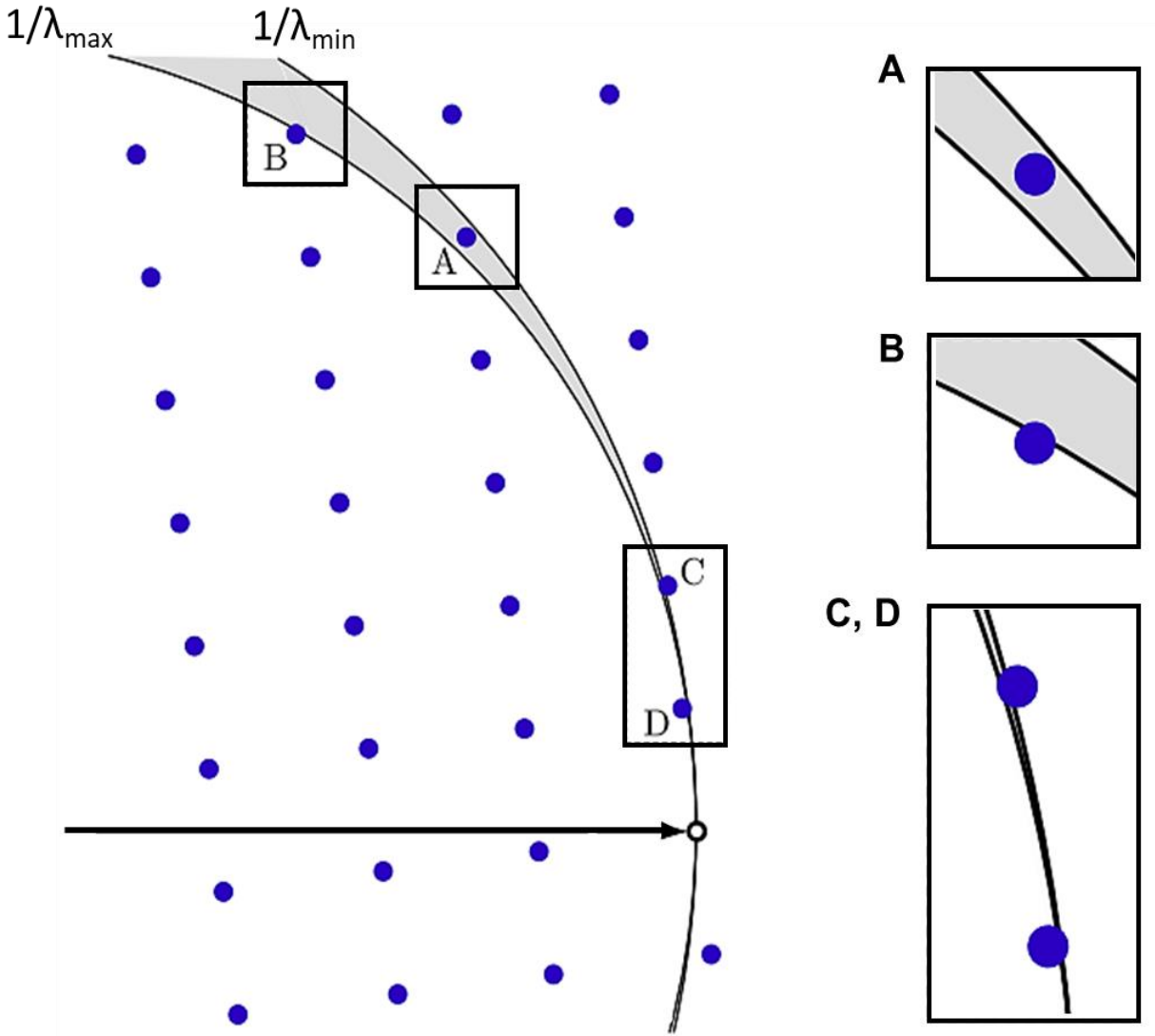
Janeth Smith Lab: https://www.youtube.com/watch?v=ux7lhFHX_s8



Data set collected at BioCARS (APS) on Phycocyanin crystals

The partiality problem in X-ray crystallography

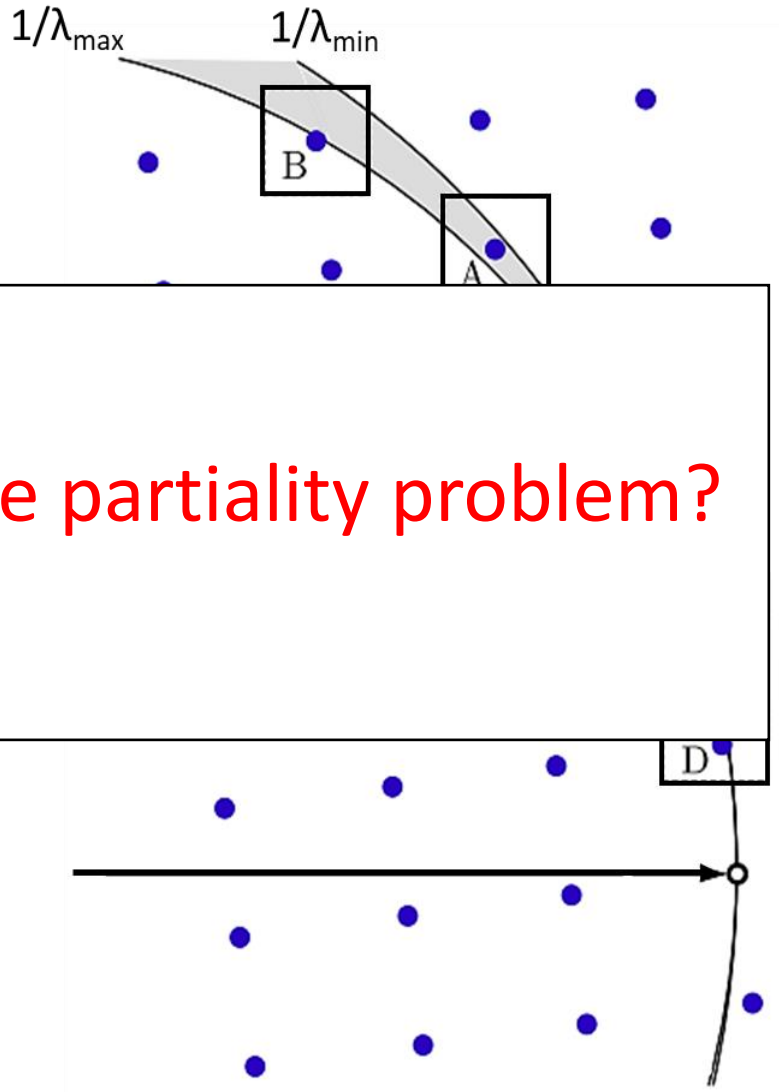
- X-rays we use to measure our samples, are not fully monochromatic.
- Ewald sphere isn´t thin but has a certain thickness.
- In a real experiment only those reflection nodes of the reciprocal space that cross the Ewald sphere are recorded.
- Only partial reflections are measured.

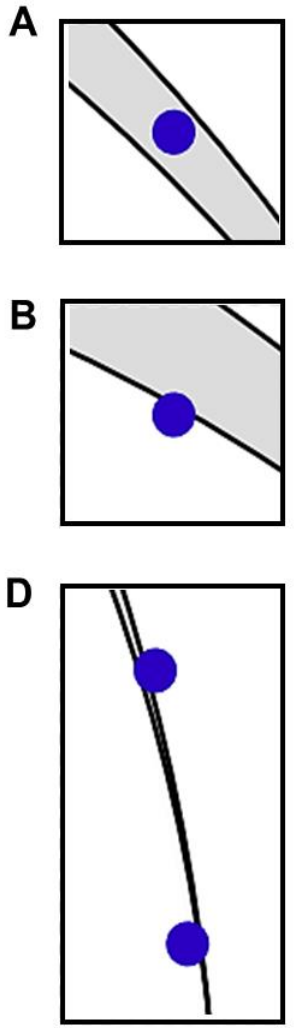


The partiality problem in X-ray crystallography

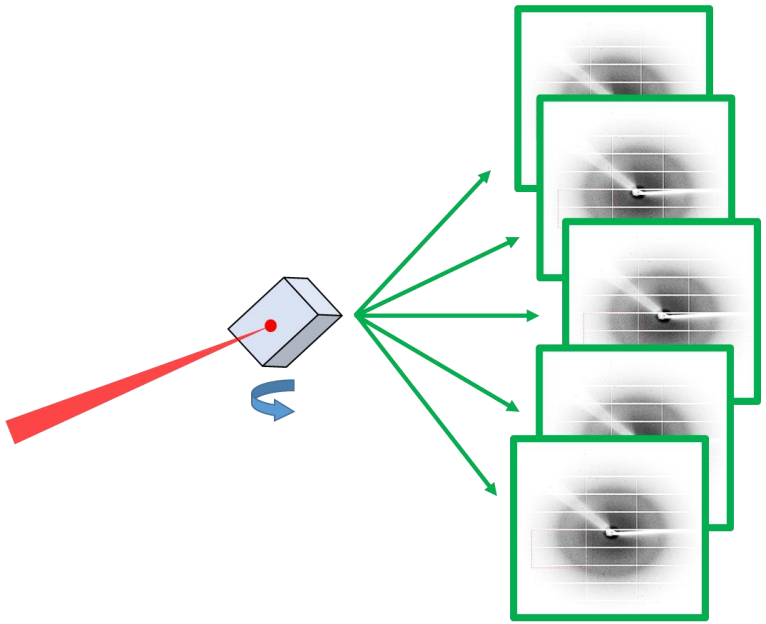
- X-rays we use to measure our samples, are not fully monochromatic.
- Ewald sphere is thickness.
- In a real experi nodes of the reciprocal space that cross the Ewald sphere are recorded.
- Only partial reflections are measured.

How do we solve the partiality problem?

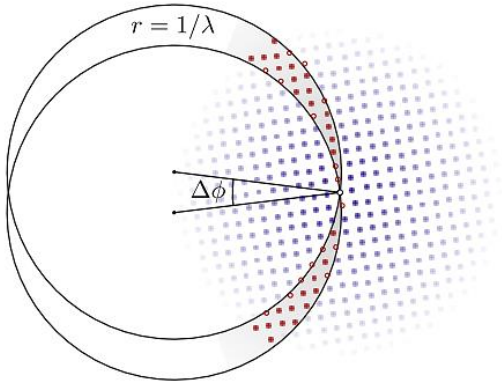




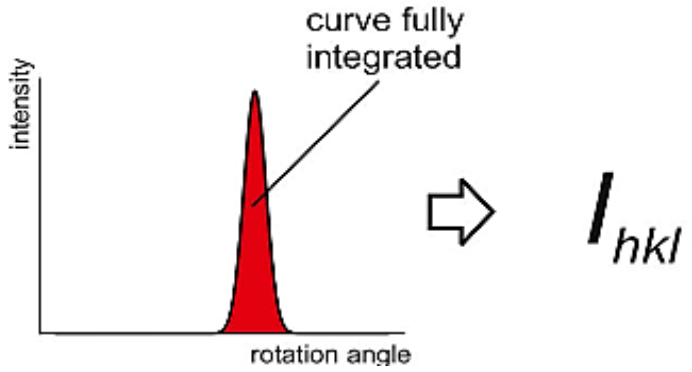
By rotating a single crystal while collecting data



Real space



Reciprocal space

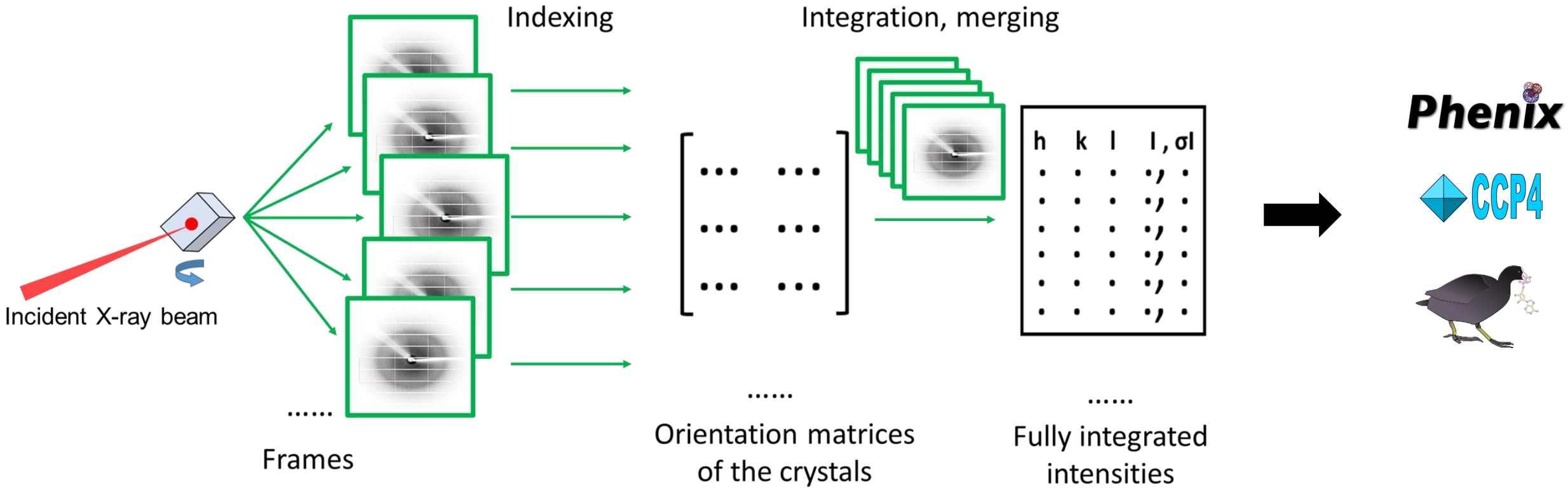


- A diffraction pattern is recorded for each small angle increment $\Delta\varphi$, typically 0.1 - 1 degree.

- Each reciprocal lattice node will cross the Ewald sphere completely and its full diffraction intensity recorded either in one rotation pattern or over several consecutive patterns

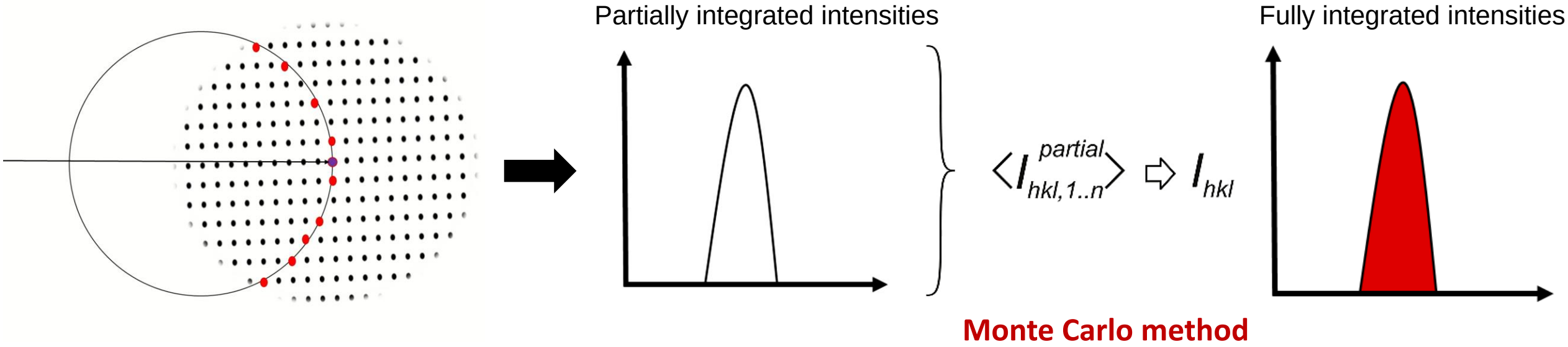
Data processing in classical rotation mode

XDS/XDSGUI



- A range of diffraction patterns (~1000 frames or even less) is available for each crystal, and each one usually contains many diffraction spots.
- The rotational relationship between the patterns is known exactly.
- Indexing is, typically, a trivial exercise in most instances.

By averaging tens of thousands of diffraction patterns collected serially



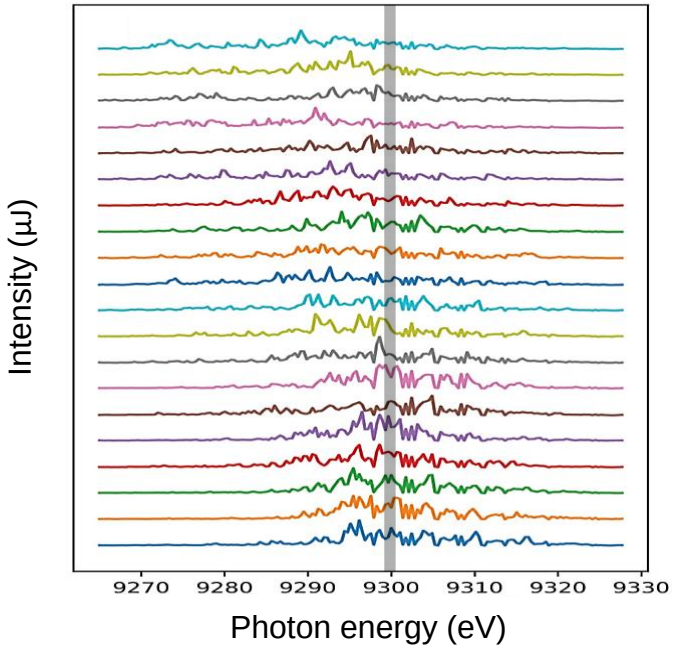
- In SFX experiments, the crystals essentially do not rotate during the exposure, so only a very small portion of the reciprocal space is sampled, so that reflections from each diffraction pattern are generally only partially recorded.

- To obtain the full intensities, we average the integrated intensities of a sufficiently large number of measurements so that the entire reciprocal space is eventually sampled.

Other considerations for SFX data processing

- Due to the stochastic nature of self-amplified spontaneous emission (SASE), there are shot-to-shot variations in XFEL pulses (intensity, energy and bandwidth).

  Naresh Kujala N, et al. *Rev Sci Instrum.* 2020. https://doi.org/10.1063/5.0019935



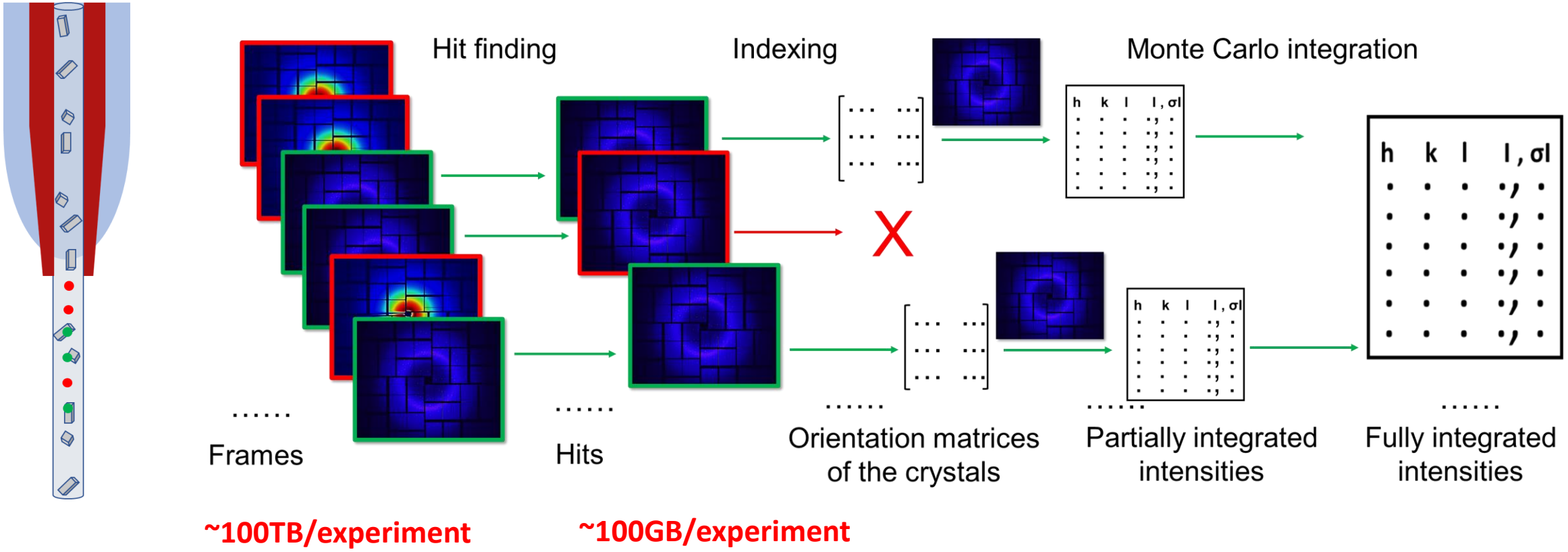
- Crystals size distribution => scaling issues

- Parameters such as **crystal-to-detector distance**, **X-ray wavelength** and **geometry of the X-ray detector** are not known precisely. A well-diffracting calibration sample such as lysozyme is required.

The AGIPD detector at the EuXFEL

- Some detectors used at XFELs consist of panels that can be moved relative to each other.

Data processing in SFX



- In SFX, not every diffraction pattern contains crystal diffraction (0.5%-10%). To reduce the disk space required to store the data and to speed up the subsequent processing, the first step in the analysis is to sort out images containing crystal diffraction by the, so called, hit finding step (“cleaning” step).
- Due to the lack of crystal rotation, there may be relatively few diffraction spots to determine crystal orientation => many diffraction patterns may be unindexable, increasing the number of crystals needed to obtain sufficient indexed diffraction patterns for Monte Carlo convergence.

Data analysis workflow SFX

Detector

Raw XTC image files containing X-ray pulse parameters, pump laser signal, diagnostics, motor positions, etc.

on-line

OnDA Monitor (OM)

1.Online monitoring
2.Live hit rate and resolution estimate
3.Live saturating pixel tracking

off-line

on-line

Cheetah

1.Hit finding (data reduction)
2.Background subtraction
3.Clean diffraction patterns and meta data saved as HDF5 or CXI
4. Statistics and preliminary analysis

on-line / off-line

CrystFEL

1.Indexing
2.Integration
3.Merging
4.Post refinement

CCP4

Phenix

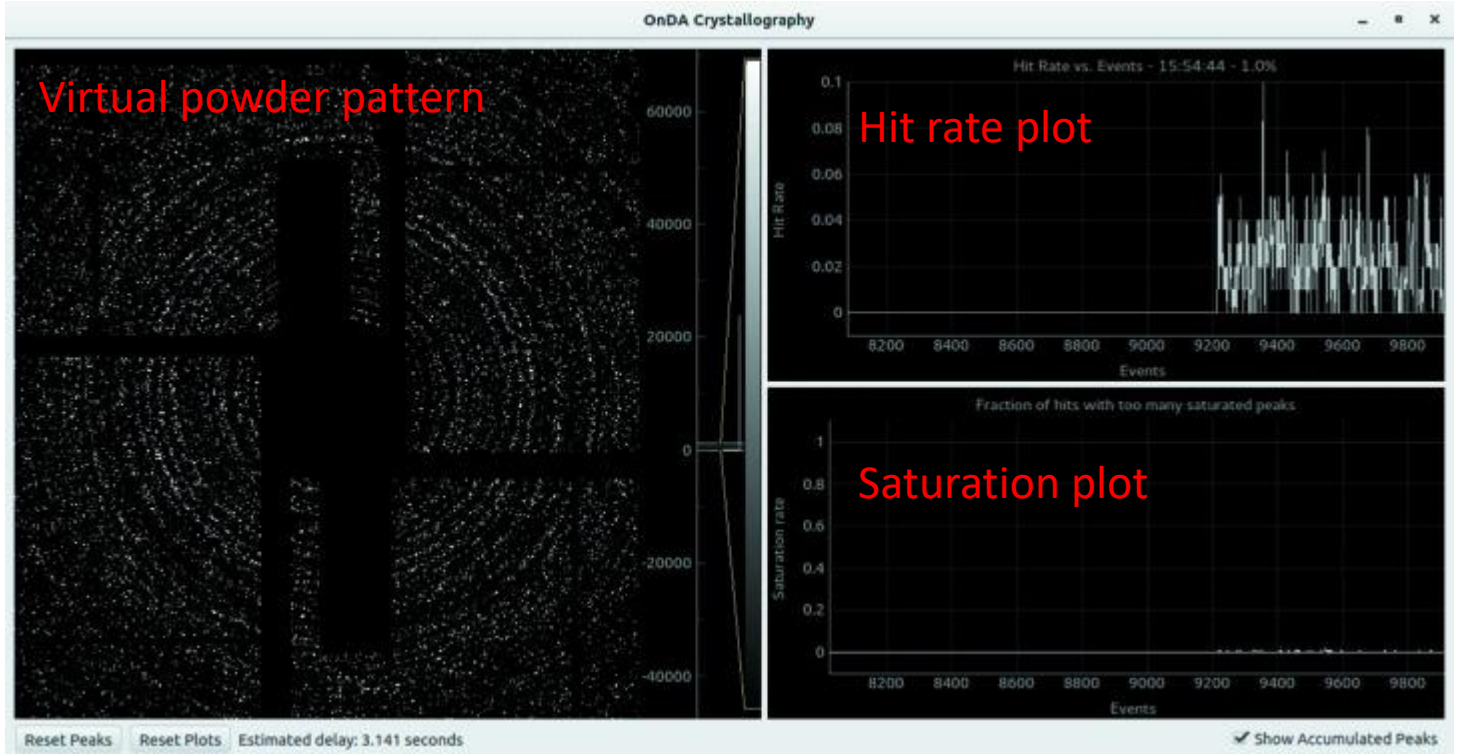
Phasing, building, refinement, validation

RCSB PDB PROTEIN DATA BANK

OnDA Monitor (OM) program for *real time* data evaluation

- **Fast online feedback** during serial X-ray diffraction and scattering experiments through real-time monitors.
- The SFX monitor allows users to keep hit rate and saturation rate under control.
- It also uses the collected data to display a constantly updating virtual powder pattern.
- The peakfinder algorithm from the Cheetah program is used to detect potential Bragg peaks.

Graphical user interface for the *OnDA* crystallography monitor



https://www.ondamonitor.com/html/index.html

Mariani V., et al., J. *Appl. Crystallogr.* 2016

Cheetah program for data reduction

1. Hit finding: Data reduction

- Detector corrections: identify and flag bad and saturated pixels, dark and common mode correction of each module and individual gain corrections for each pixel

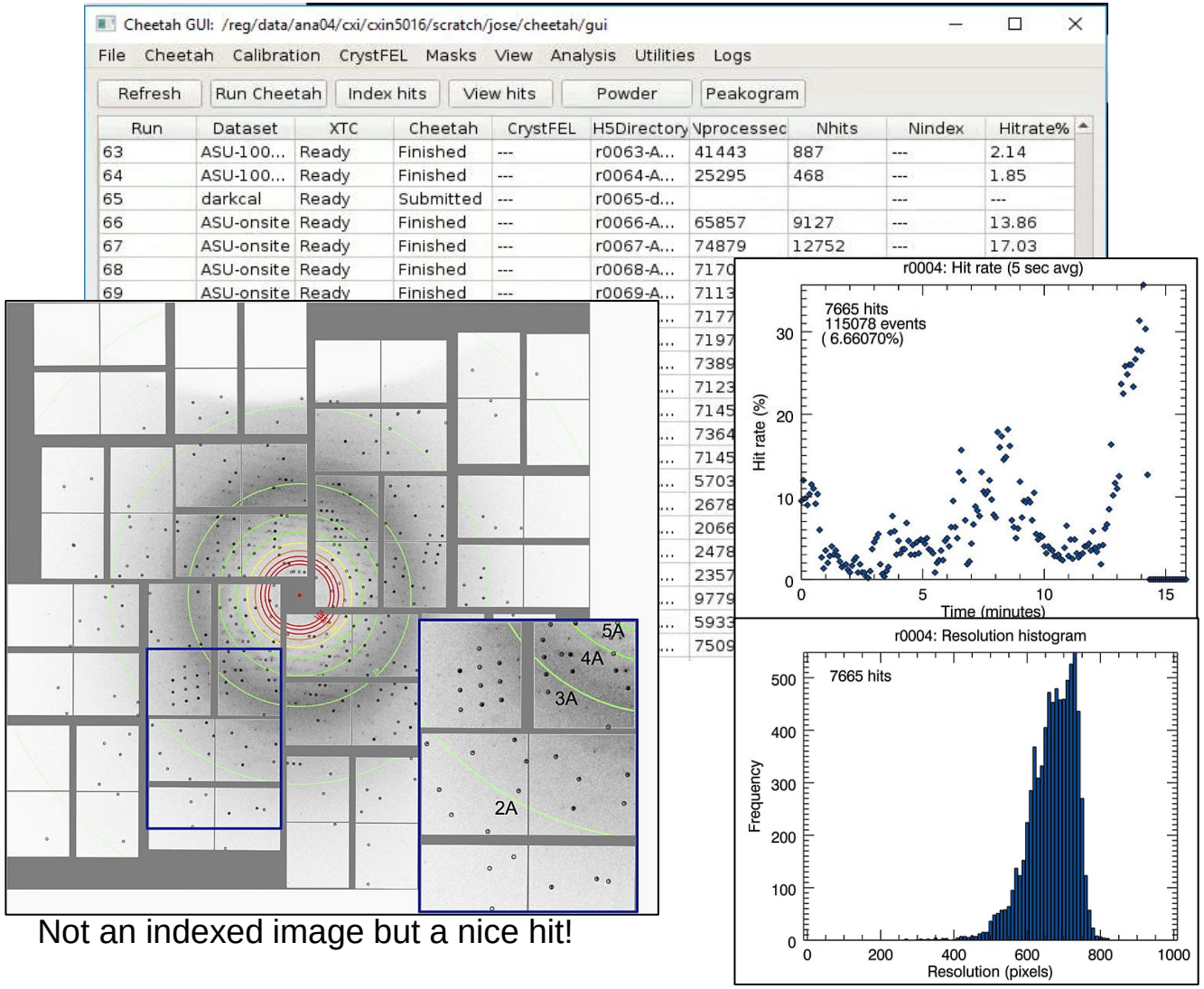
- Search for Bragg peaks in every image using the algorithm *peakfinder8*.

2. Data translation

- XTC data is converted to a facility independent format (HDF5 or CXI)

3. Data organization

- Summarize statistics per run.



Not an indexed image but a nice hit!

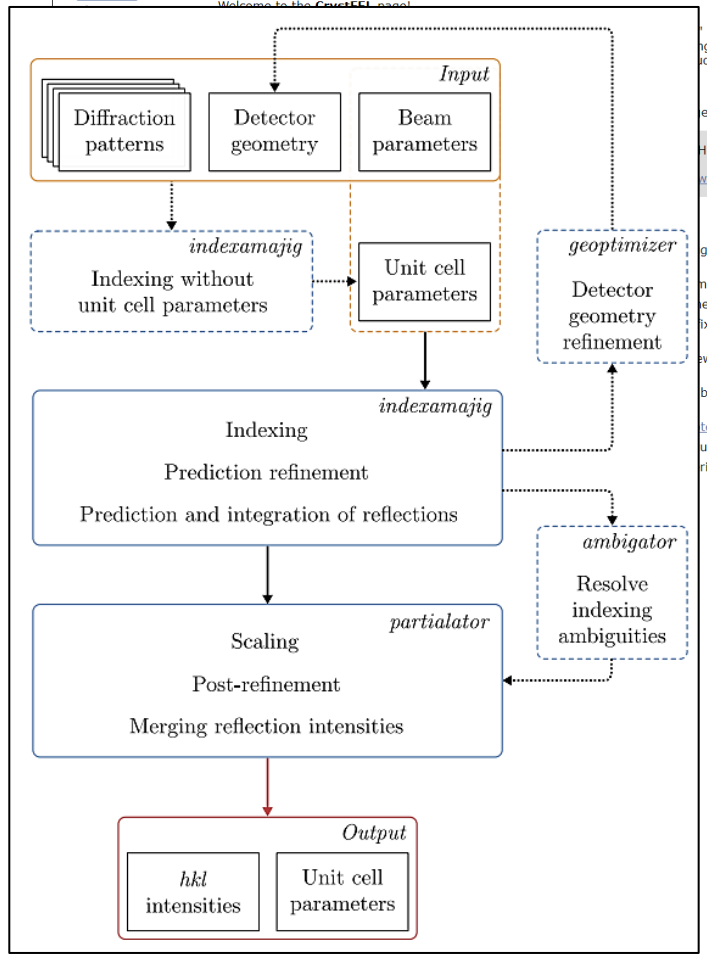
https://www.desy.de/~barty/cheetah/Cheetah/Welcome.html

Barty, A. et al. J Appl. Cryst. 2014

CrystFEL: data processing for serial crystallography

CrystFEL version 0.10.2 released in February 2023

- Tools for **indexing** and **integrating** diffraction patterns, **scaling** and **merging** intensities
- Simulating patterns, calculating figures of merit for the data and visualizing the results.
- CrystFEL converts hkl files into mtz files that can be fed into Phenix or CCP4
- Tutorials and webinars available



https://www.desy.de/~twhite/crystfel/index.html

White, TA. et al. J Appl. Cryst. 2012

Why microcrystals and serial data collection instead of large crystals and rotation data acquisition?

Nanocrystals and microcrystals are gaining popularity within the structural biology community.

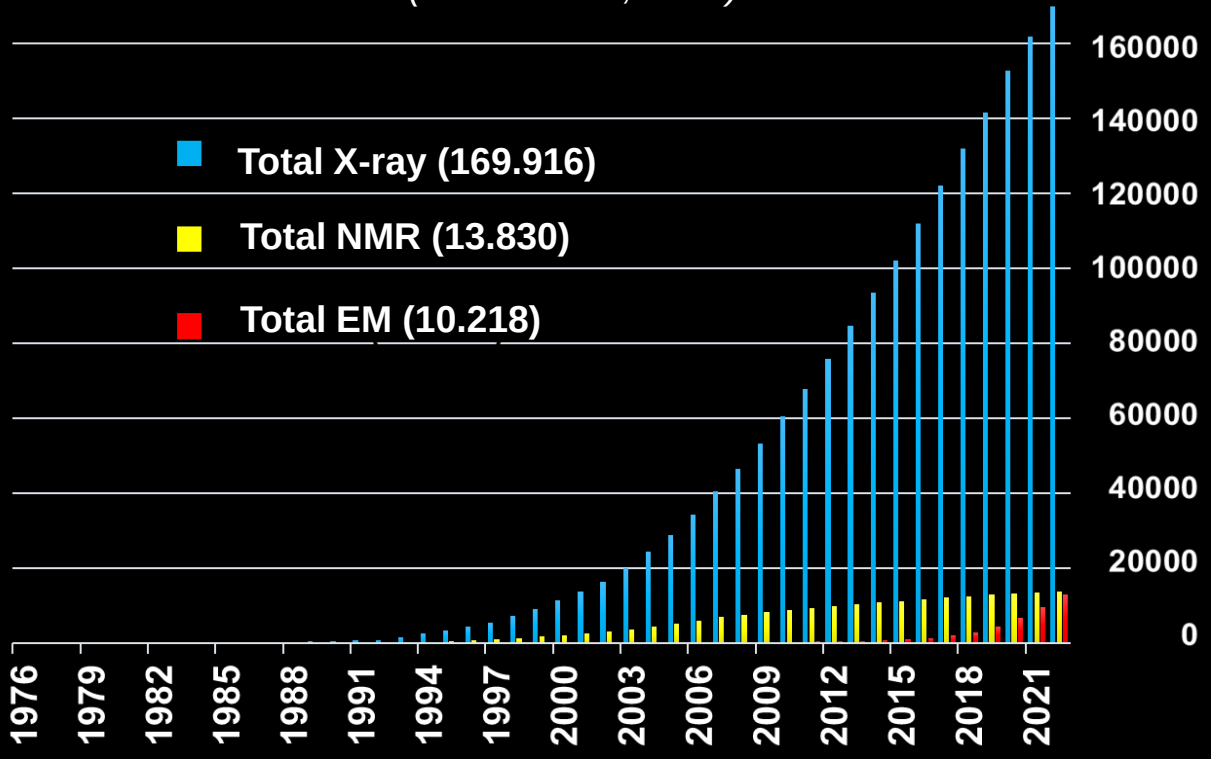
- ✓ No need to optimize crystallization conditions, which used to be time-consuming
- ✓ No cryo-protectant needed
- ✓ Determine room temperature free radiation damage crystal structures
- ✓ Study protein dynamics of reaction mechanisms => **molecular movies** of proteins at work

The vast majority of X-ray diffraction data at synchrotrons are measured from samples frozen at ~100K

Slightly warmer than the surface of Pluto (33-55 K)

Too cold to sustain life as we know it

Overall growth of structures deposited per year in the PDB (November 1, 2022)

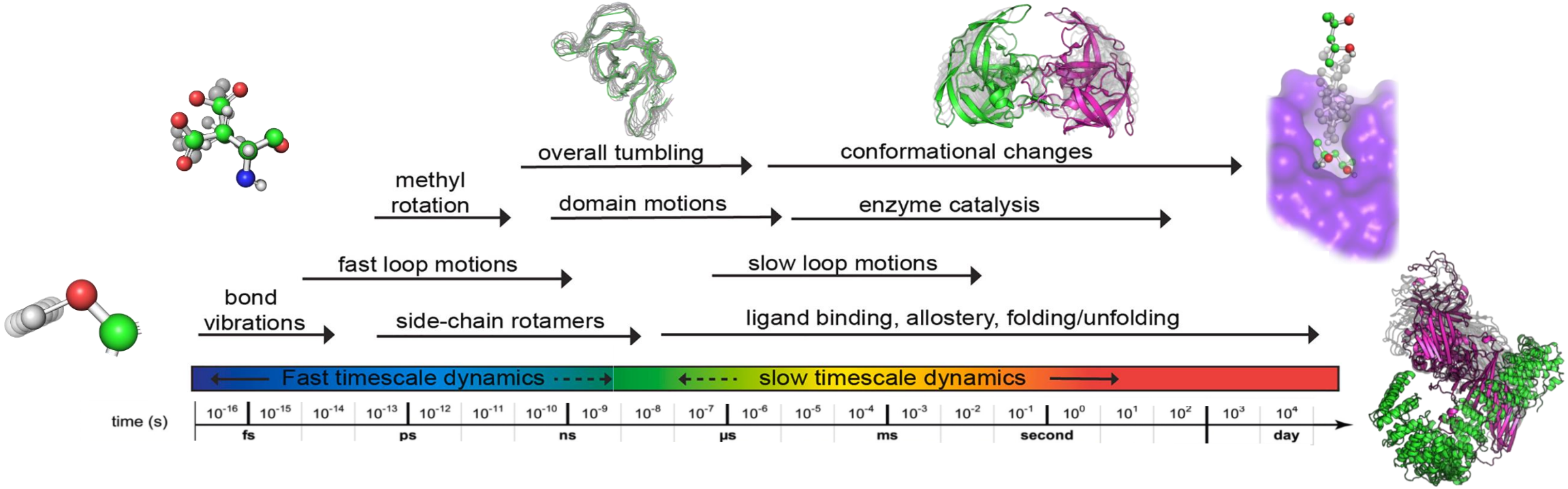


February 22, 2022
NASA/JHUAPL/SwRI

Life is dynamic, and function is not compatible with **cryogenic conditions**…

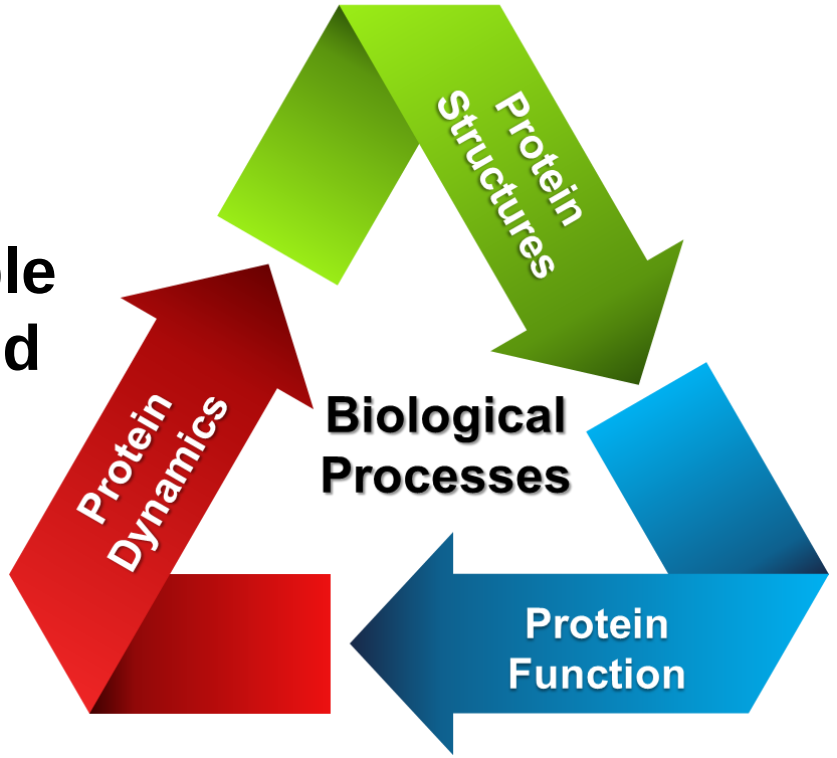
A **pending challenge** in structural biology is to determine time-resolved crystal structures at **physiological temperature and pressure**.

Protein dynamics and time scales in biology

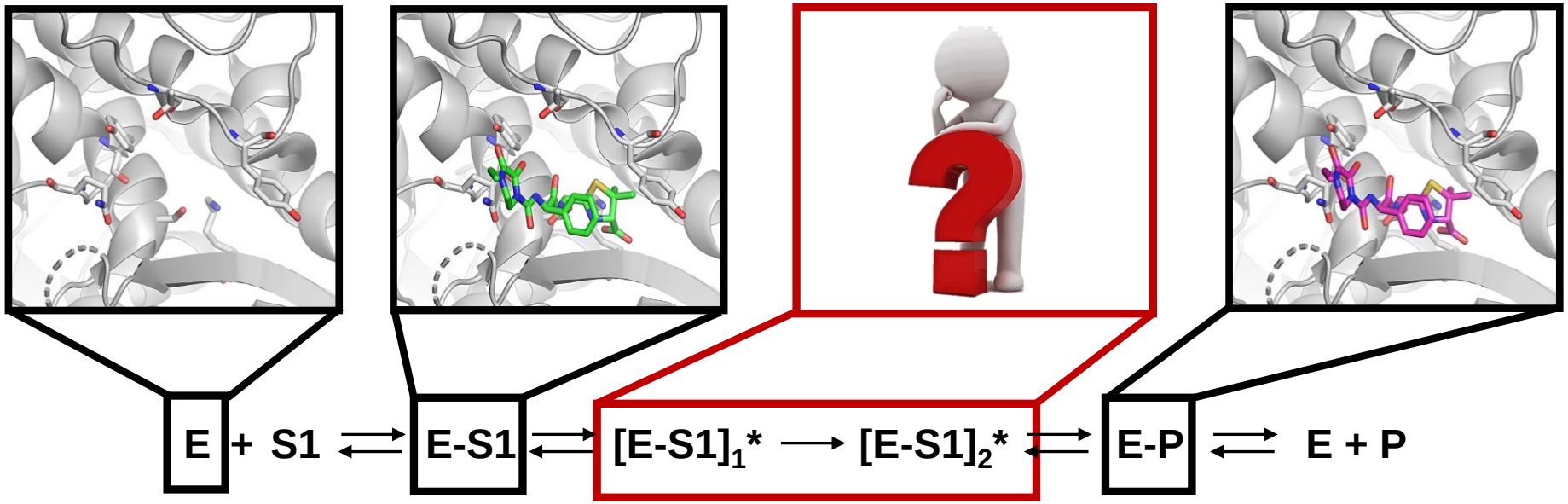


Time scales span from femtoseconds for bond breaking/formation, or bond vibrations all the way to milliseconds for most enzymatic reactions including allosteric regulation.

A complete understanding of the function of proteins requires of experiments able to probe both structural and dynamic properties!!!



Traditional MX @ synchrotrons limits dynamics studies



- Large-crystals
- Freeze-trapping technique misses dynamic motions and can perturb the protein energy landscape

Diffusion time limits the reaction initiation

The key event in a time-resolved experiment is the **diffusion** of the trigger of the reaction (ligand or light) throughout the crystals, which must be **much faster** and **homogenous** than the reaction of interest.

Hypothetical 3D protein crystal
(50% protein, 50% solvent)

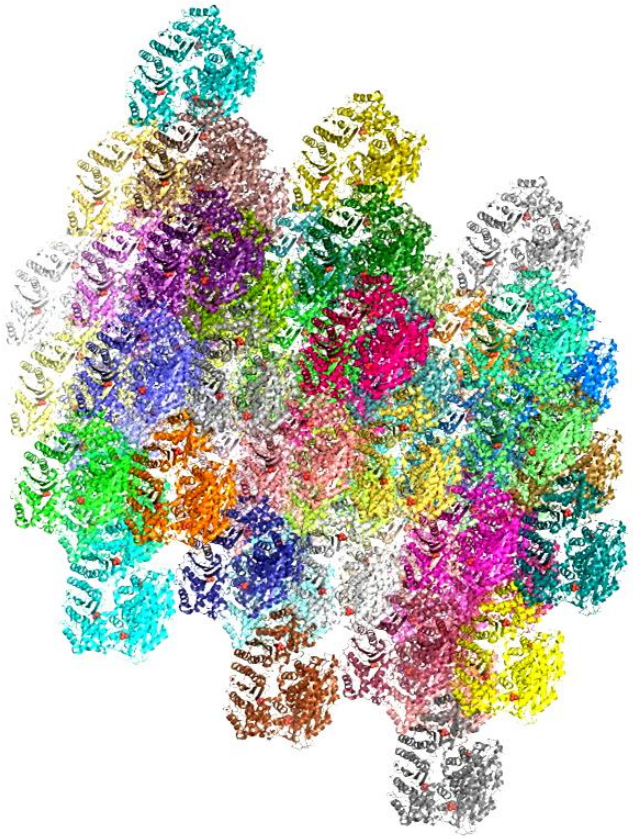


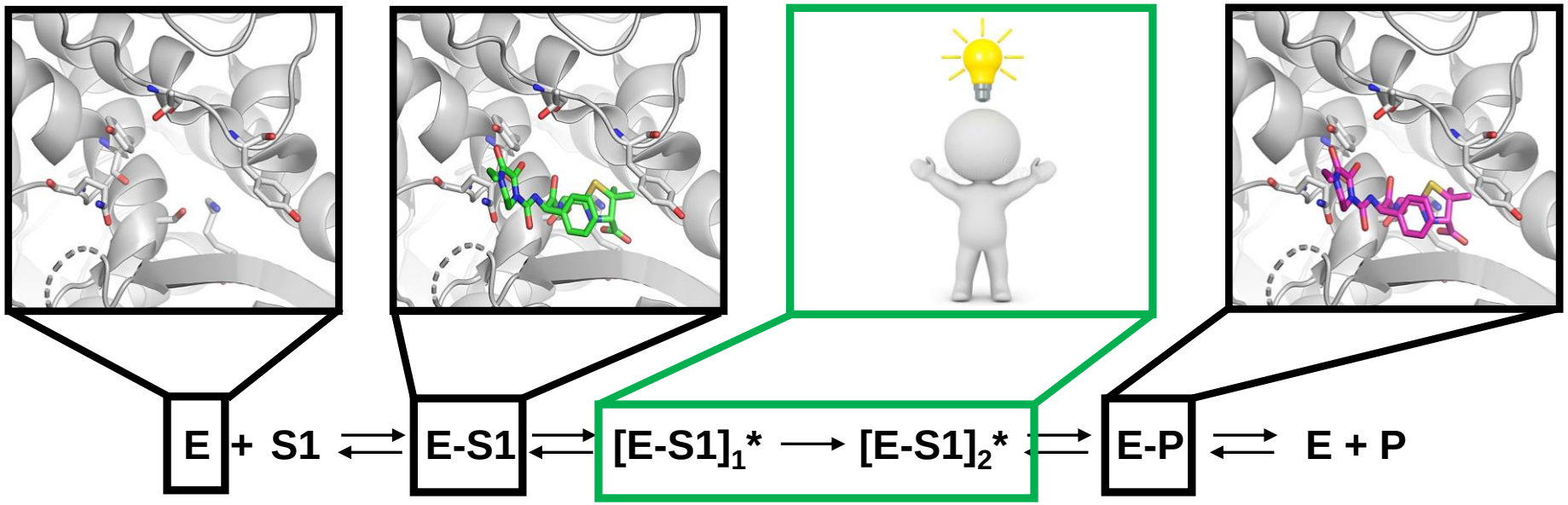
TABLE 1: Diffusion times τ_D for various crystal sizes from calculation, simulation, and experiment.

	Crystal size	τ_D
Large crystals	$400 \times 400 \times 1600\ \mu m^3$	16 s (3)
		24 s [60]
		<1 min [59]
	$300 \times 400 \times 500\ \mu m^3$	9.5 s
Small crystals	$10 \times 20 \times 30\ \mu m^3$	15 ms
	$3 \times 4 \times 5\ \mu m^3$	1 ms
	[a] $1 \times 2 \times 3\ \mu m^3$	150 μs
	$0.5 \times 0.5 \times 0.5\ \mu m^3$	17 μs
	[b] $0.1 \times 0.2 \times 0.3\ \mu m^3$	1.5 μs

[a] With much smaller crystals mixing times might be slower than diffusion times.
[b] *S. oneidensis* ccNIR crystals would have about 4000 unit cells.

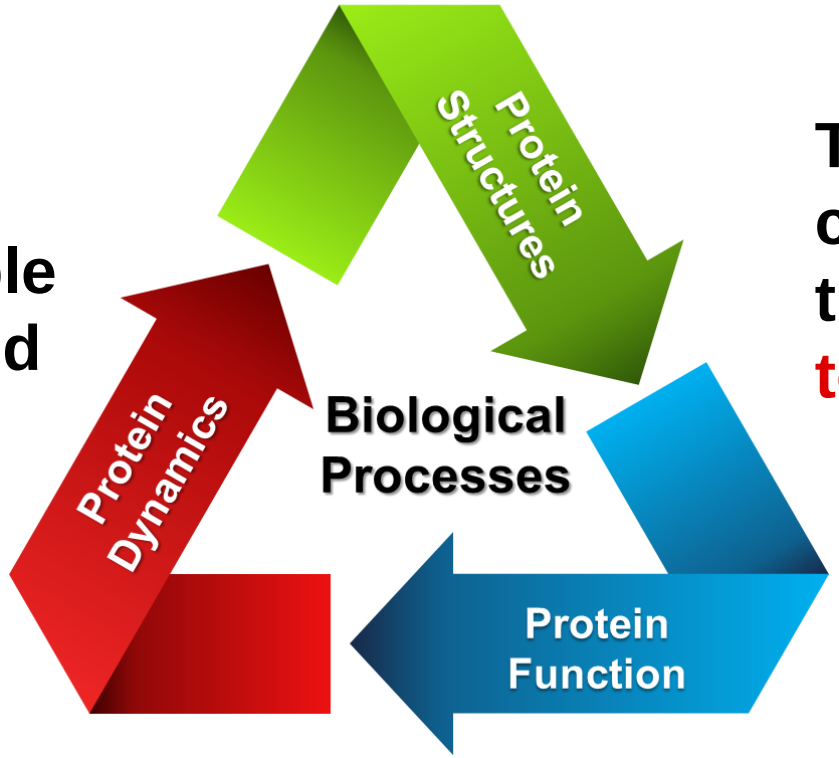
Schmidt, M. Advances in Condensed Matter Physics, 2013

Time resolved serial crystallography enables function and dynamics

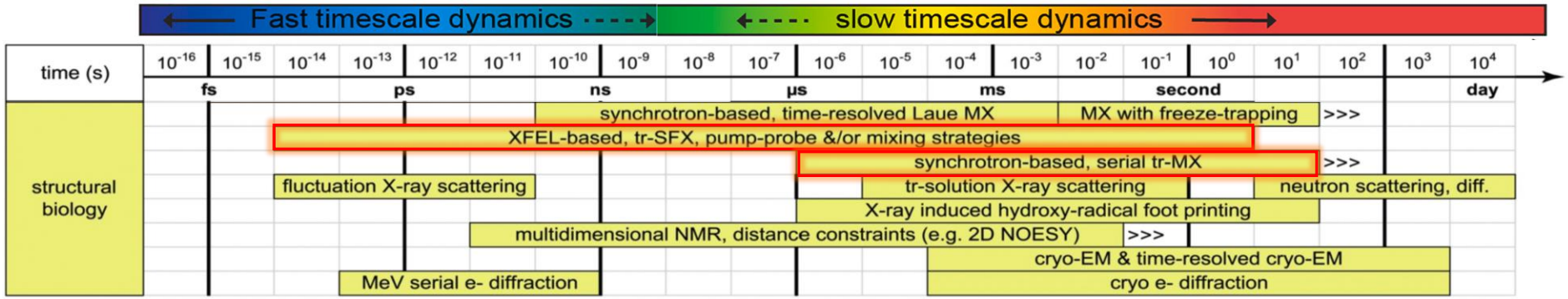


- The use of **micro- or submicron crystals** along with the recently **developed rapid mixing strategies** it is possible to diffuse small molecules through crystals a lot faster and uniformly.

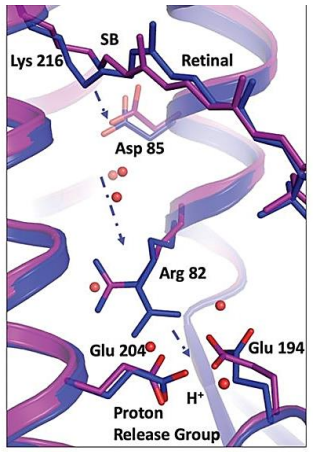
A complete understanding of the function of proteins requires of experiments able to probe both structural and dynamic properties!!!



Time resolved X-ray crystallography provides the greatest spatial and temporal resolution

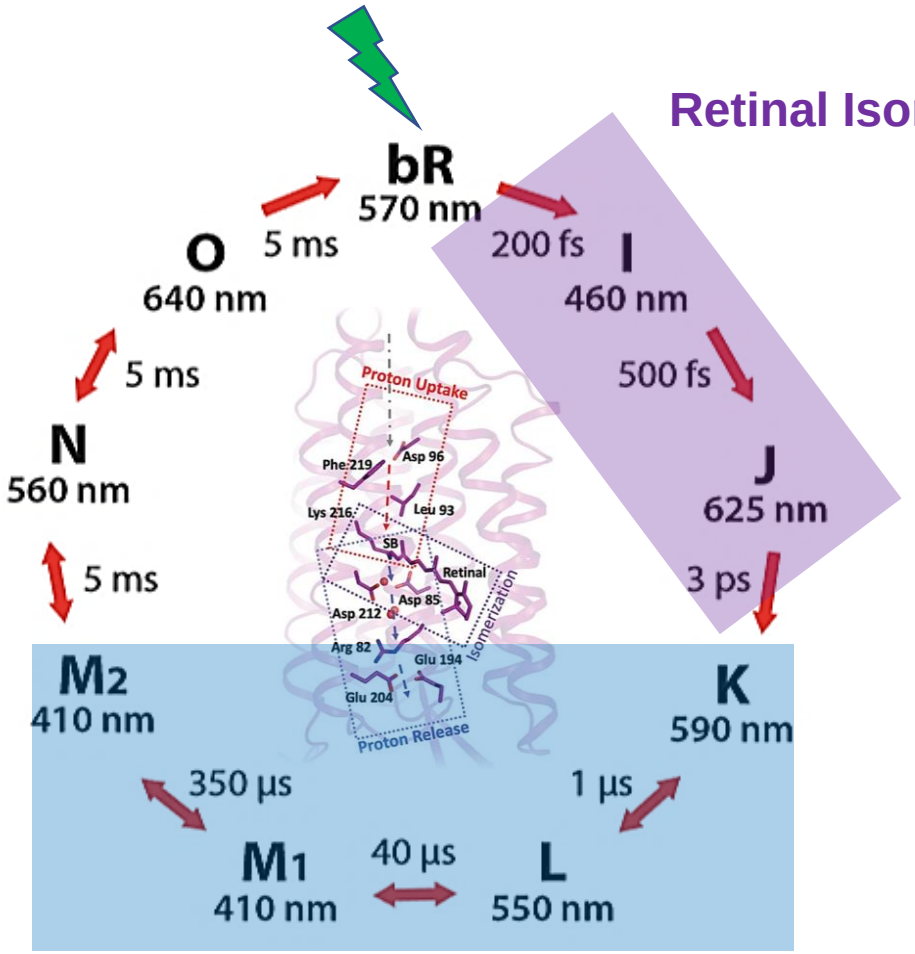


TR-SFX to decipher the mechanism of a biological pump



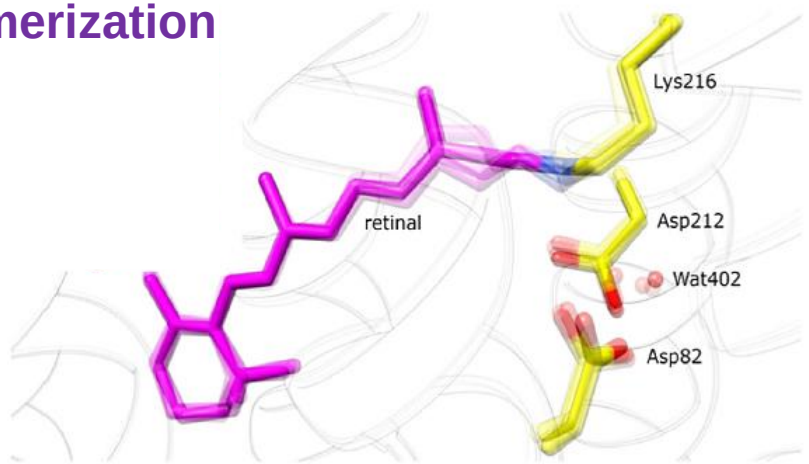
Nango et al. Science 2016
1st XFEL exp. (SACLA), Δt= ns-ms

Captured the primary proton transfer events



Proton Released

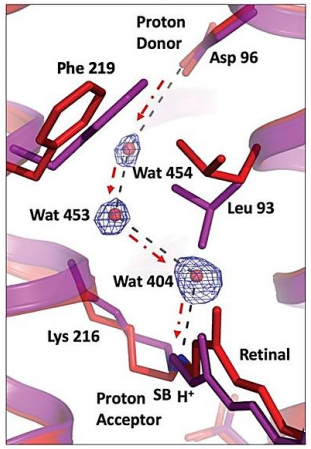
Retinal Isomerization

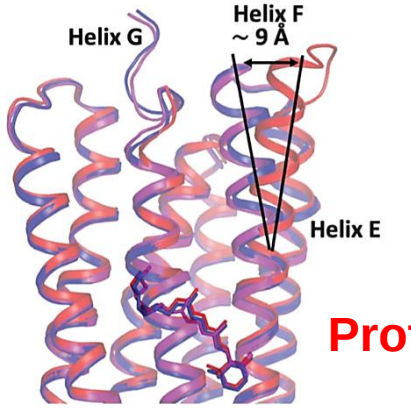


Nogly et al. Science 2018
2nd XFEL exp. (LCLS), Δt= fs-ps

Captured early events of retinal *trans/cis* isomerization

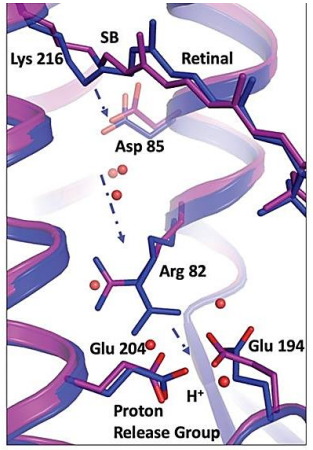
The synergy XFEL-Synchrotron to "Film" the first Molecular movie of bR





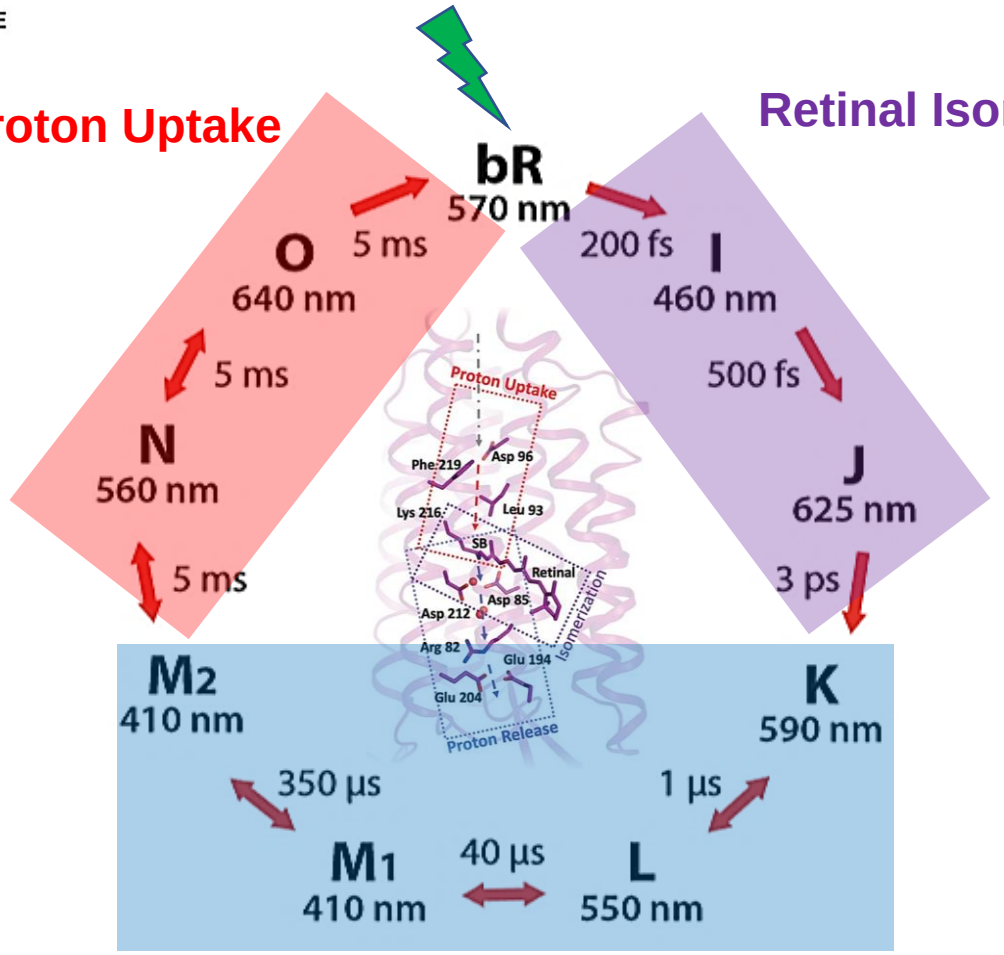
Weinert et al. Science. 2019
SMX exp. (SLS), Δt= ms

Captured the latest events



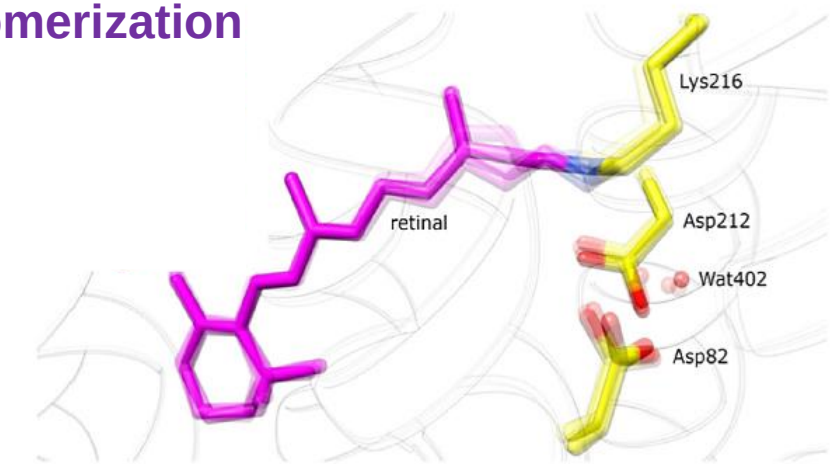
Nango et al. Science 2016
1st XFEL exp. (SACLA), Δt= ns-ms

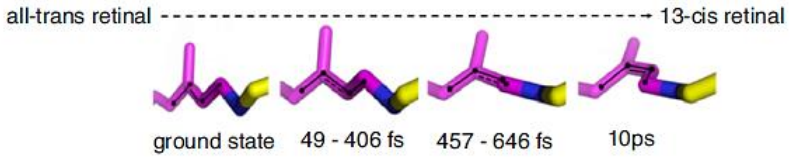
Captured the primary proton transfer events



Proton Released

Retinal Isomerization

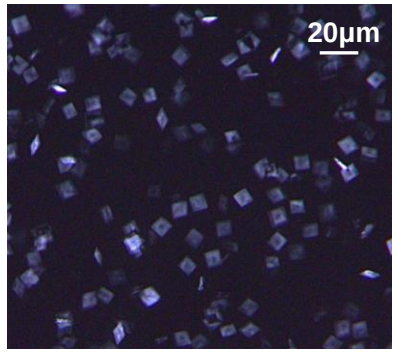




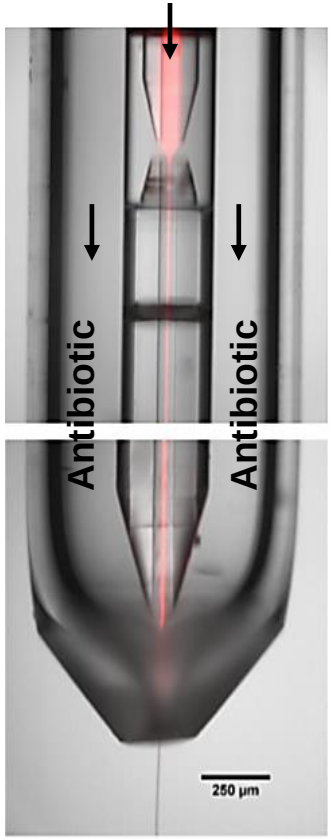
Nogly et al. Science 2018
2nd XFEL exp. (LCLS), Δt= fs-ps

Captured early events of retinal *trans/cis* isomerization

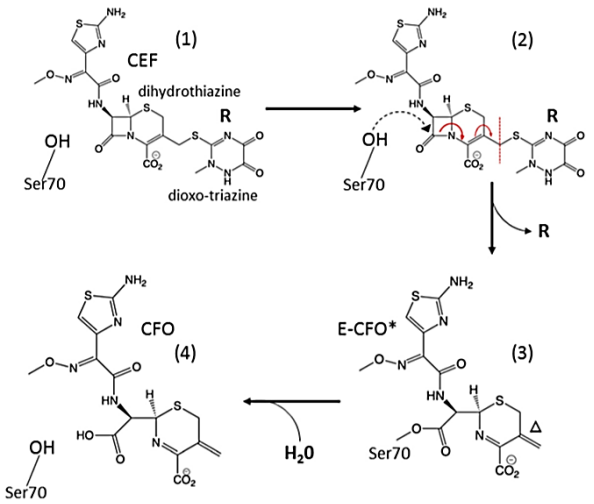
Watching the inactivation of an antibiotic in *real time*

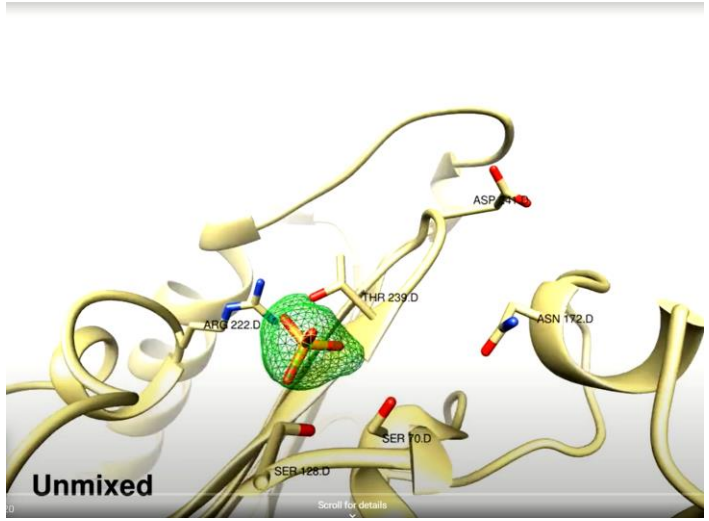


BlaC crystals



Catalytic reaction of an enzyme (β-lactamase C) with a β-lactam antibiotic (ceftriaxone) inside microcrystals using MISC TR-SFX

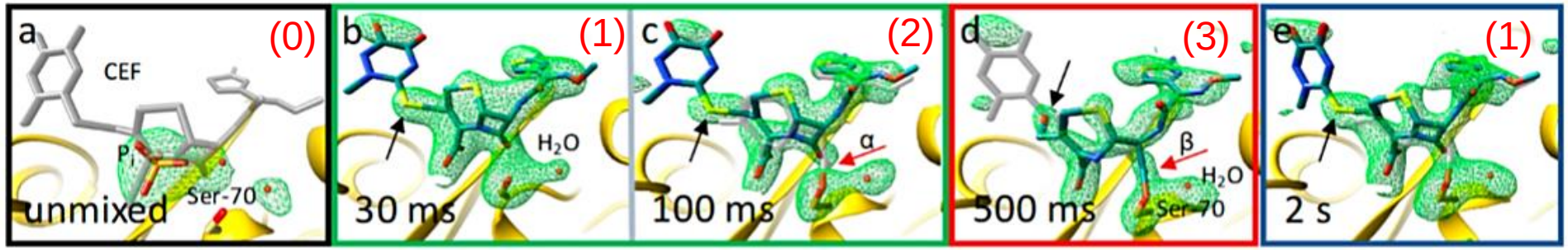




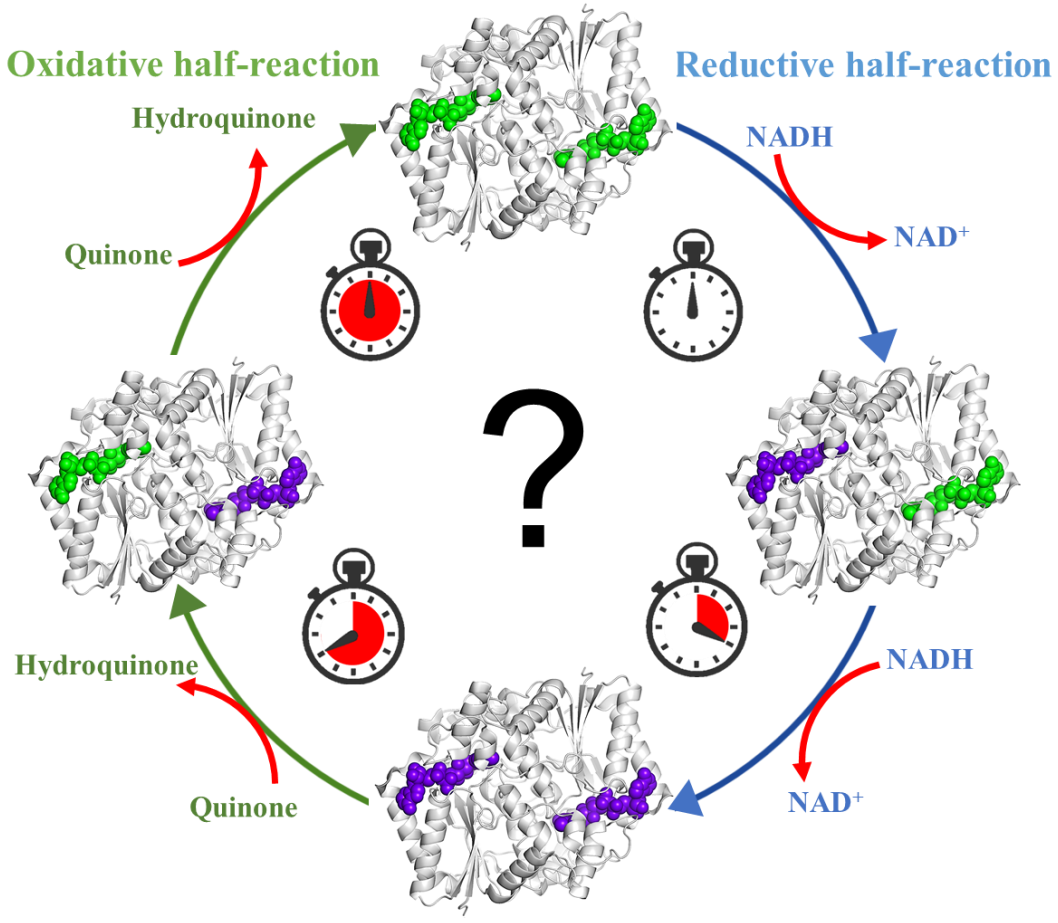
Kupitz C, et al. Structural Dynamics 2016

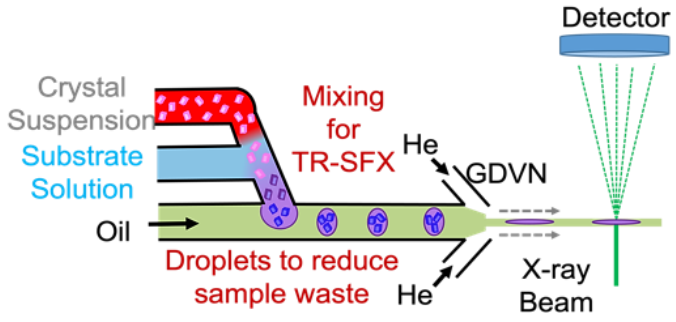
Olmos JL, et al. BMC Biology 2018

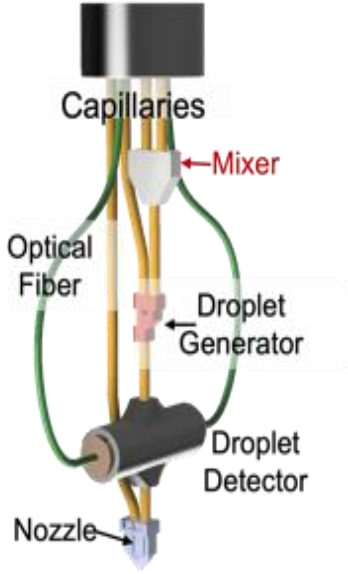
Suraj P, et al. *IUCrJ 2021*

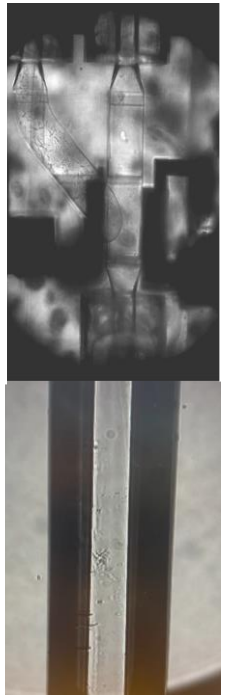


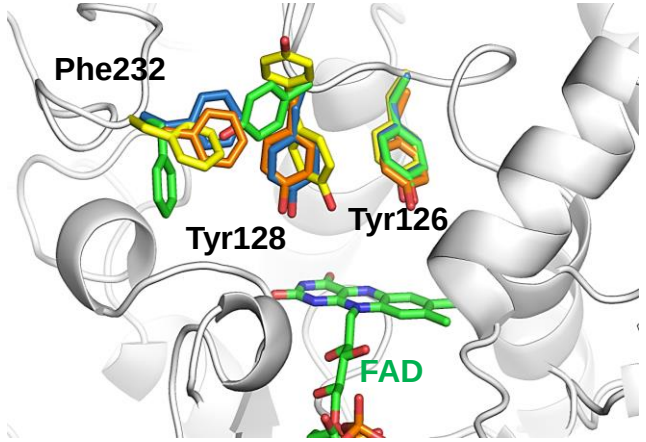
New insight into the redox mechanism of NQO1 by SFX

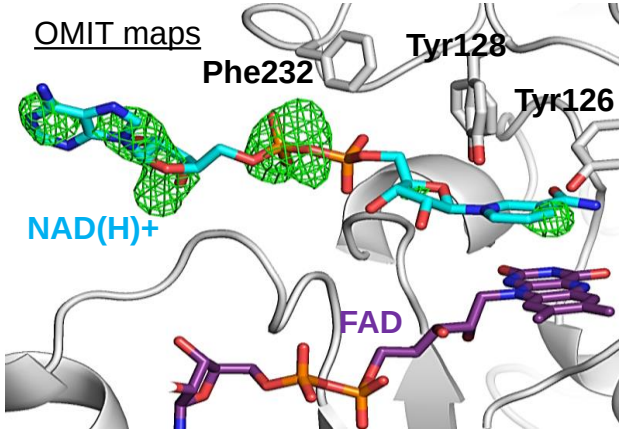


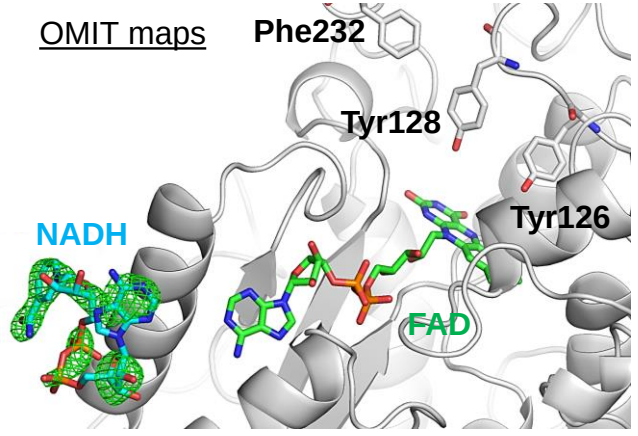












Doppler D, et al., Lab on a Chip. 2023 (*in review*)
Grieco A, et al. *Manuscript in preparation*

More beamtime at EuXFEL in Sept 2023

Acknowledgements

Juan Hermoso
Alice Grieco

Milagros Medina
Marta Martínez-Júlvez

Carlos Martí-Gastaldo

Petra Fromme
Alexandra Ros
Sabine Botha

Ana Cámara-Artigas

Katerina Doerner
Joachim Schulz
Richard Bean

Marius Schmidt

Roeland Boer

Rodolfo Miranda
Wojciech Gawelda

Robert Fischetti
Vukica Srajer

Daniele de Sanctis
Shibom Basu

Víctor de la Peña

Henry Chapman
Oleksandr Yefanov

Angel L. Pey
Irene Luque-Fernández

Felipe Gándara