

# Ada Lovelace Centre and Diamond Light Source

Paul Quinn

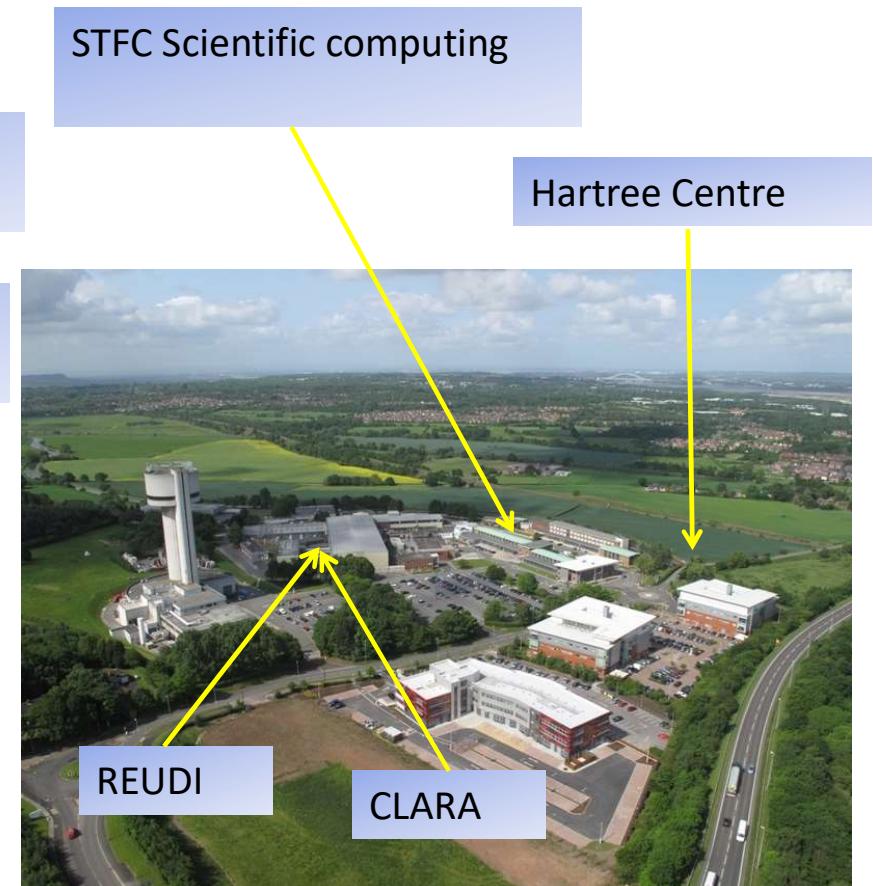
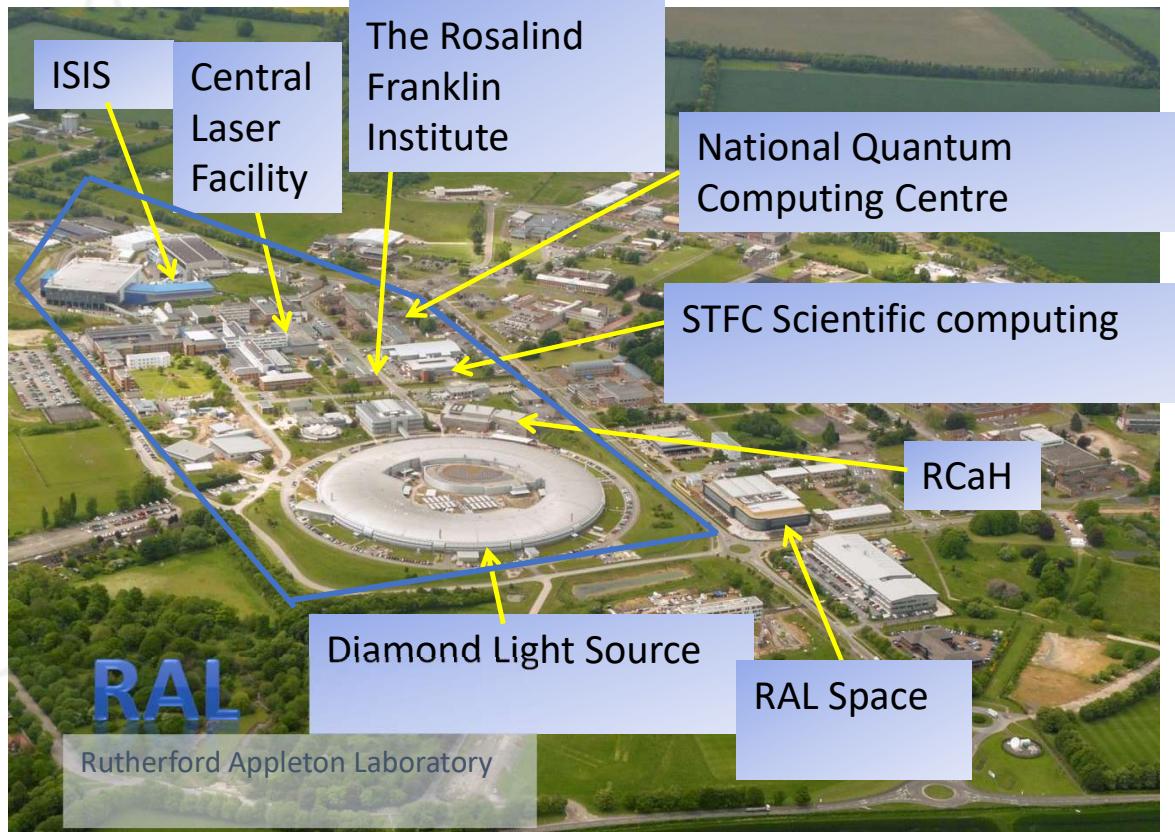
Director ALC

Julia Parker

Diamond Science Group Leader – Imaging and Microscopy



# Harwell and Daresbury Campus





Science and  
Technology  
Facilities Council

Scientific Computing

Paul Quinn

[Paul.quinn@stfc.ac.uk](mailto:Paul.quinn@stfc.ac.uk)

# Ada Lovelace Centre





## Ada Lovelace Centre

A centre of expertise in scientific computing



### Purpose

To maximise the scientific impact of the STFC large facilities – CLF, Diamond, ISIS and Scientific computing



### Vision

Transformative impact through challenge-driven collaboration

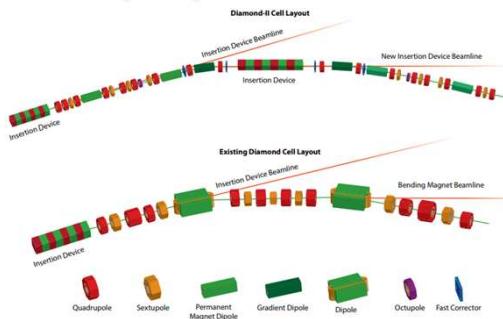


### Mission

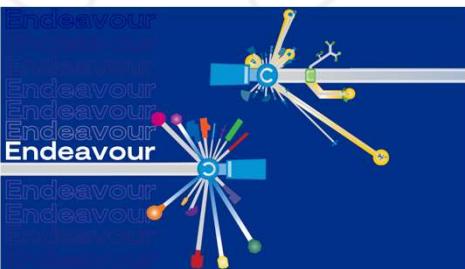
To create and deliver innovative scientific computing solutions to drive scientific and operational impact at large facilities.

# New Facilities and Upgrades

[UK XFEL](#)

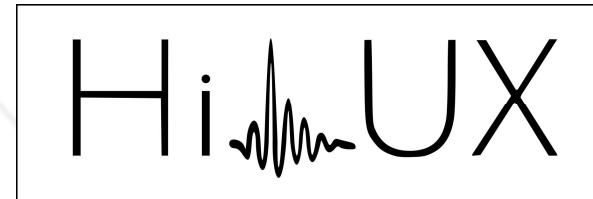


[Diamond-II - Diamond Light Source](#)



[ISIS The Endeavour Programme \(stfc.ac.uk\)](#)

[CLF HiLUX upgrades to Ultra and Artemis](#)



[CLF-EPAC](#)



PetaWatt laser and X-ray source

- x10-100 fold increases in power, flux, brightness
- Faster Timescales
- New instruments/techniques



Relativistic Ultrafast Electron Diffraction & Imaging



Themes

- Computational Engineering
- Computational Biology
- Computational Maths
- AI for Science
- Materials Modelling
- Data Engineering
- Cyber Security
- Open Science
- Software Engineering
- Platforms and Services
- Infrastructure

Programmes

STFC large facilities

ALC

Data infrastructure for  
the physical sciences

PSDI

Community software

CoSeC

Computing resources

IRIS



# From Responsive Projects to Programmes

- Seeking to develop challenge led programmes
  - Engaging with facilities
  - Bringing multiple disciplines to bear on a challenge
- Bridge across activities on site
  - Ptychography (ISIS, RFI, CLF,DLS), Tomography (ISIS,DLS,CLF)
  - Data management infrastructure
  - Materials modelling
- Set directions
  - We can't simply be responsive or wait for facilities to decide to do something

# ALC – transitioning to a programme



Deliver an AI and maths driven transformation



Enhance science and operations with simulation



Challenge led imaging



Cross-cutting computational biology

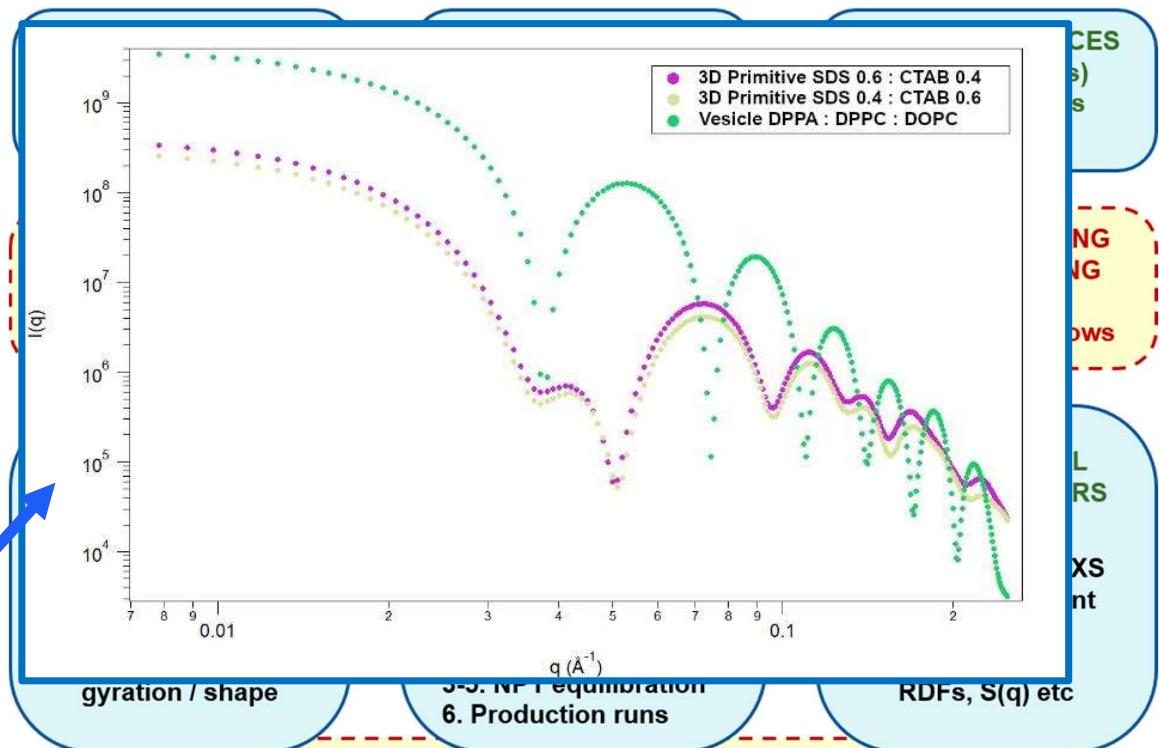
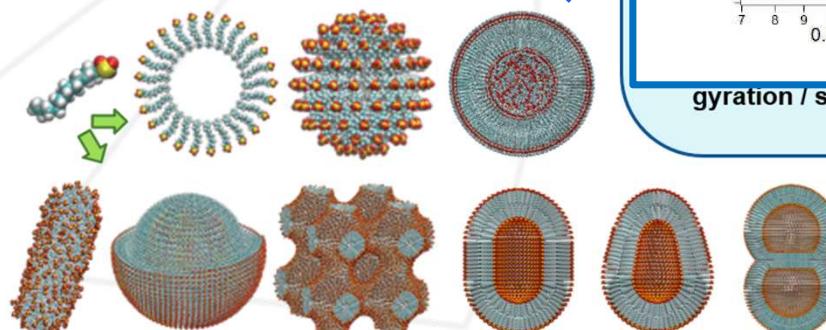


Platforms and services for scientific computing delivery

# Soft Matter simulations, analysis and link to SANS/SAXS - Shapespyer

## Project overview

- Generation, simulation & analyses of pre-equilibrated molecular structures for functional soft matter research
- Target structures / trajectories for academic and industrial R&D; Linking with SANS, SAXS experiments
- Provide automated workflows Python library, APIs, Bash scripts
- Partners / collaborators:  
*Maggie Holme, Chalmers, Sweden*  
*Hanna Barriga, Karolinska Institutet, Sweden*  
*Tom Headen, ISIS, UK*  
*SasView Team, International*



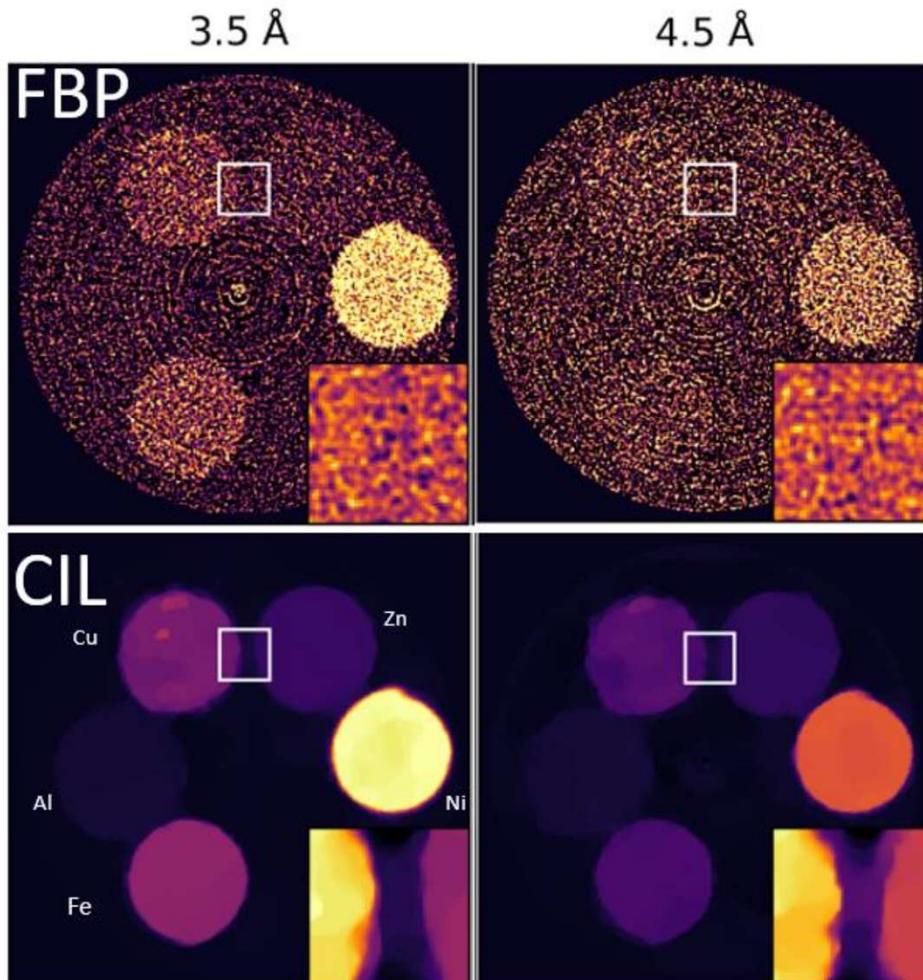
MC via DL\_MONTE & DPD via DL\_MESO  
[www.scd.stfc.ac.uk/Pages/Shapeshyer.aspx](http://www.scd.stfc.ac.uk/Pages/Shapeshyer.aspx)  
[gitlab.com/simnavi/shapeshyer](https://gitlab.com/simnavi/shapeshyer)

Andrey Brukhno (Co-I), Computational Chemistry, SCD

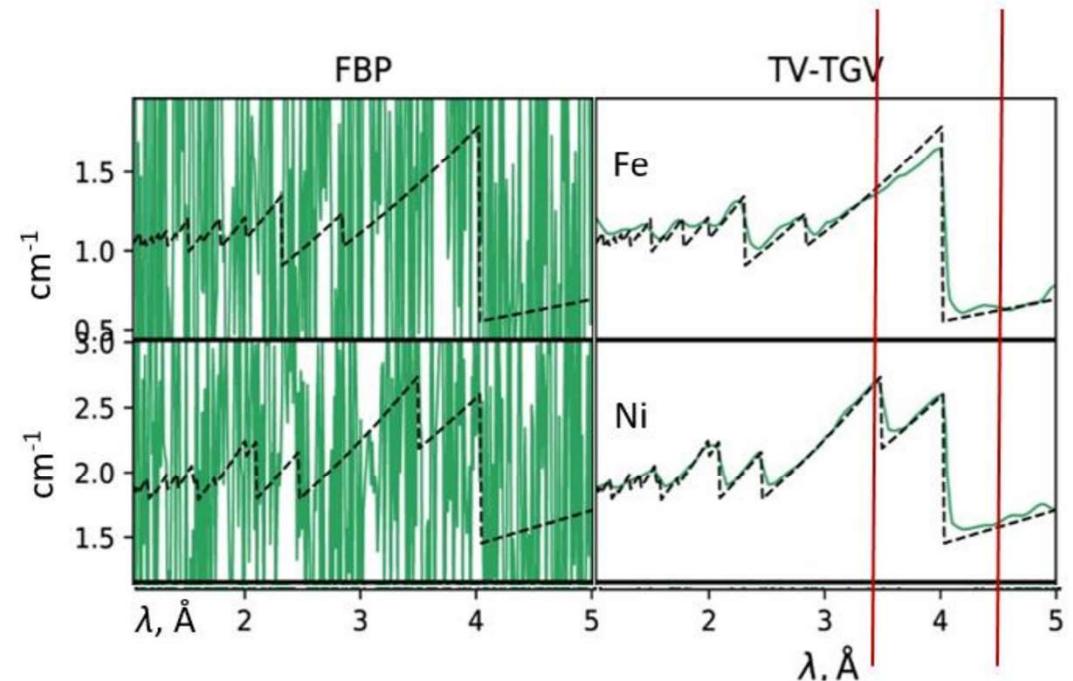
PI: James Doutch, ISIS, Co-I's: Tim Snow, Diamond, Michael Seaton, John Purton, SCD



# Energy-resolved neutron CT



- Data from ISIS/IMAT
- Proposed spatio-spectral TV-TGV regularization
- Enables clear identification of Bragg edges in 3D



Ametova et al. 2021: *Crystalline phase discriminating neutron tomography using advanced reconstruction methods*, J. Physics D, <https://doi.org/10.1088/1361-6463/ac02f9>

Key:

# FitBenchmarking



Table:

	WeightedNLLSCostFunc																					
	gsl					levmar			ralfit									scipy-ls				
	lmder	lmder: j:scipy 2-point	lmsder	lmsder: j:scipy 2-point	nmsimplex	levmar	levmar: j:scipy 2-point	gn	gn: j:scipy 2-point	gn_reg	gn_reg: j:scipy 2-point	hybrid	hybrid: j:scipy 2-point	hybrid_reg	hybrid_reg: j:scipy 2-point	dogbox	dogbox: j:scipy 2-point	lm-scipy: j:scipy 2-point	lm	trf	trf: j:scipy 2-point	
ENGINX 193749 calibration, spectrum 651, peak 19	inf (inf) inf (inf) <sup>3</sup>	inf (inf) inf (inf) <sup>3</sup>	54.4 (1) 0.6348 (55.1)	54.4 (1) 0.6349 (55.11)	56.45 (1.038) 0.9912 (86.03) <sup>1</sup>	99.41 (1.827) 0.01152 (1) (2.048)	99.41 (1.827) 0.02359 (490.5)	54.49 (1.002) 54.49 (1.002) 5.651 (490.5)	54.52 (1.002) 54.83 (475.9) 5.606 (486.5)	54.52 (1.002) 5.735 (497.7) 1.533 (133)	54.4 (1) 1.635 (141.9) 1.56 (135.4)	54.4 (1) 1.642 (142.5) (384.2) <sup>1</sup>	52.6 (1.151) 4.427 (472.9)	62.62 (1.151) 4.909 (426) <sup>1</sup>	54.4 (1) 0.4637 (40.24) (34.75)	54.4 (1) 0.4004 (26.25) (28.74)	54.4 (1) 0.3025 (0.3311)	54.4 (1) 0.4004 (26.25) (28.74)	54.4 (1) 0.3025 (0.3311)			
ENGINX 193749 calibration, spectrum 651, peak 20	56.38 (1) 0.4192 (44.11)	56.38 (1) 0.4101 (43.15)	56.38 (1) 0.1857 (19.54)	56.38 (1) 0.1887 (19.85)	64.84 (1.15) 0.8394 (88.33) <sup>1</sup>	224.5 (3.982) 0.009503 (1) (2.063)	56.38 (1) 0.01961 (18.81)	56.38 (1) 0.1787 (18.85) (31.67)	56.38 (1) 0.301 (31.67)	56.38 (1) 0.3002 (31.59) (38.99)	56.38 (1) 0.3705 (38.98)	56.38 (1) 0.3681 (38.74)	120.5 (2.137) 4.495 (472.9) <sup>1</sup>	104.9 (1.861) 4.801 (505.2) <sup>1</sup>	56.38 (1) 0.4264 (44.87) (38.05)	56.38 (1) 0.3616 (51.06) (63.09)	56.38 (1) 0.4852 (0.5995)	56.38 (1) 0.4852 (0.5995)	56.38 (1) 0.4852 (0.5995)			
ENGINX 193749 calibration, spectrum 651, peak 23	76.09 (1) 0.7439 (79)	76.09 (1) 0.7365 (78.22)	76.09 (1) 0.4968 (52.77)	76.09 (1) 0.5093 (54.1)	115.1 (1.512) 0.8105 (86.09) <sup>1</sup>	718 (9.436) 0.009416 (1) (2.062)	718 (9.436) 0.01942 (610.6)	76.11 (1) 5.749 (610.6)	76.11 (1) 5.444 (578.1) (578.2)	76.14 (1.001) 5.519 (586.2) (92.86)	76.09 (1) 0.8743 (92.86)	76.09 (1) 0.8799 (93.46) (92.99)	76.09 (1) 0.8755 (92.66)	76.09 (1) 0.8725 (92.66) (504.8) <sup>1</sup>	77.46 4.753 5.074 (538.9) <sup>1</sup>	76.7 (1.008) 0.7502 (79.68) (52.27)	76.09 (1) 0.4921 (53.62) (56.13)	76.09 (1) 0.5049 (0.5285)	76.09 (1) 0.5049 (0.5285)	76.09 (1) 0.5049 (0.5285)		
ENGINX 193749 calibration, spectrum 651, peak 5	14.58 (1.002) 0.2981 (31.86)	14.58 (1.002) 0.2973 (31.78)	14.58 (1.002) 0.7617 (81.41)	14.58 (1.002) 0.8031 (85.84)	16.16 (1.111) 0.8115 (86.73) <sup>1</sup>	24.82 (1.706) 0.009356 (1) (2.079)	14.55 (1) 5.066 (541.4)	14.55 (1) 4.981 (532.3)	14.68 (1.009) 5.435 (580.9)	14.68 (1.009) 5.384 (575.5) (84.2)	14.91 (1.025) 0.7878 (82.36)	14.91 (1.025) 0.7705 (82.36) (82.34)	14.91 (1.025) 0.7704 (82.34)	22.86 4.645 5.21 (556.9) <sup>1</sup>	22.86 (1.571) 0.7969 (85.17) (496.5) <sup>1</sup>	14.58 (1.002) 0.2998 (32.04) (28.21)	14.58 (1.002) 0.2639 (24.48) (30.96)	14.58 (1.002) 0.2291 (24.48) (30.96)	14.58 (1.002) 0.2291 (30.96)	14.58 (1.002) 0.2291 (30.96)		
ENGINX 193749 calibration, spectrum 651, peak 6	26.76 (1141) 0.2948 (25.44)	26.76 (1141) 0.254 (21.92)	inf (inf) inf (inf) <sup>3</sup>	inf (inf) inf (inf) <sup>3</sup>	25.79 (1101) 0.7929 (68.43) <sup>1</sup>	40.12 (1.712) 0.01933 (1.669)	23.46 (1.001) 5.159 (443.6)	23.46 (1.001) 5.197 (448.6)	23.47 (1.002) 5.179 (447)	23.43 (1) 5.41 (466.9)	23.43 (1) 5.334 (460.4)	23.43 (1) 5.413 (467.1)	23.43 (1) 5.347 (461.5) (305.9) <sup>1</sup>	29 (1.237) 3.545 (305.9) <sup>1</sup>	28.25 (1.206) 4.163 (359.3) <sup>1</sup>	26.74 (1.141) 0.3064 (26.44) (40.96)	26.72 (1.141) 0.4746 (40.96) (30.4)	26.68 (1.139) 0.3492 (30.4) (30.86)	26.68 (1.139) 0.3492 (30.86)	26.68 (1.139) 0.3492 (30.86)		
ENGINX 236516 vanadium, bank 1, 10 brk	8.455e+04 (1) 0.9798 (3.102)	8.455e+04 (1) 0.9797 (3.101)	8.455e+04 (1) 1.312 (4.154)	8.455e+04 (1) 1.351 (4.277)	6.362e+06 (75.24) 13.75 (43.53) <sup>1</sup>	8.455e+04 (1) 0.3159 (1)	8.455e+04 (1) 1.801 (5.7)	8.455e+04 (1) 1.367 (4.328)	8.455e+04 (1) 1.04 (3.293)	8.455e+04 (1) 1.054 (3.337)	8.455e+04 (1) 1.535 (4.858)	8.455e+04 (1) 1.55 (4.907)	8.455e+04 (1) 1.503 (4.759)	8.455e+04 (1) 1.501 (4.751) (3.101 (9.817)	8.455e+04 (1) 3.136 (9.929)	8.455e+04 (1) 1.324 (4.19)	8.455e+04 (1) 1.291 (4.086)	8.455e+04 (1) 3.911 (12.38)	8.455e+04 (1) 3.79 (12)	8.455e+04 (1) 3.79 (12)		
	2.478e+04		2.478e+04		2.478e+04		2.478e+04		2.478e+04		2.478e+04		2.478e+04		2.478e+04		2.478e+04		2.478e+04			

Clicking a result in the tables will give more details, such as graphs of the fit against the data and the parameters that the minimizer found.

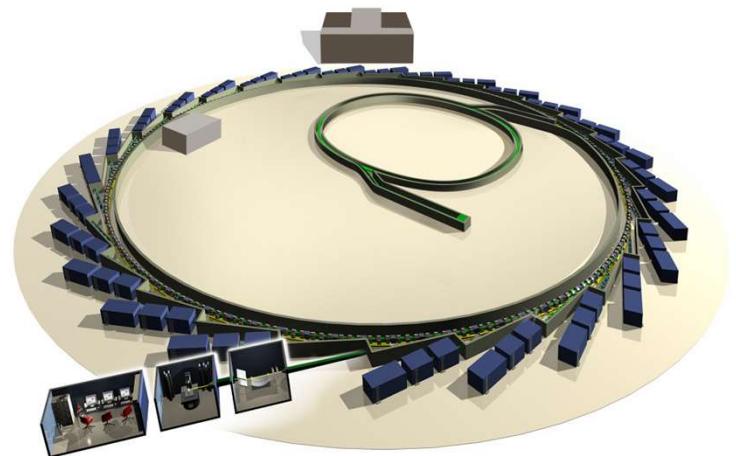
Clicking the problem names will take you to details of the best minimizer.

Clicking the software name will take you to FitBenchmarking Read the Docs documentation for the selected software.

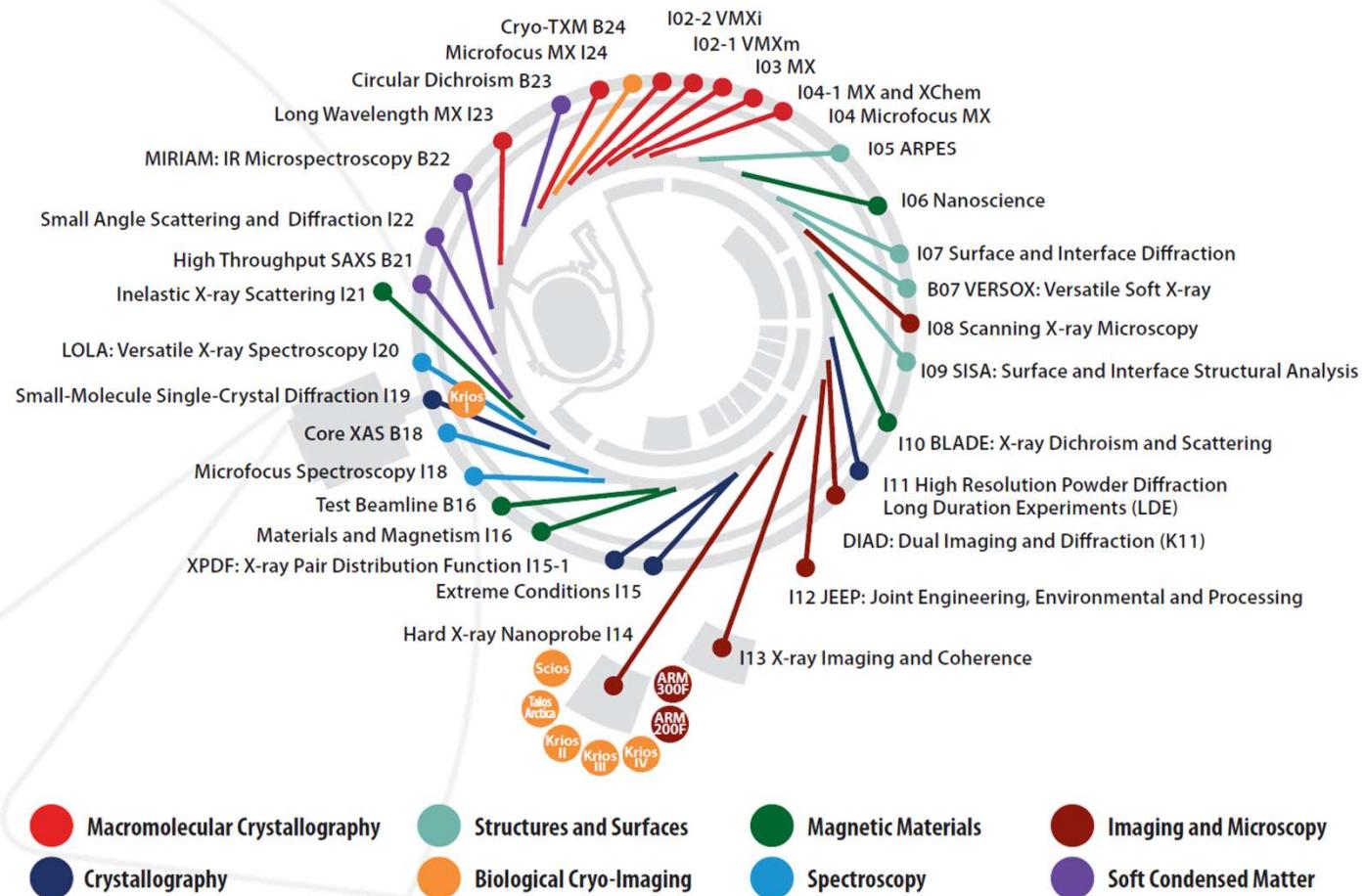
Hovering over each minimizer name will display its matching algorithm types out of those selected in options.

# Diamond Light Source

- Is the UK's national synchrotron facility
- 33 operational beamlines for imaging, spectroscopy, diffraction and more
- Also home to electron microscopy facilities for biology and physical sciences (eBIC and ePSIC)



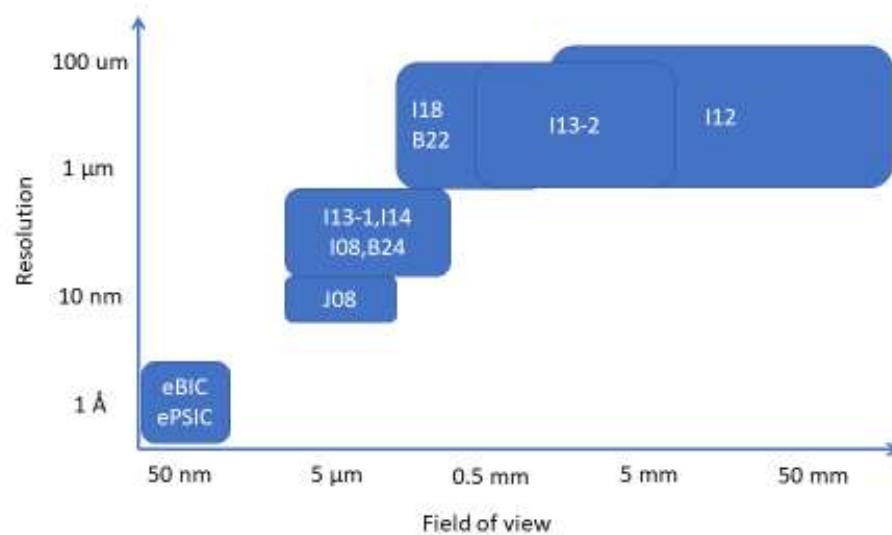
# Diamond Beamlines and Science Groups



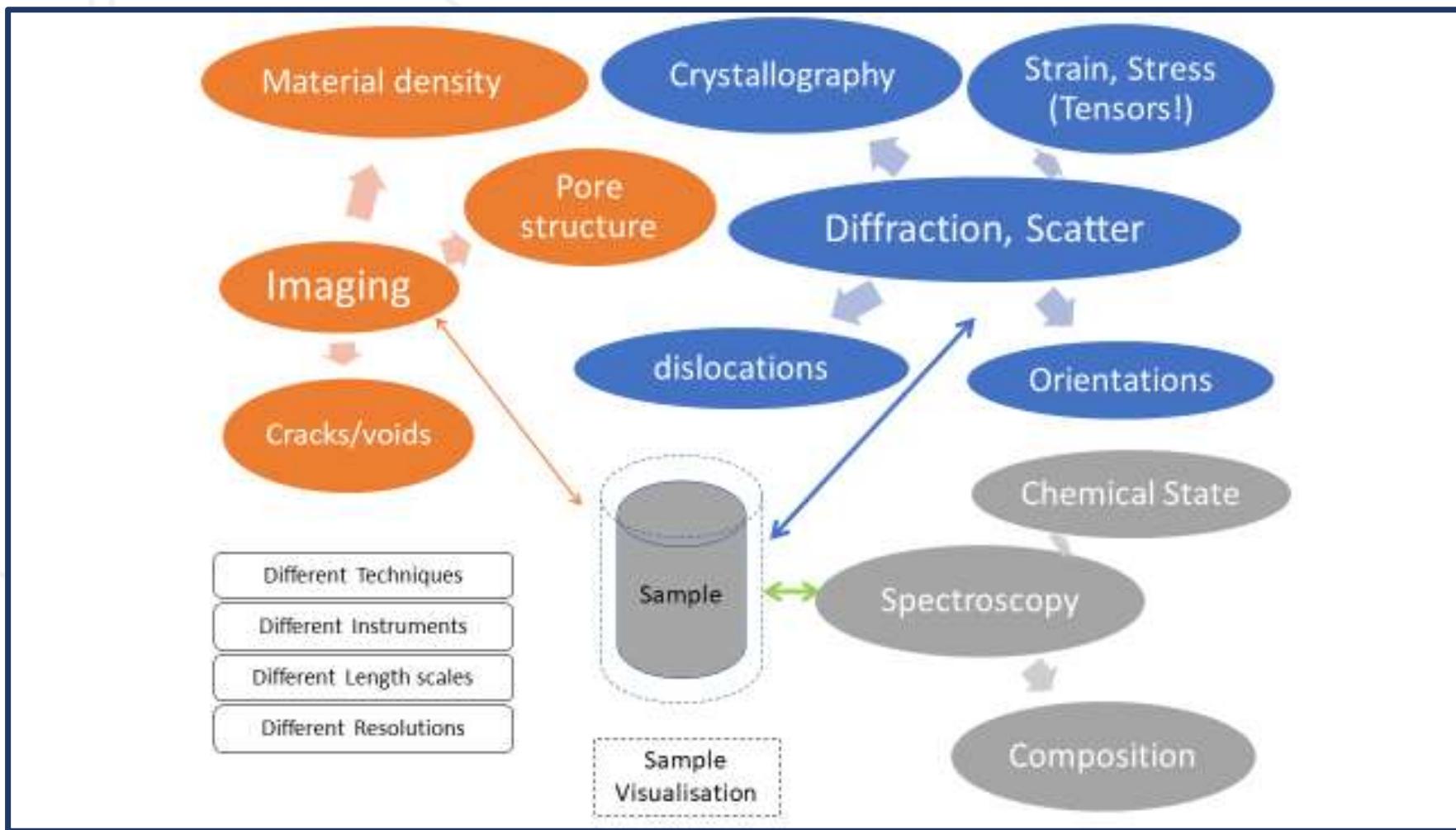
# Imaging and Microscopy at Diamond

- Scanning Transmission X-ray Microscope & Soft X-ray spectro- and tomo-ptychography (**I08/J08**)
- X-ray Imaging and Coherence (**I13-1, I13-2**, in collaboration with UoMan)
- Hard X-ray Nanoprobe (**I14**)
- **I12 –JEEP** high energy imaging/diffraction/scattering & **DIAD** (Dual Imaging & Diffraction)
- **ePSIC** (Electron Microscopy in collaboration with UoX/Johnson Matthey)
- **<https://www.diamond.ac.uk/Instruments/Imaging-and-Microscopy.html>**

Beamlines cover a range of resolutions and fields of view



# Imaging and Microscopy





Some X-ray Imaging Examples...



# Across length scales, resolutions and techniques

## Fall out particles from Fukushima Incident



Major nuclear accident at the Fukushima Daiichi nuclear power plant in Japan which began in 2011 following an earthquake and tsunami  
Loss of power and cooling – failure of containment and release of radioactive contaminants  
Aim to understand spread and provenance of contaminants – for modelling sequence of failure events and understanding clean up



- Combine techniques for element specific information
- I<sub>13</sub>-2: μ-X-ray Tomography, μ-XRF Tomography

ARTICLE

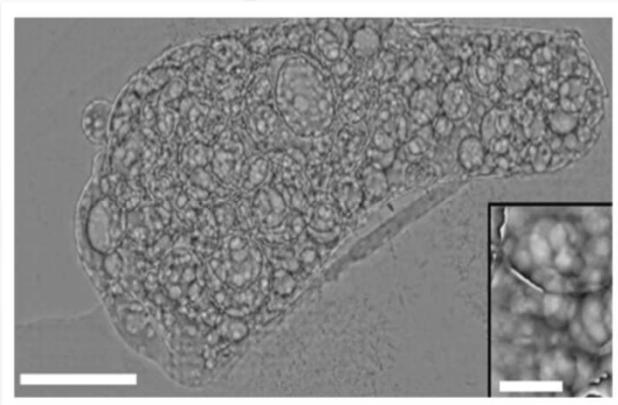
<https://doi.org/10.1038/s41467-019-10937-z>

OPEN

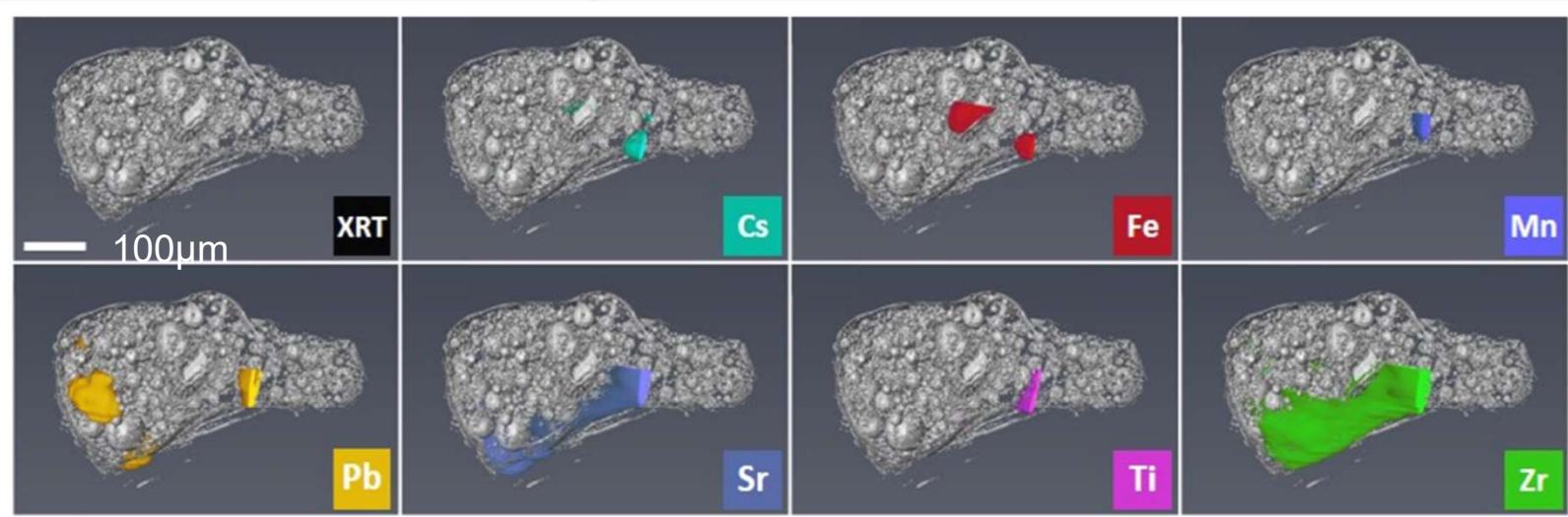
Provenance of uranium particulate contained within Fukushima Daiichi Nuclear Power Plant Unit 1 ejecta material

Peter G. Martin<sup>1</sup>, Marion Louvel<sup>2</sup>, Silvia Cipiccia<sup>3</sup>, Christopher P. Jones<sup>1</sup>, Darren J. Batey<sup>3</sup>, Keith R. Hallam<sup>1</sup>, Ian A.X. Yang<sup>1</sup>, Yukihiko Satou<sup>4</sup>, Christoph Rau<sup>3</sup>, J. Fred W. Mosselmans<sup>1</sup>, David A. Richards<sup>1</sup> & Thomas B. Scott<sup>1</sup>

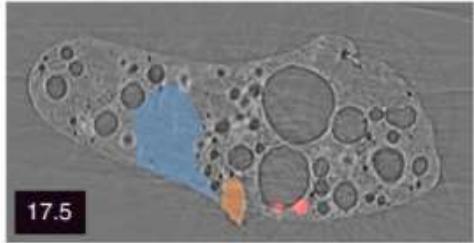




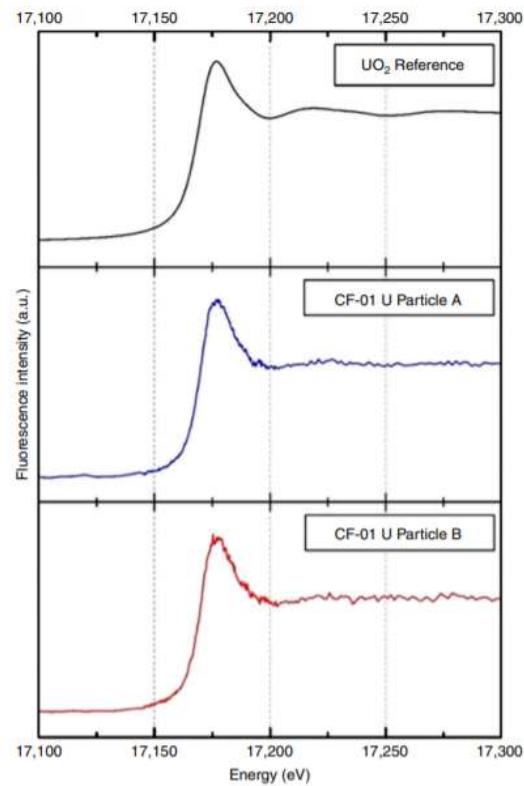
Morphology (Tomography)  
Glassy bubble structure: Temperature



P. G. Martin et al., *Nature Communications* (2019), DOI:[10.1038/s41467-019-10937-z](https://doi.org/10.1038/s41467-019-10937-z)



- I18  $\mu$ -XANES
- Uranium XANES on two U composition particles within the Si based particle
- Exist as  $\text{UO}_2$  – identical to reactor fuel



P. G. Martin et al., *Nature Communications* (2019), DOI:[10.1038/s41467-019-10937-z](https://doi.org/10.1038/s41467-019-10937-z)

## Single-Cell Chemistry of Photoactivatable Platinum Anticancer Complexes

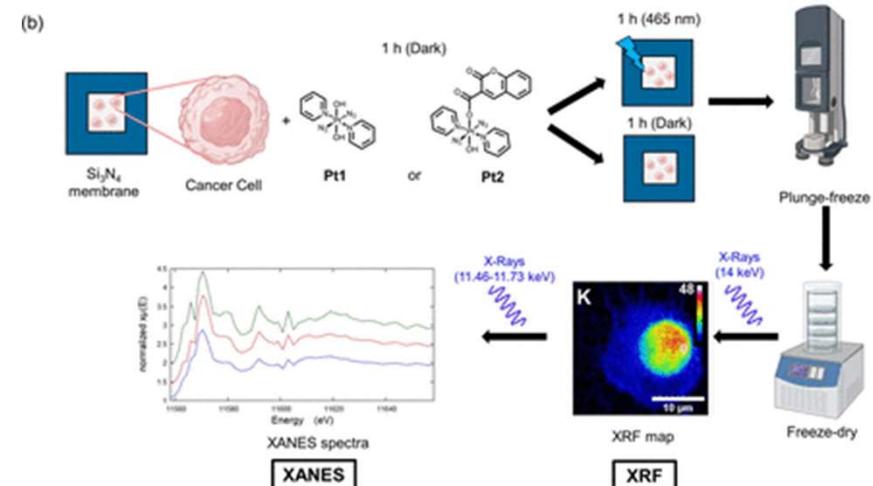
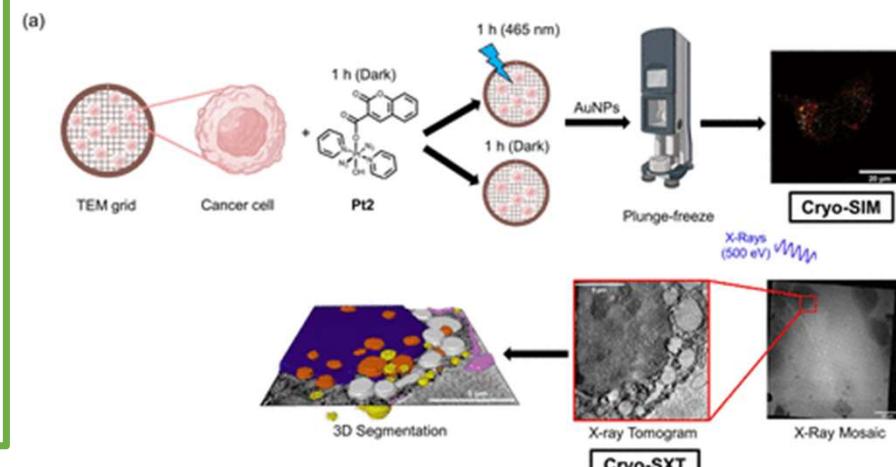
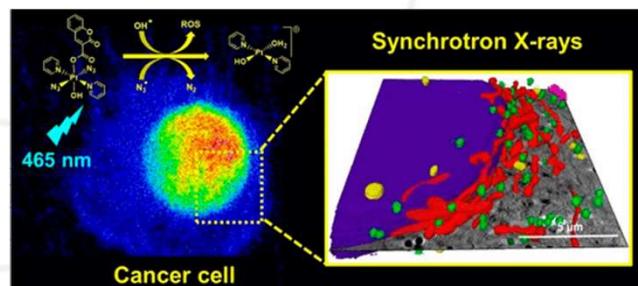
Elizabeth M. Bolitho, Carlos Sanchez-Cano, Huayun Shi, Paul D. Quinn,\* Maria Harkiolaki,\* Cinzia Imberti,\* and Peter J. Sadler\*

Cite This: *J. Am. Chem. Soc.* 2021, 143, 20244–20240

Read Online

Combined cryo-SXT(B24) with n-XRF, and n-XANES for the first time to investigate the anticancer mechanism of action of photoactivatable platinum anticancer complexes in cancer cells in their near-native state. Observed higher levels of platinum in cells that had been treated and irradiated.

XANES results allowed following the photoreduction process of Pt(IV) to Pt(II) during irradiation.





## Nanoscale correlative X-ray spectroscopy and ptychography of carious dental enamel

Cyril Besnard <sup>a,\*</sup>, Ali Marie <sup>a</sup>, Sisini Sasidharan <sup>a</sup>, Petr Buček <sup>b</sup>, Jessica M. Walker <sup>c</sup>, Julia E. Parker <sup>c</sup>, Thomas E.J. Moxham <sup>a,c</sup>, Benedikt Daurer <sup>c</sup>, Burkhard Kaulich <sup>c</sup>, Majid Kazemian <sup>c</sup>, Richard M. Shelton <sup>d</sup>, Gabriel Landini <sup>d</sup>, Alexander M. Korsunsky <sup>a,\*</sup>

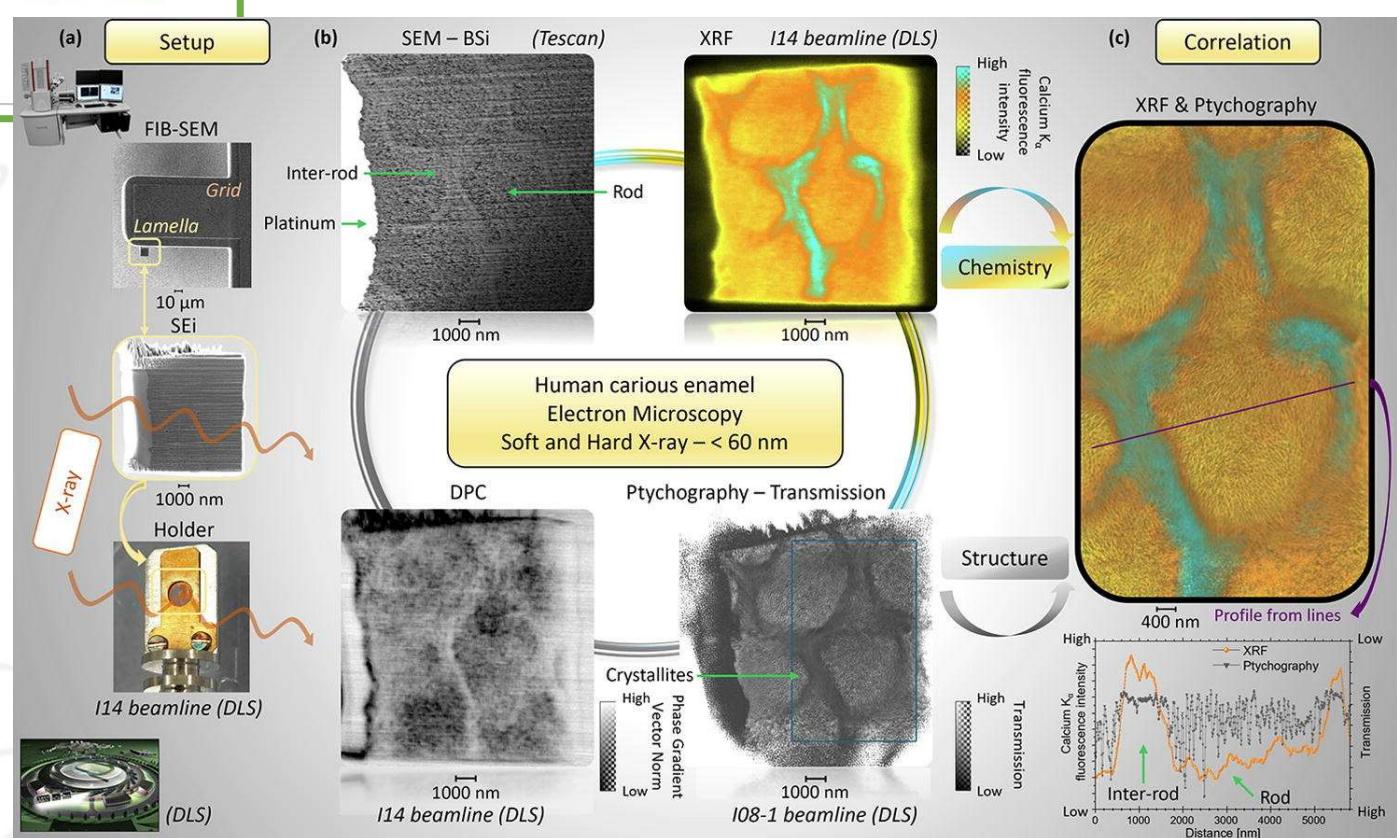
<sup>a</sup> MIREM, Department of Engineering Science, University of Oxford, Parks Road, Oxford, Oxfordshire OX1 3PJ, UK

<sup>b</sup> TESCAN-UK Ltd., Wellbrook Court, Girton, Cambridge CB3 0NA, UK

<sup>c</sup> Diamond Light Source Ltd., Didcot, Oxfordshire OX11 0DE, UK

<sup>d</sup> School of Dentistry, University of Birmingham, 5 Mill Pool Way, Edgbaston, Birmingham, West Midlands B5 7EG, UK

- Correlative XRF, differential phase contrast imaging (I14), soft X-ray ptychography (I08) and SEM of tooth enamel FIB section
- Reveals nanoscale variation in chemistry and structure across rod and inter-rod regions
- Increased demineralisation in the highly orientated rods, compared to the inter-rod 'filler' material



- Multi technique approach revealed unprecedented detail of the enamel structure and mechanism of demineralisation
- Approach applicable to not only dental research but other hierarchical biomaterials structures such as bone repair

# ITT Challenges

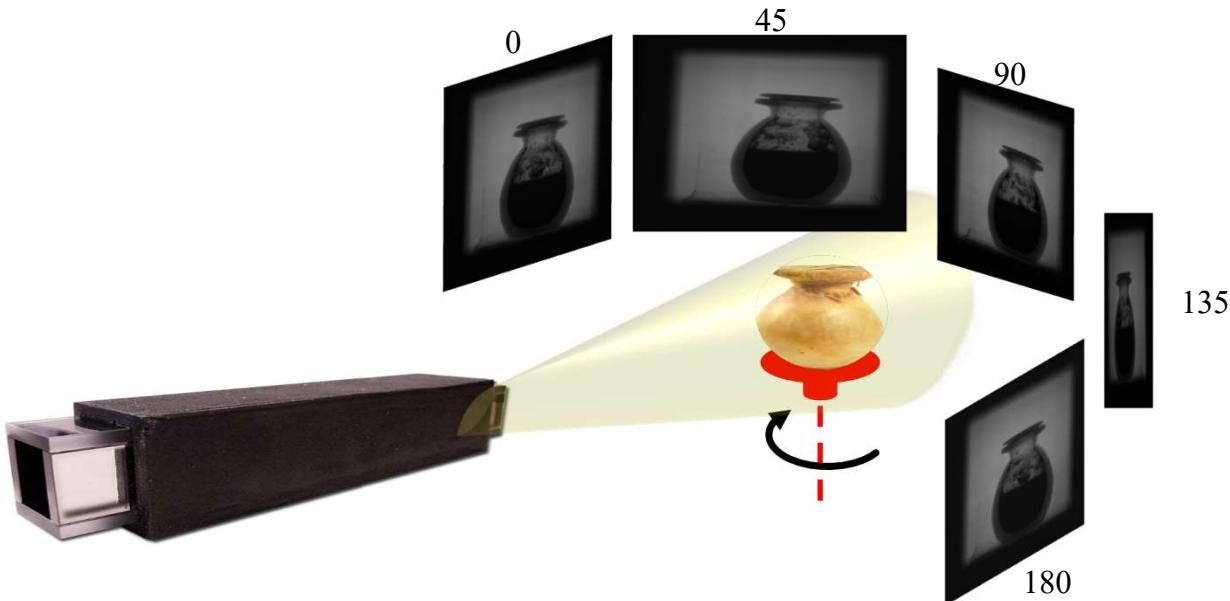
1. Efficient Tomography
2. Optimised scanning - Decision making in experiments
3. Merging data from different modalities
4. ML for ptychography?

# Challenge 1 – Efficient tomography



# X-ray and Neutron Imaging

Produce real-space representations of objects or phenomena in 1D/2D/3D

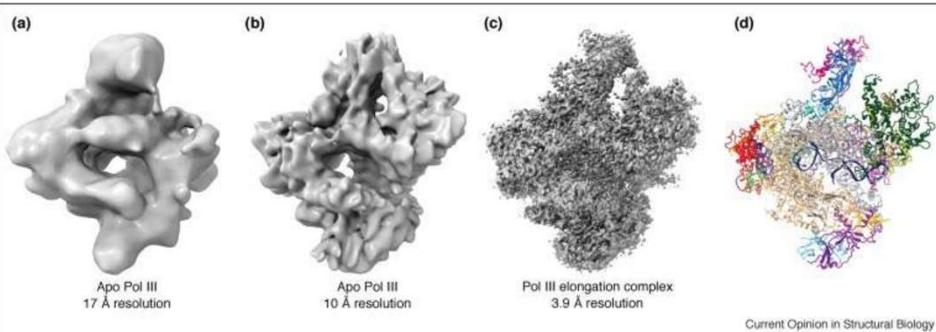
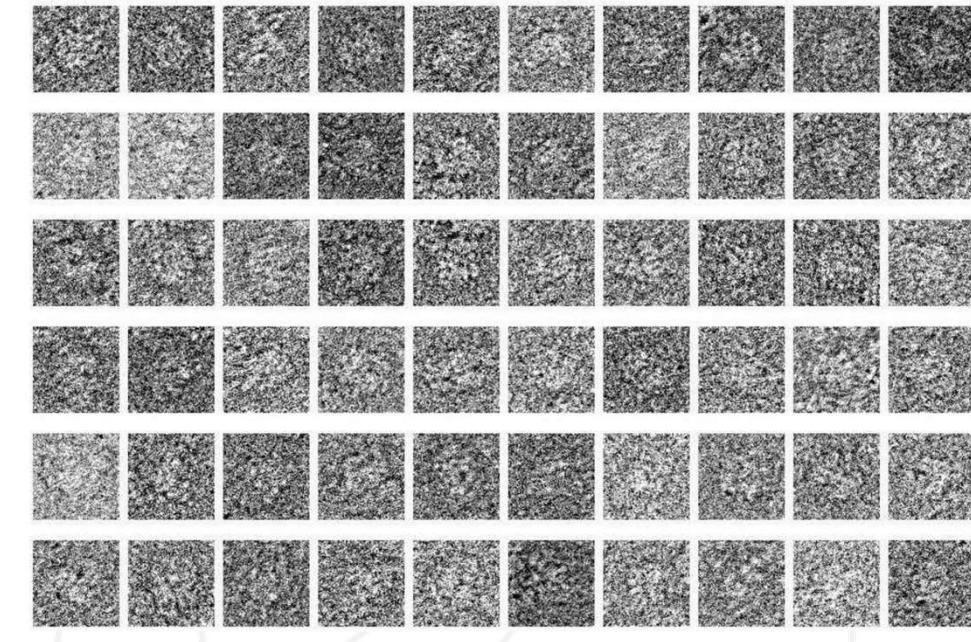


- Radiography or Projection (2D) (x,y)
- Tomography (3D) (x,y,z)
- Time-lapse radiography/tomography (3D/4D) (x,y,z, time)
- Energy-selective (“colour”) imaging (3D/4D) (x,y,z, energy)

## Dose Fractionation

- The total dose required to achieve statistical significance for each voxel of a computed 3D reconstruction is the same as that required to obtain a single 2D image of that isolated voxel at the same level of statistical significance. Hegerl and Hoppe [doi:10.1515/zna-1976-1241](https://doi.org/10.1515/zna-1976-1241). S2CID 3539651
  - Potential Impact
    - Example DIAD at Diamond -
      - Acquires 3D data in 40 mins (around 1 second per projection) but in practice could acquire in 0.01 s per projections or even 0.0001s for a total acquisition of 10-20 seconds

It works...



The Nobel Prize in Chemistry 2017

Jacques Dubochet, Joachim Frank, Richard Henderson

Share this: [f](#) [G+](#) [t](#) [+](#) [e](#) { 991 }

## The Nobel Prize in Chemistry 2017

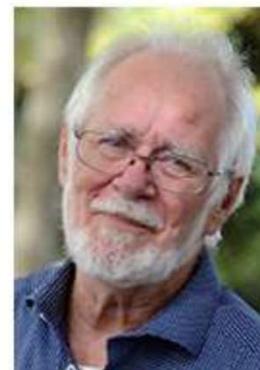


Photo: Félix Imhof © UNIL  
[CC BY-SA 4.0]

**Jacques Dubochet**

Prize share: 1/3



Photo: B. Winkowski ©  
Columbia University  
Medical Center

**Joachim Frank**

Prize share: 1/3



Photo: MRC Laboratory of  
Molecular Biology  
**Richard Henderson**

Prize share: 1/3

The Nobel Prize in Chemistry 2017 was awarded to Jacques Dubochet, Joachim Frank and Richard Henderson "for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution".

Also works in ptychography!!



# Great so why isn't it widely used for tomography

- In practice everyone takes a statistically significant 2D images
  - Why ?
    - They like to see a nice images of each projection ?
    - Don't have robust, verified approaches for dose fractionation
      - The "statistically significant" bit is also a bit vague – signal to noise, resolution etc ?
    - Spectral or Vector tomography
- ⑩ But according to dose fractionation we could be 100-1000's times less dose or 100-x1000's times faster.
- When you do an experiment there are different corrections and artefacts before/after a reconstruction
  - Distortion correction
  - Dark field (detector response removal)
  - Zingers. ( hot or dead pixels)
  - Misalignment of projections (jitter in the position of each projection)
  - Offset centre of rotation (the image is rotating but not about the centre of the image )
  - Ring Removal. (different pixel responses result in stripes or rings in the reconstructed image – some techniques try to filter the projection or sinogram )
  - The detector - e.g. a scintillator and camera – tends to have a background and intrinsic noise
    - Where we've seen fractionation work has been in photon-counting (no detector noise)

# What's missing ?

Is the way of taking data the best way to do this with the lowest dose ?

Is the model correct ?

How does the fractionation affect quantification ?

Is it the variety of data ?

How do we incorporate the corrections with the noisy data ?

Need a turn key solution





# Faster experiments

- How do I take good tomography data most efficiently ?
  - We want to take the data with the lowest X-ray dose that produces a good result
  - We want to take to know when we have enough data or how we can adapt a scan to optimally sample the data

How do we take the lowest X-ray dose ?

- **Dose Fractionation**

- The total dose required to achieve [statistical significance](#) for each [voxel](#) of a computed [3D reconstruction](#) is the same as that required to obtain a single [2D](#) image of that isolated voxel at the same level of statistical significance. Hegerl and Hoppe [doi:10.1515/zna-1976-1241](https://doi.org/10.1515/zna-1976-1241). [S2CID 3539651](#)
- In practice everyone takes a statistically significant 2D image because of how we perform tomography reconstructions but according to dose fractionation we could be 100-1000's times more efficient and faster.
- But there tend to be other step-wise processing of data and the reconstruction
  - Distortion
  - Zingers
  - Misalignment of projections or offset centre of rotation
  - Phase contrast effects + modalities
  - Spectral/diffraction tomography ?

Here I'm assuming that we acquire 1 good image and then divide up the dose over the projections -  
Maybe there is a better way ?



How do I take the X-ray data efficiently?



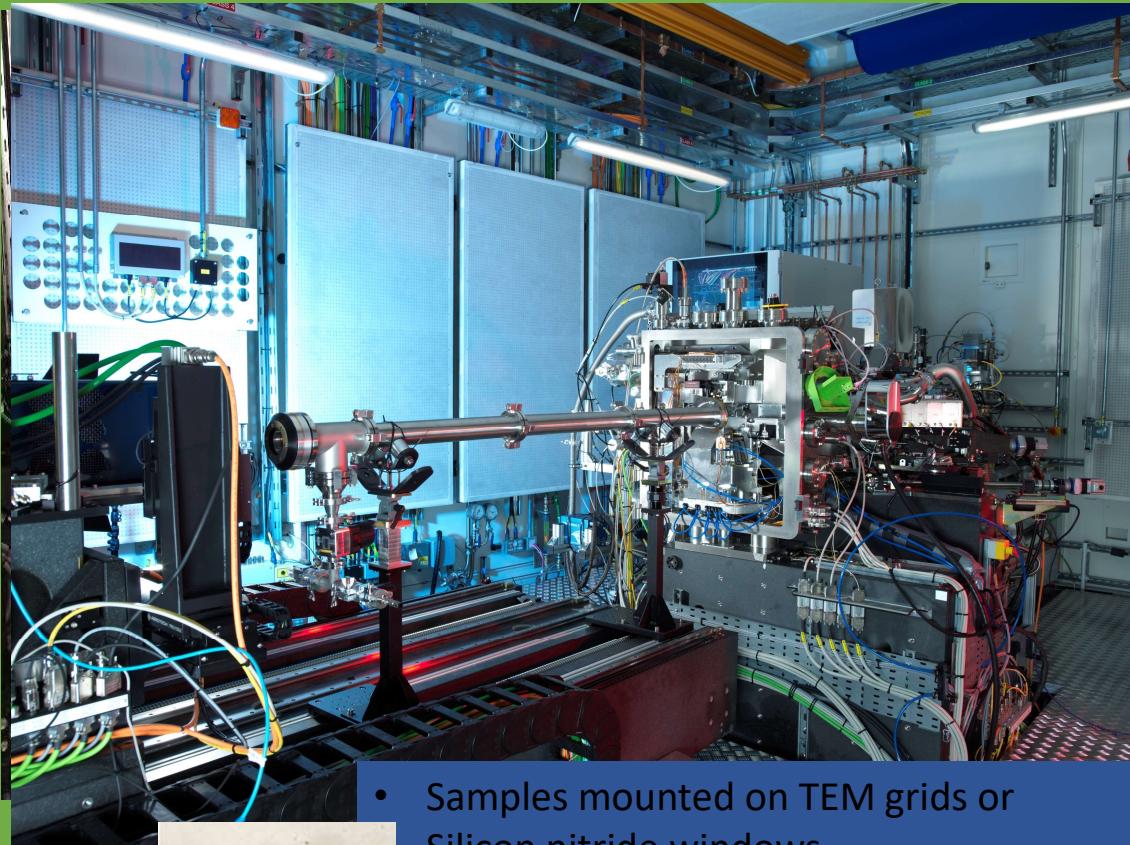
# Faster experiments

- Can you "learn" ptychography ?
  - We would like to make a decision on if/how to develop AI solutions for ptychography
  - Traditional computing is much slower than experiment acquisition so how do we bridge this ?
- Examples of Machine learning "inference" used but always trained on the local data or image libraries (cats, dogs, flowers etc) . Some progress with PINN's but ultimately nothing actually in use for real applications
  - Normal reconstruction always run to confirm
- The question is can we trust the result or have artefacts been generated ?
- If you know, for example each Fourier transform pattern is 256x256 and 16 bits do I need 256x256x65536 patterns or more to train an all knowing, completely accurate inference model ?
- What about down sampling to 64x64, for example to get a very good estimate ?
- Can you provide a principled estimation of model uncertainties ?
- Is there anything that can be said mathematically about this ? What criteria do we need to infer a result ?Compressibility, the route needed to build an accurate ptychography inference model ?

## Challenge 2: Decision making in experiments

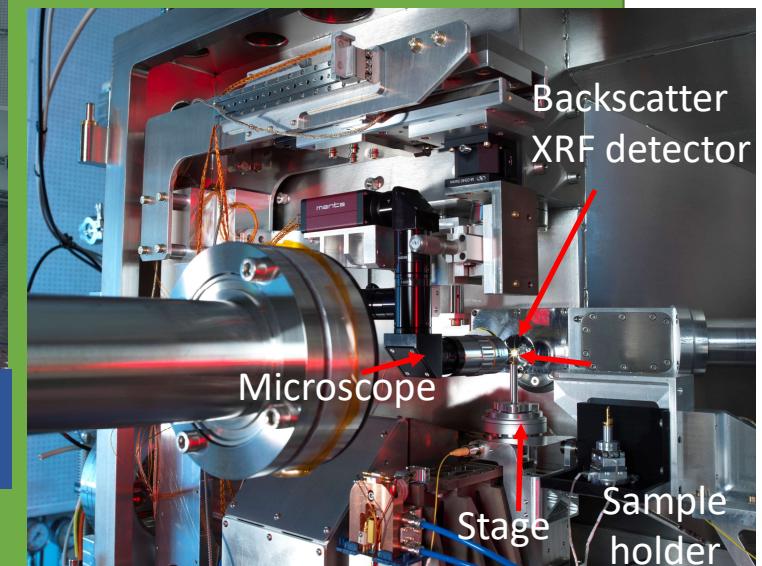


## X-Ray Nanoprobe

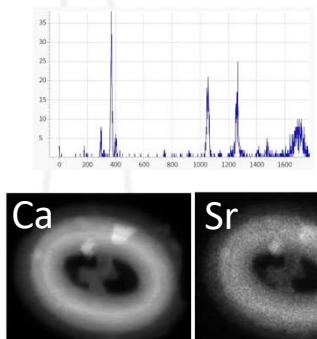


- Samples mounted on TEM grids or Silicon nitride windows

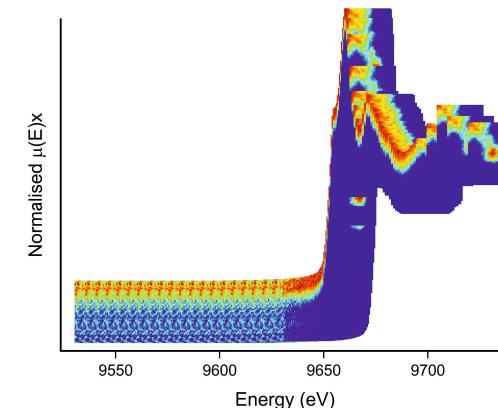
A 50nm X-ray beam for XRF mapping, spectroscopy, diffraction, imaging



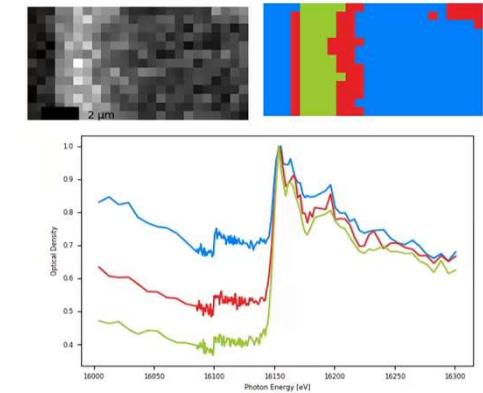
# *What can I do with it?*



## X-ray fluorescence -Elemental composition



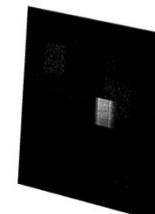
## X-ray spectroscopy -Chemical speciation



## Detector



## Diffraction/SAXS -Crystal structure, strain, orientation -Mineral phase identification



## Phase contrast imaging (DPC) and ptychography

- Structural features and density
- Image context for composition

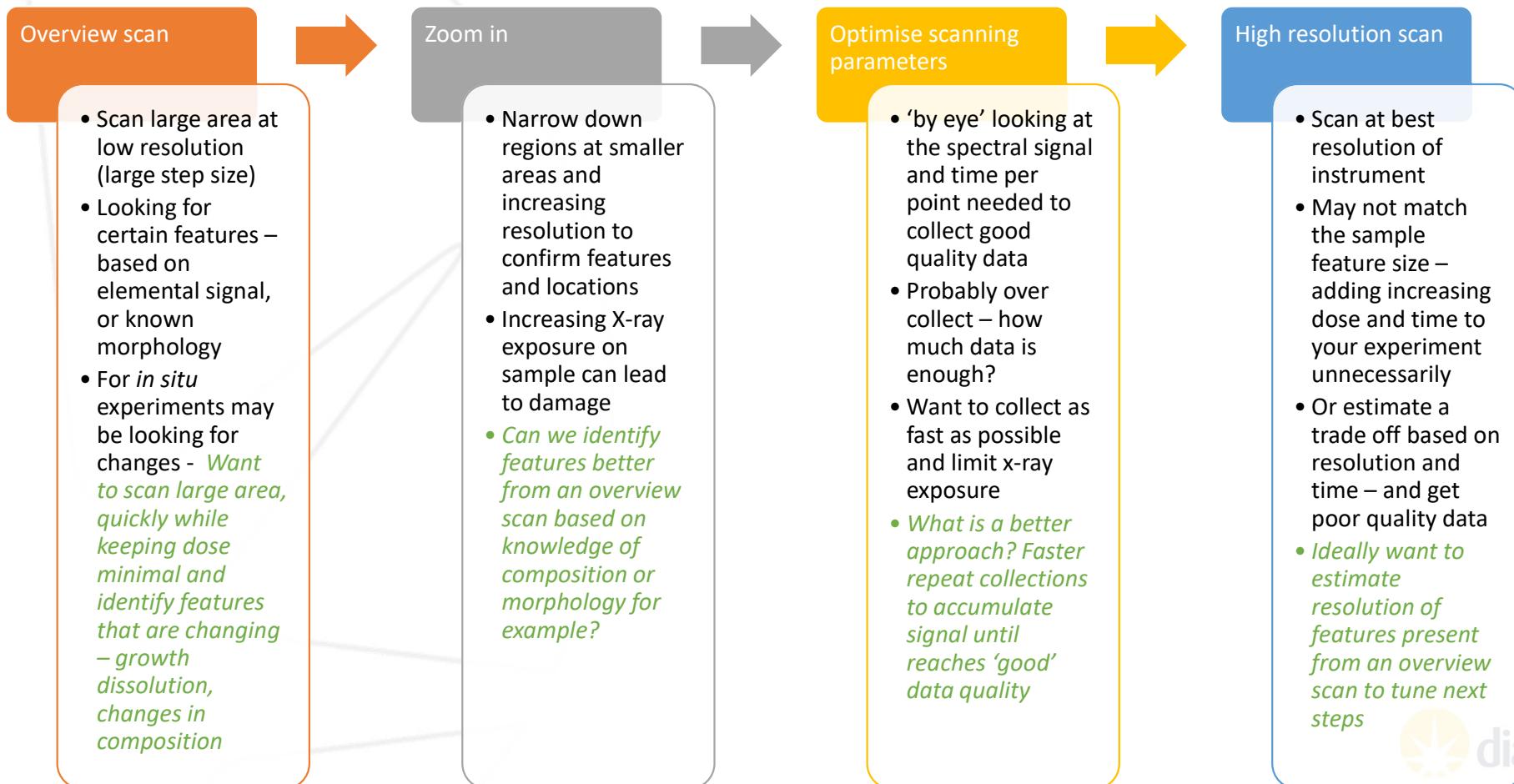


Computation

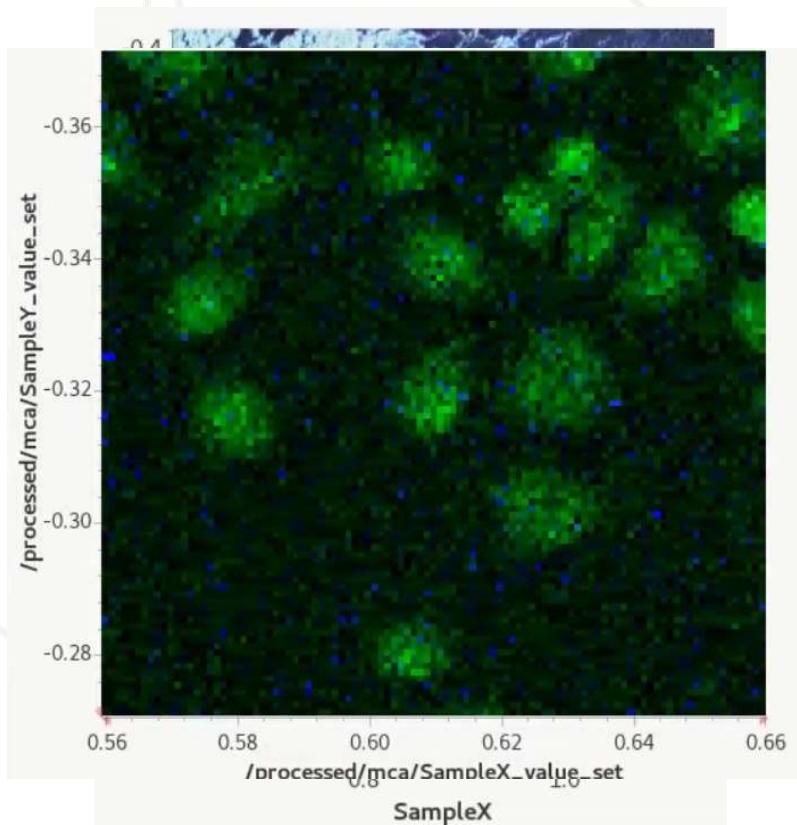
 diamond

# Challenge 2: Decision making in experiments

## A typical X-ray scanning probe experiment



# Challenge 2: Decision making in experiments



## What would an experiment look like?

Example: Finding cells

Cells on a window – treated with a metal containing drug (Pt)

Overview scan

Large area scan XRF

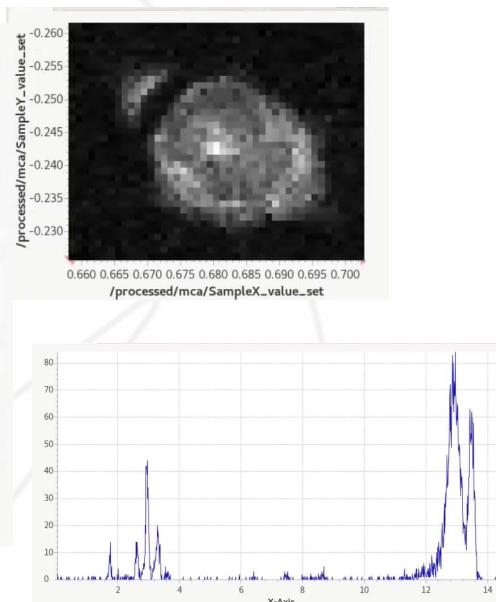
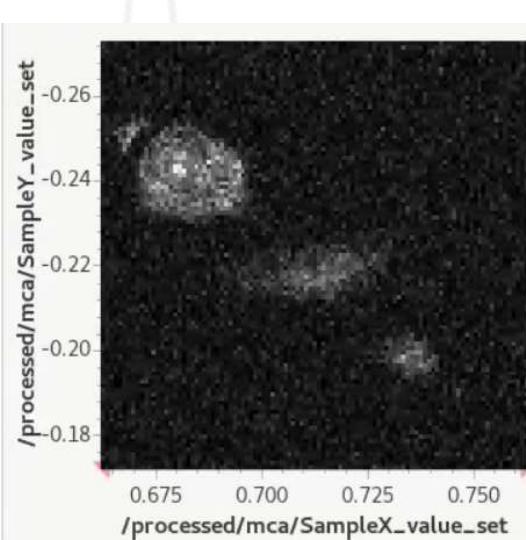
- Coarse step size (several  $\mu\text{m}$ )
- use Potassium signal to ‘see’ cells, Platinum signal much weaker

- Can we collect this overview scan to be able to
  1. identify potential cell locations
  2. Optimise the collection time
  3. Identify the highest resolution of feature size

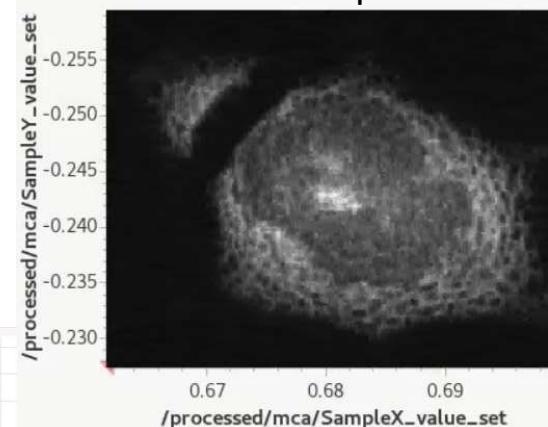
Zoom in

Optimise scanning parameters

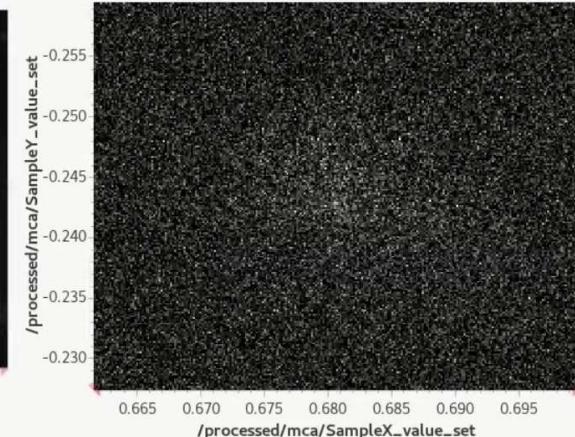
High resolution scan



Potassium map



Platinum map

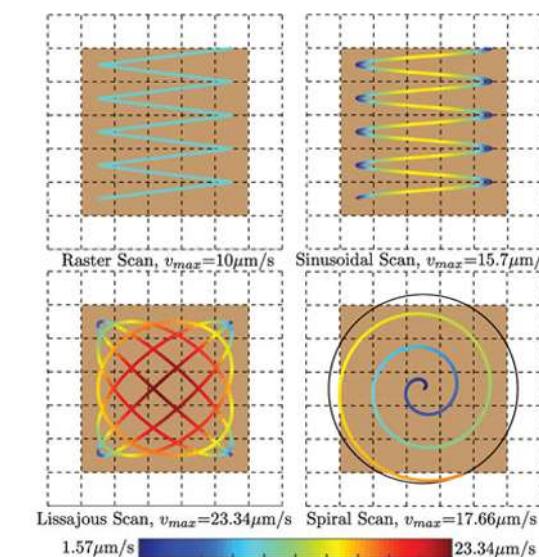
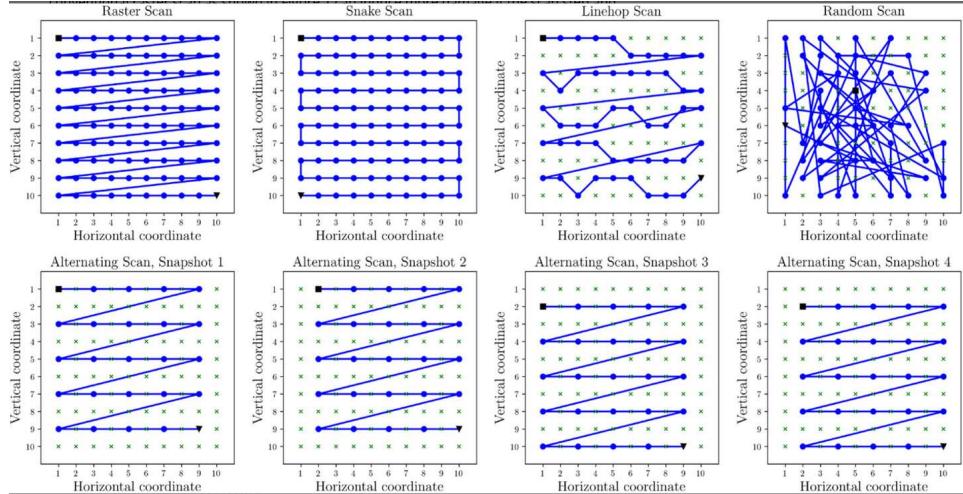


To obtain this high resolution image the cell has been mapped 3-4 times  
Extra time and dose to the sample

Would like to map multiple cells automatically while minimising time and dose.

# Scanning limitations...

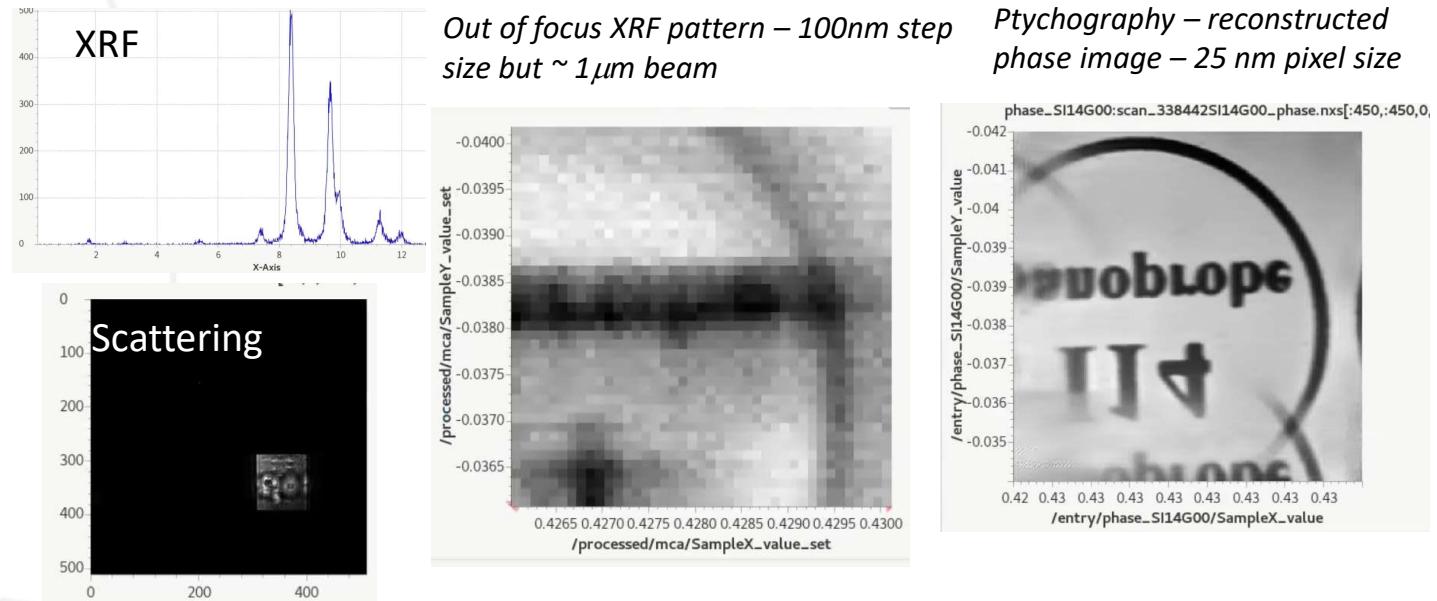
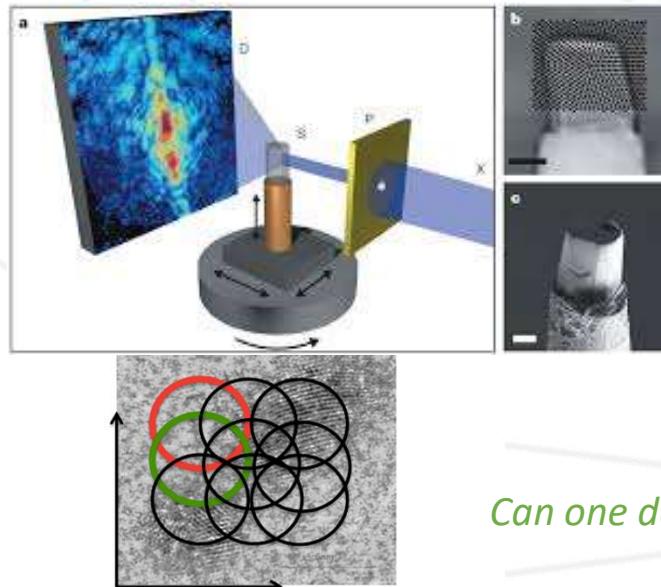
- Scan in constant velocity mode
- Typically raster or snake scan
- Limitation on speed of motion (as well as detector frame rates)
- Other scan paths are possible – but NOT random position
- Have to consider stage velocity
- Scanning pattern can also determine how damage from X-ray dose is distributed in the sample
- X-ray focal spot size – fixed (50nm)
- Can translate sample along beam direction (out of focus) to change spot size on the sample (sort of)



# Challenges 3: Merging data from different modalities

Aim = resolution enhancement and faster experiments

- Multi modal data sets collected on same instrument, simultaneously or sequentially, data collection times or resolutions do not always match e.g. ptychography data higher resolution than XRF data on same sample
- Collect ptychography data out of focus, need to repeat a scan at focus to get high resolution XRF on the same region



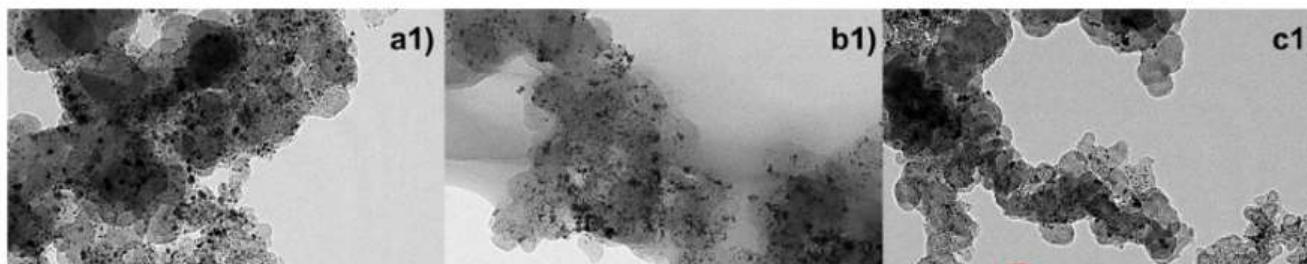
Can one data set be used to 'improve' the other? Reducing the time taken for a experiment

# Challenges 3: Merging data from different modalities

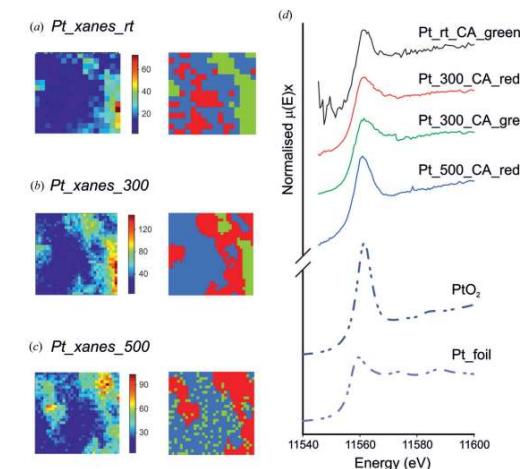
Aim = resolution enhancement and faster experiments

Different instruments- have different sensitivities and spatial resolutions eg TEM and X-ray nanoprobe:

- TEM imaging gives a high resolution image – X-ray nanoprobe is lower resolution but larger FOV and spectral sensitivity
- Image in TEM, then same area in X-ray nanoprobe for spectroscopic mapping – see aggregates of smaller particles



*Can the prior knowledge from high resolution TEM imaging be used to resolve the spectroscopy mapping?*



# Challenges 3: Merging data from different modalities

Aim = resolution enhancement and faster experiments

From 2D to 3D information?

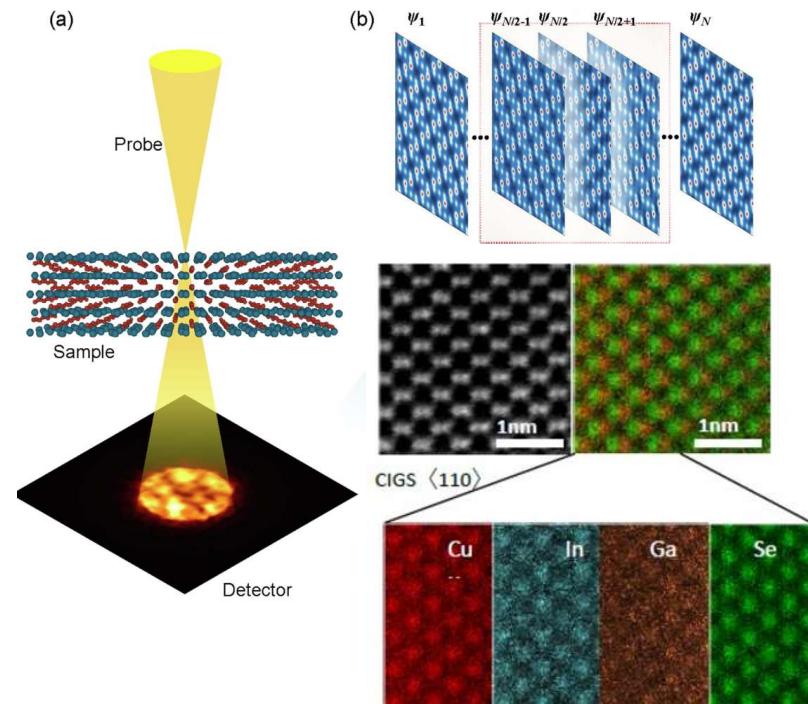
TEM: Multislice Ptychography

The 3D transmission function provides a detailed picture of the precise atomic locations in the sample but limited information regarding the chemical species of the atoms.

Energy-dispersive X-ray (EDX) spectroscopy give chemical information but no depth information – 2D projection

Ptychography data and xray spectroscopy data are acquired simultaneously and we are able to recover a 3D map of electron scattering probability - believe it should be possible to recover three dimensional information regarding the emission location of the xray photons.

*problem = map the 2D projected x-ray fluorescence onto the 3D scattering probability map.*



Related problem for X-ray nanoprobe – can have an 3D phase or elemental volume – but collecting 3D diffraction data is much slower

A diffraction pattern is collected in transmission – so a projection through the volume

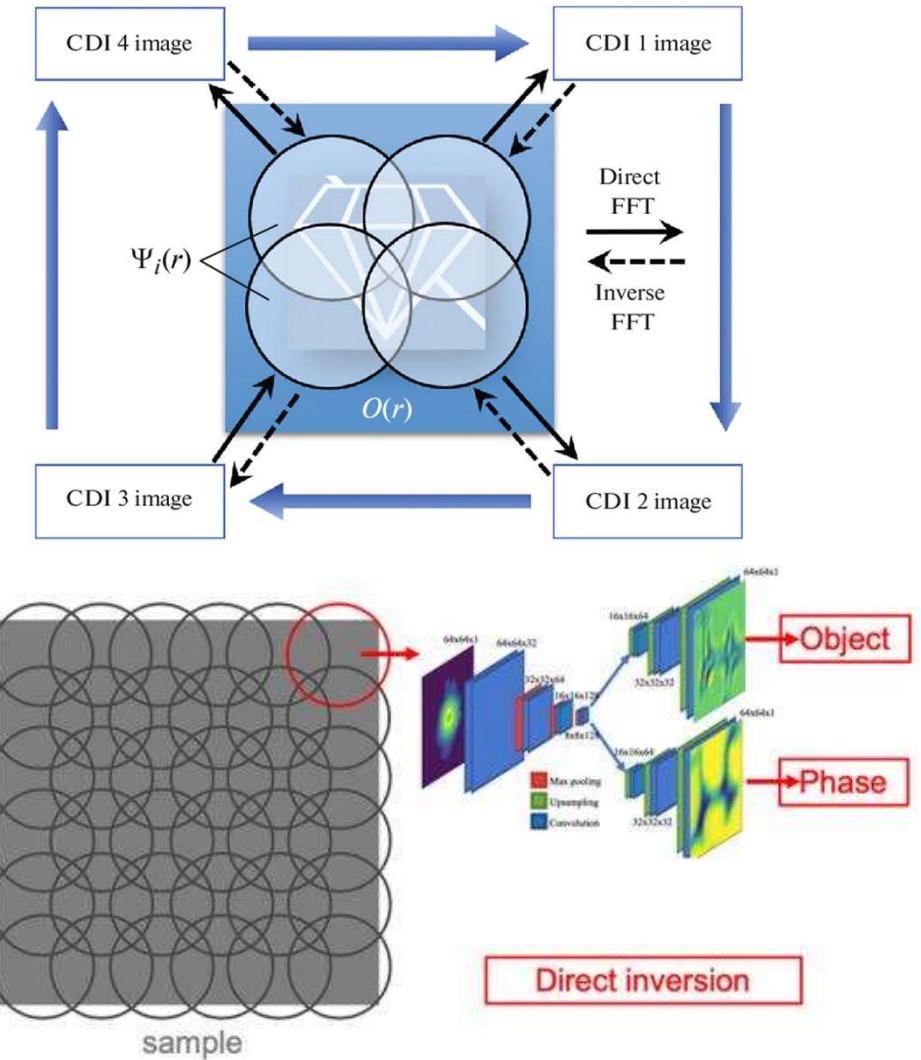
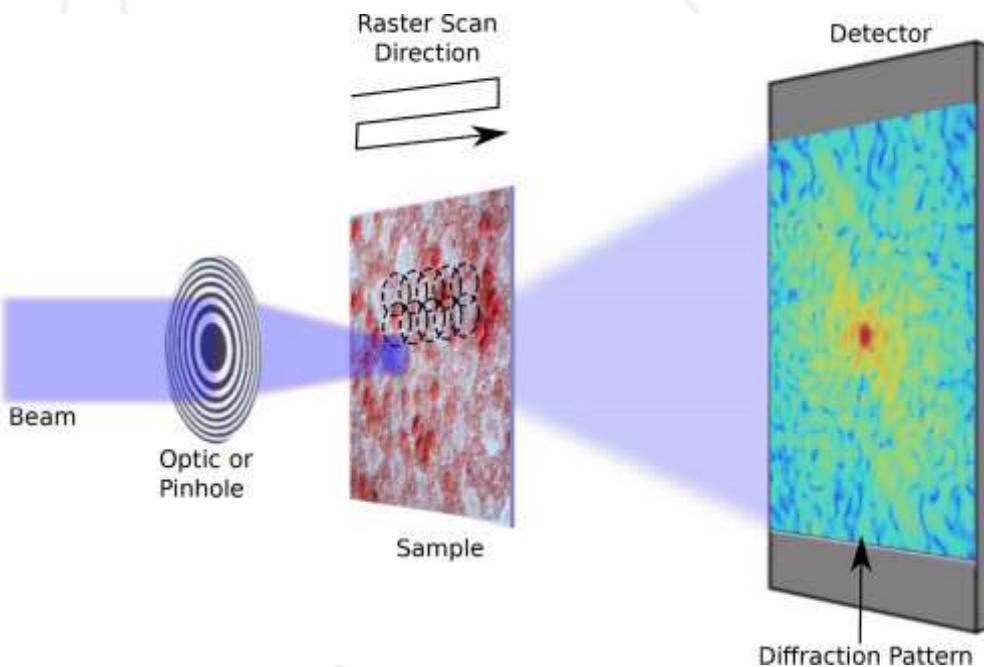
# Challenge 4 – Machine learning ?



# Challenge 4 – Machine learning ?



# Ptychography



predict the sample amplitude and phase from data.



# The problem

We want to use ptychography more broadly, but it takes much longer to reconstruct an image than it does to take the data.

In 3D can take 30 mins for acquisition, 24 hours for reconstruction...

How can we jump the time to solution forward ?

Where do we focus ?

# The problem

We want to use ptychography more broadly, but it takes much longer to reconstruct an image than it does to take the data.

In 3D can take 30 mins for acquisition, 24 hours for reconstruction...

How can we jump the time to solution forward ?

Is this ML or some hybrid ?

## Ptychography challenges

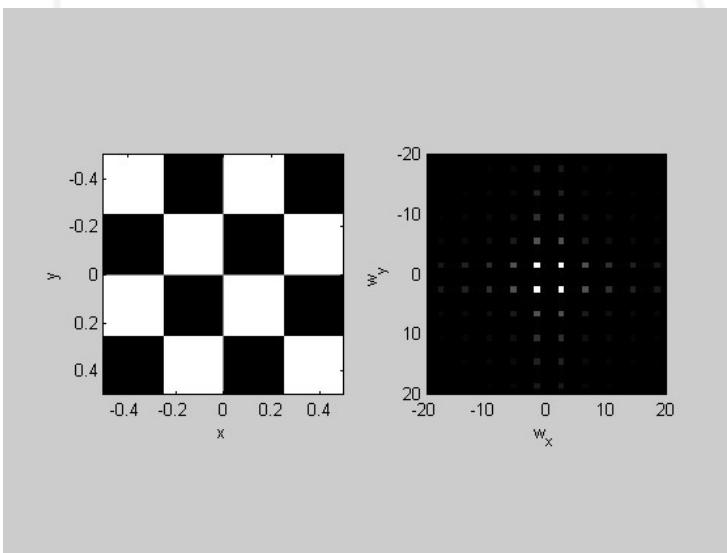
We would like to make a decision on if/how to develop AI solutions for ptychography

- Can you "learn" ptychography ?
  - Inference is x100 times faster than conventional approach BUT only demonstrated on local problem
    - I.e. train model on the data you have and you can show it can infer similar data
- Can you trust an AI solution ?
  - **The question is can we trust the result or have artefacts been generated ?**
  - Can you provide a principled estimation of model uncertainties ?
  - Is there anything that can be said mathematically about this ?
  - What criteria do we need to infer a result, how much data would be needed to build a perfect model ?
- If you can't trust it - what the alternative ?



# Question – ptychography small beam case

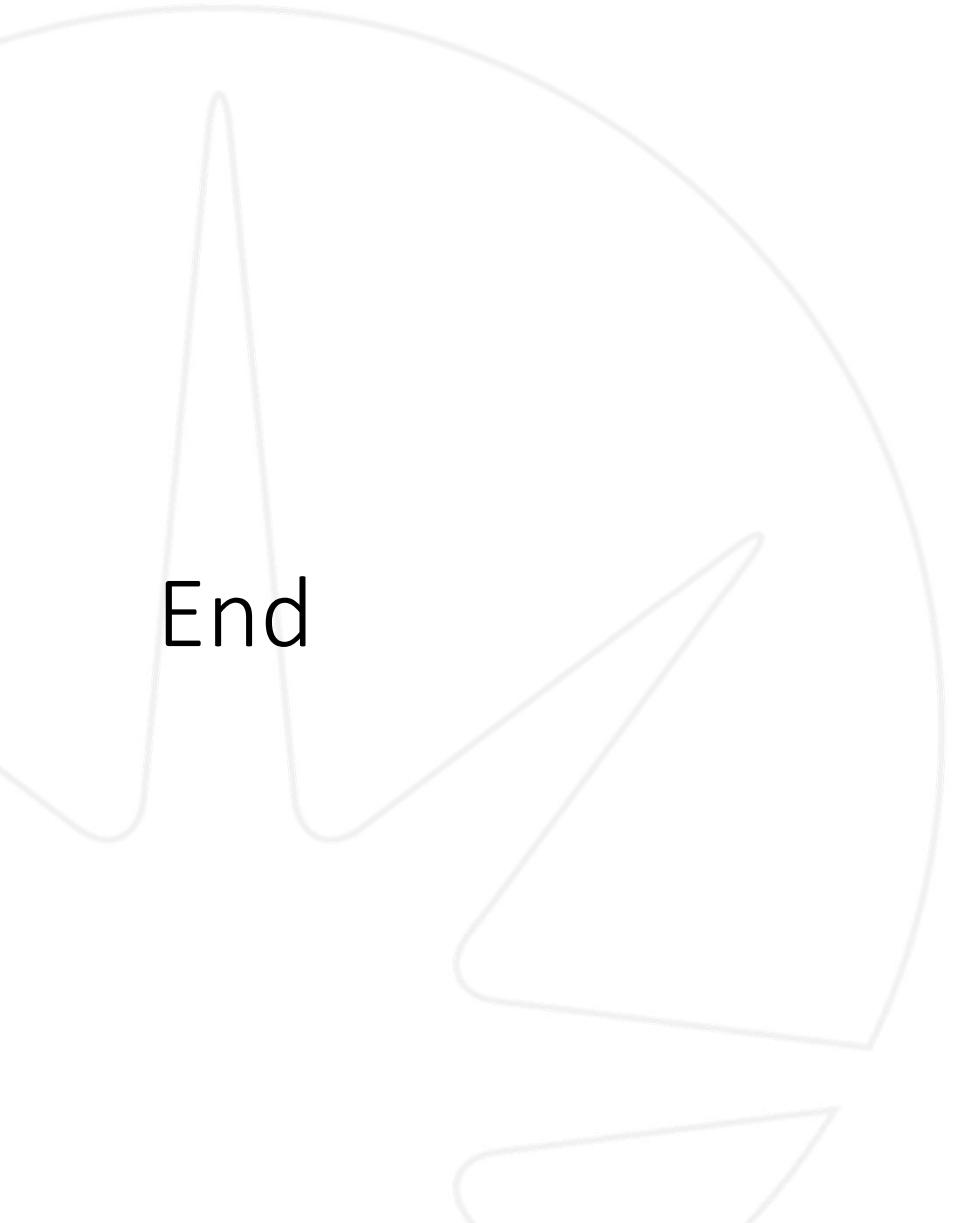
- At a nanoprobe the beam is 50 nm
- The resolution of ptychography could be 25 nm



- If the beam is small is it easier to "learn" or invert from FFT to phase and amplitude ?
- Can I say when this has failed ?

# question – ptychography from image libraries

- Online libraries for AI – million images
- If I train an AI on a million images is it enough ?
  - What are the limitations ?



End





# The Materials Project

## The Materials Project by the numbers

MATERIALS

154,718

REGISTERED USERS

500,000+

INTERCALATION ELECTRODES

4,351

CITATIONS

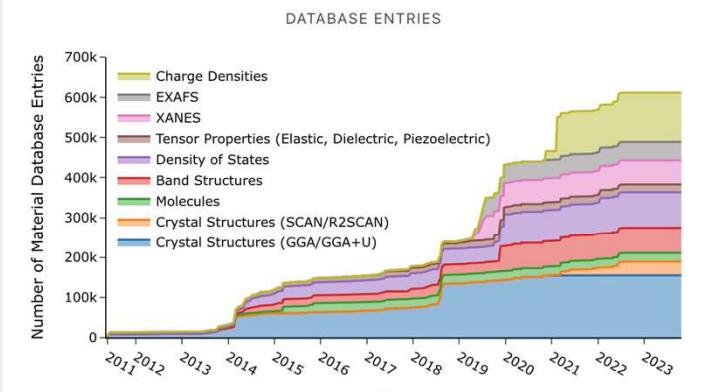
42,000+

MOLECULES

172,874

CPU HOURS/YEAR

100 million



web-based access to  
computed information  
on known and  
predicted

Computationally  
expensive DFT, TD-DFT  
calculations

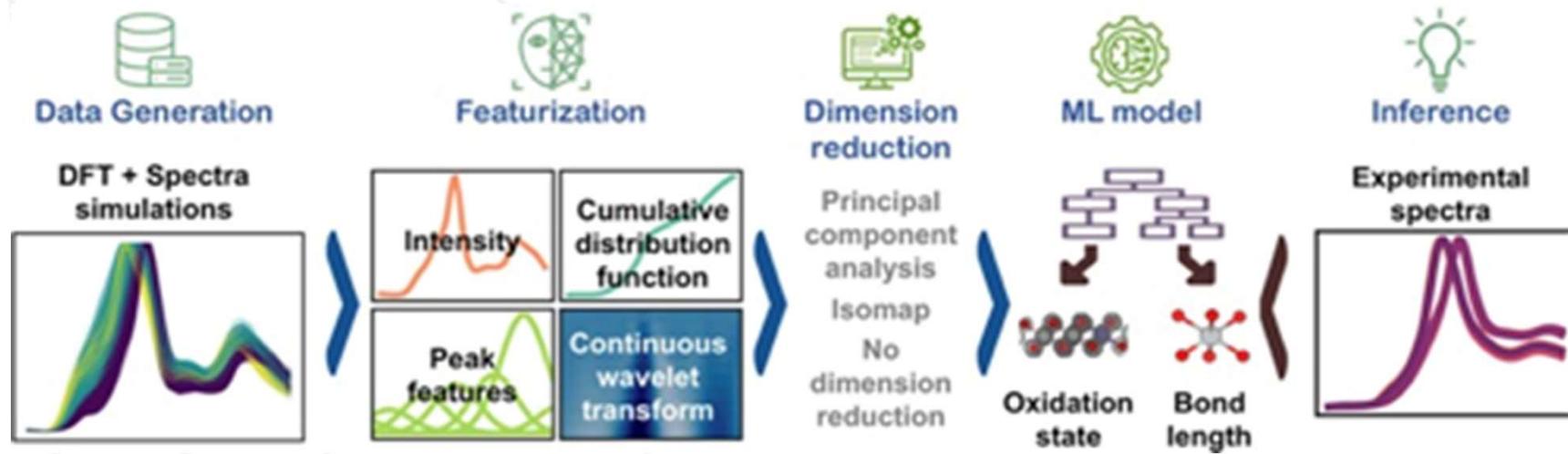
## Access data directly using the Materials Project API

Anyone who registers with the Materials Project automatically receives free and unlimited access to our API. The API lets you pull data directly from our database and use it in your own applications. Python users can download the [MPRester](#) API client to make pulling data from the API even easier. [Read more on our API page.](#)

```
from mp_api.client import MPRester
with MPRester(api_key="your_api_key_here") as mpr:
    # retrieve SummaryDocs for a list of materials
    docs = mpr.summary.search(material_ids=["mp-149", "mp-13"])
```

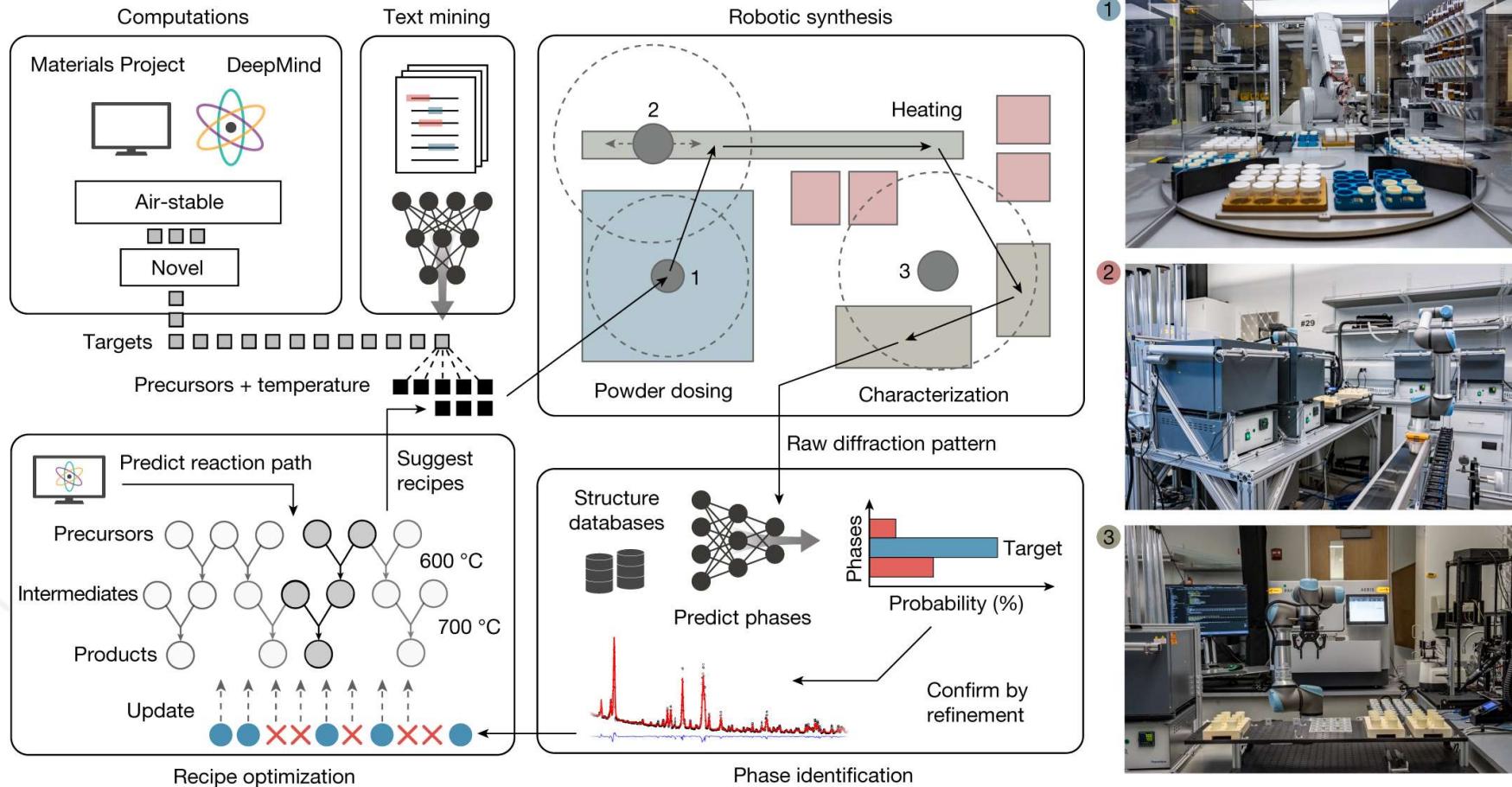


# From spectra to structures - inversion



<https://doi.org/10.1021/acs.chemmater.3c02584>

# Autonomous materials discovery with the A-Lab.

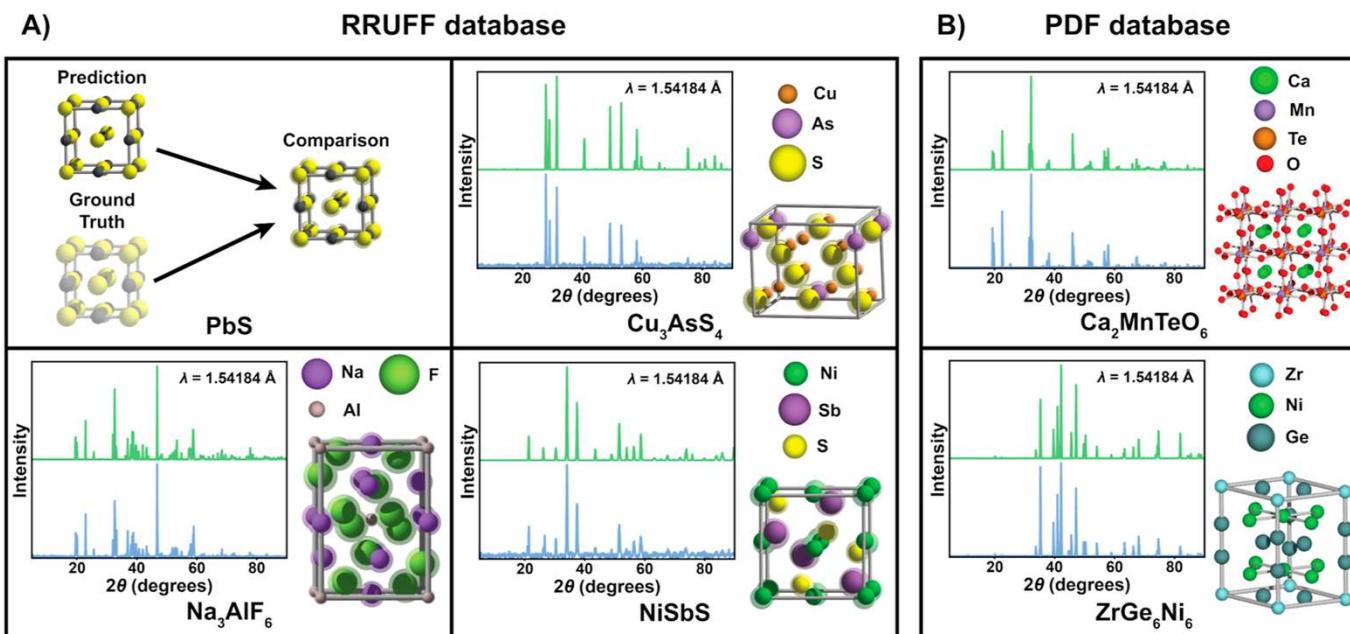


Based on the  
DeepMind  
discoveries

Szymanski, N.J., Rendy, B., Fei, Y. et al. An autonomous laboratory for the accelerated synthesis of novel materials. *Nature* **624**, 86–91 (2023).  
<https://doi.org/10.1038/s41586-023-06734-w>



# Crystal Structure Determination from Powder Diffraction Patterns with Generative Machine Learning



Not a replacement for Rietveld refinement (traditional optimisation)

A high-quality initial structural model.

67% accuracy

Does not explicitly refine preferred orientation, atomic positions, or fractional occupation

# Identifying phases and features

Many experiments want to automatically identify features or changes -

For us it's about decision making on resources to solve problems -

- For spectral examples - Individual structural phases or element features occur at set positions
- The default is for people to know try examine spectra using AI to match spectral features to individual features
- You can alternatively do linear algebra and solve N individual features into a spectra - but that needs some probabilistic approach or ranking procedure to say combinations are probable ?
- What is the best approach for spectral data ?