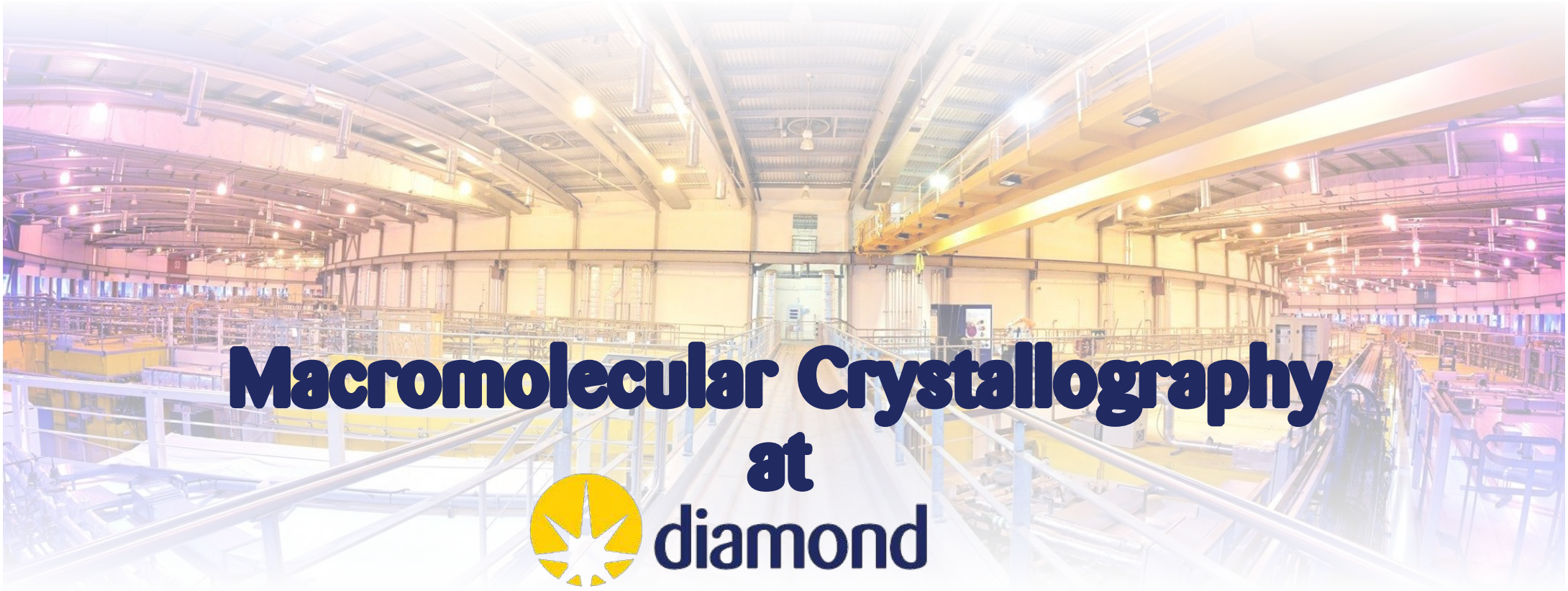
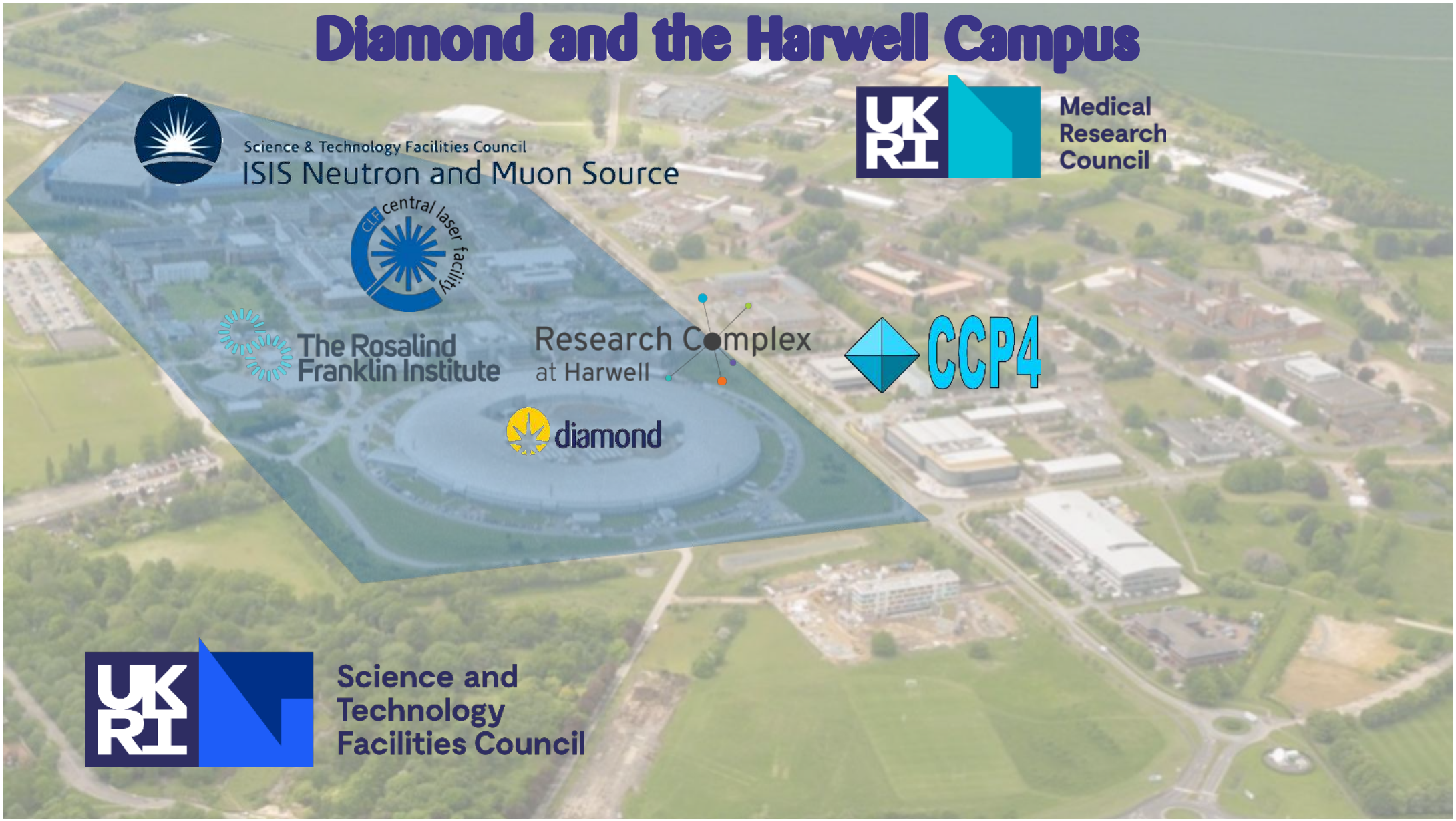


# Macromolecular Crystallography at diamond

Marco Mazzorana

Diamond – CCP4 workshop, 29/11/2021

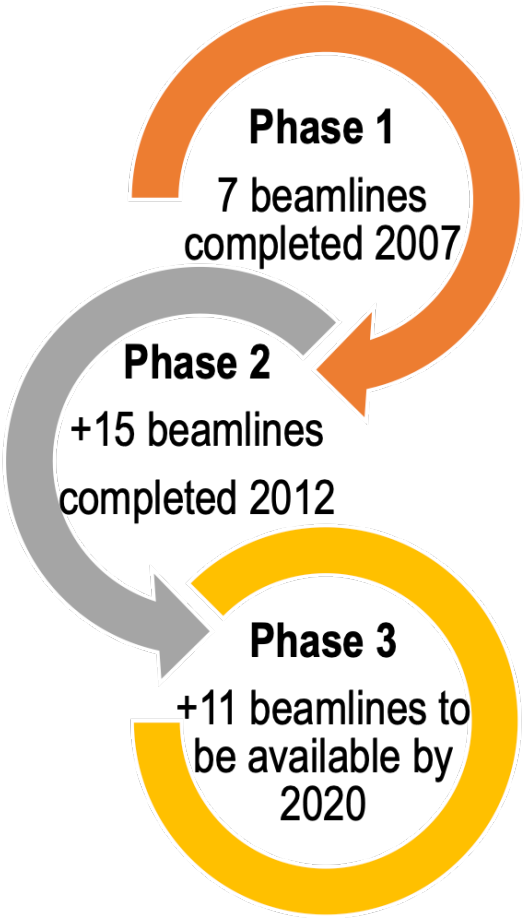

Diamond and the Harwell Campus
Science & Technology Facilities Council
ISIS Neutron and Muon Source
central laser facility
CLF
The Rosalind Franklin Institute
Research Complex at Harwell
diamond
UKRI
Medical Research Council
CCP4
UKRI
Science and Technology Facilities Council

# What is Diamond

## The UK national synchrotron

(86% STFC, 14% Wellcome Trust)

# 15+ years of development




# Many different instruments…

33 beamlines (8 science groups)

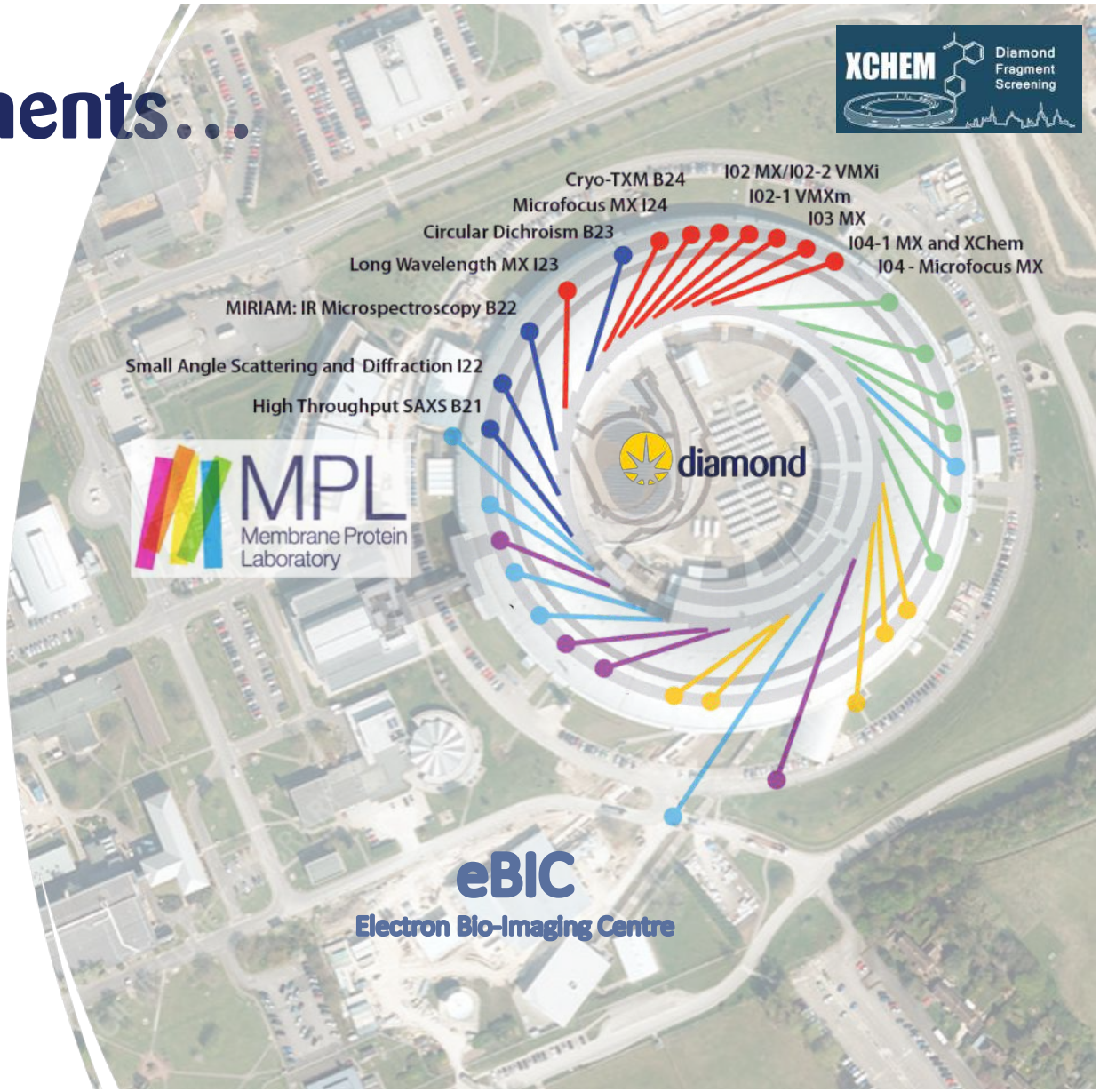
2 national cryo-EM facilities

12k+ users

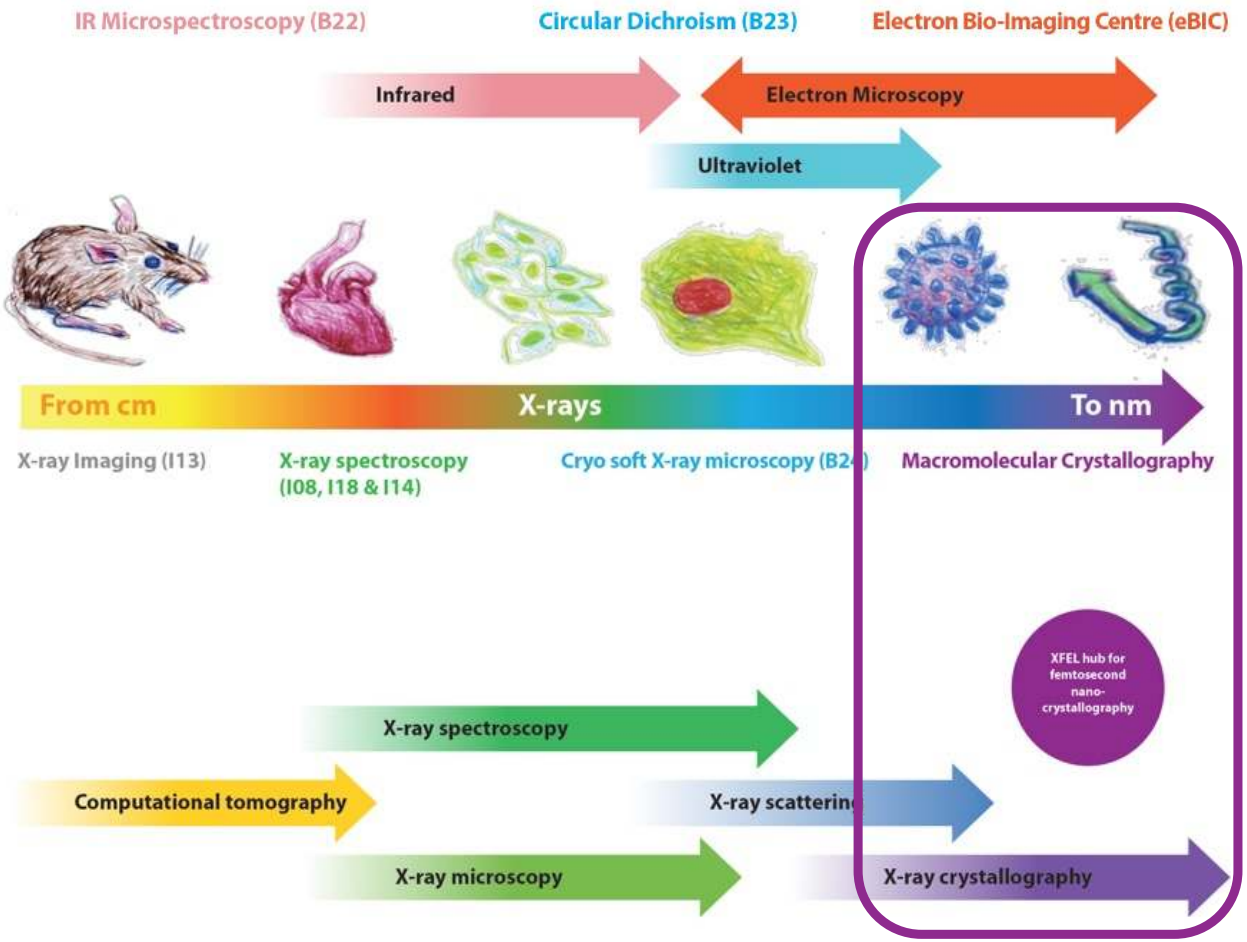
100+ companies

Structural biology-dedicated labs
(MPL, crystallization facility, XChem)

UK-XFEL Hub

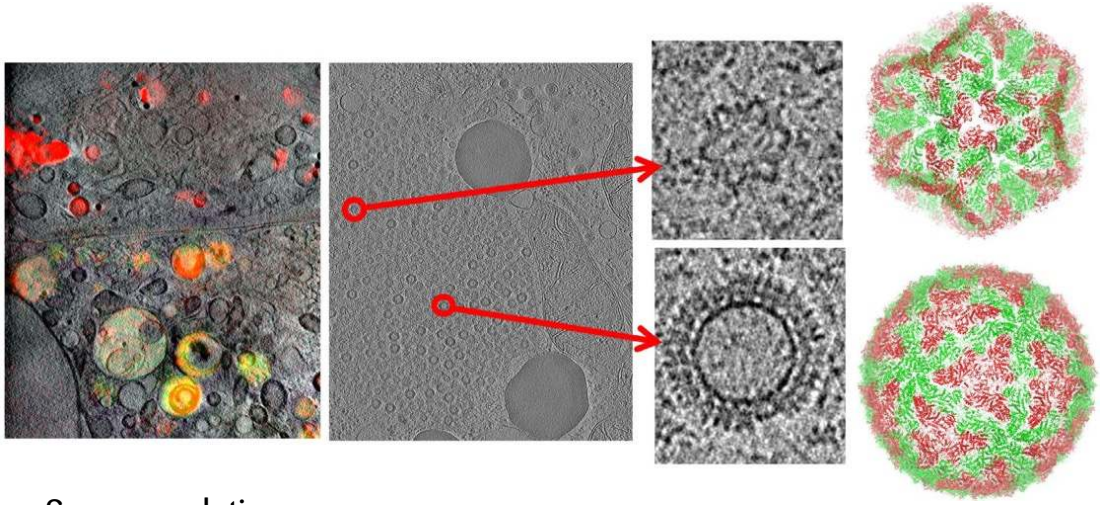


# …to investigate the structure of biological systems




diamond

# Combination of techniques



Super resolution cryo-light microscopy

Cryo X-ray tomography

Cryo-electron tomography

Atomic models from MX fit to averaged particles of two intermediates

Understand interactions and action across a large range of spatial and time domains (often weak and transient)

Reveal inner workings of the cell with a full dynamic picture of living organisms

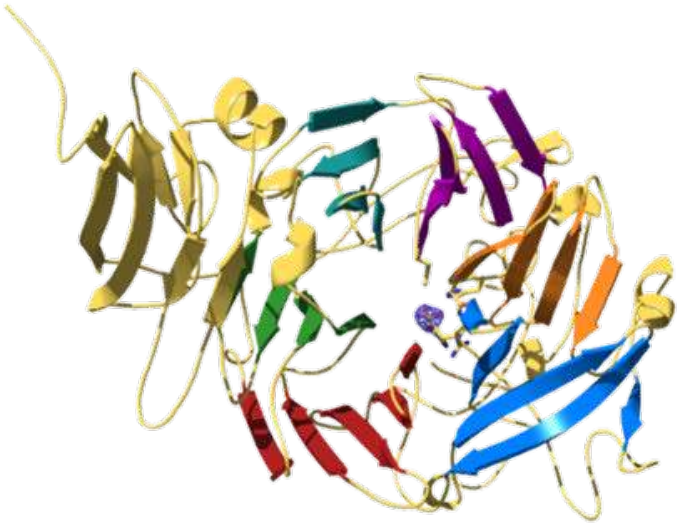
Molecular description of cellular organization can transform fundamental science and understanding of disease processes

# What crystallography tells us about macromolecules

## Shape

Crystal packing

Putative interaction surfaces



## Electrostatics

Interactions with other molecules

Docking sites

## Fold

Putative function

Evolutionary relationships

## Protein-ligand complexes

Functional sites

Druggability

## Local structure

Catalytic mechanism

Conformation determinants

diamond

# Challenges and synergies

**High Throughput**
- PADs
- Automation
- Pipelines
- Serial Crystallography
- Access (rapid/freq)

**Microfocus**
- Tuneable
- Rapid
- Variable focus
- High flux
- *In situ*

**Long Wavelength**
- Tuneable beamlines
- I23 dedicated
- Smart data collection strategies
- Expert systems

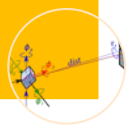
***In situ***
- VMXi
- I24
- I03
- Screening
- Data collection
- CLIII

**Experimental Phasing**
- Long wavelength
- Multi-axis strategies
- Inverse beam
- Wedged MAD
- Pipelines

**Biological Containment**
- *In situ*
- Containment Level II & III
- Schedule 5

**Complementary information**
- Humidity control
- Spectroscopy
- Tomography
- Dynamics
- Fluorescence

**Dials Project**
- Diffraction image processing for SR Electron and XFEL data
- Collaborative effort
- Led and managed by DLS



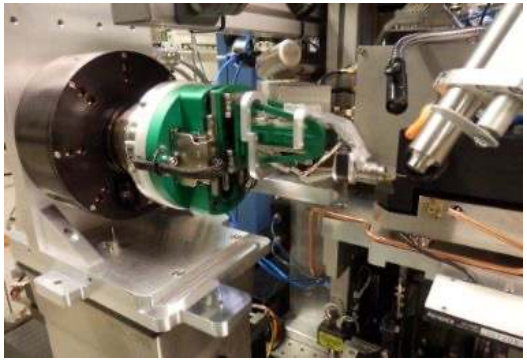
**Data Management**
- Experiment database
- Synchweb
- Data archive
- Data access
- Data reprocess

# Doing smaller things better: microfocus

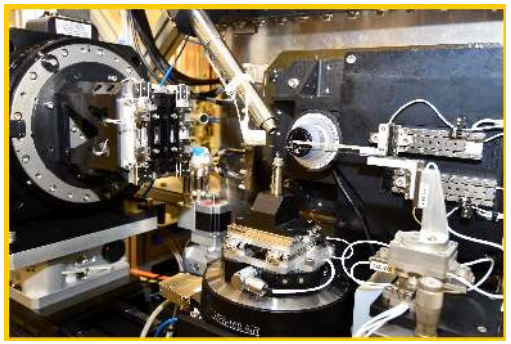
- Tuneable
- Rapid
- Variable focus
- High flux
- *In situ*

Microfocus

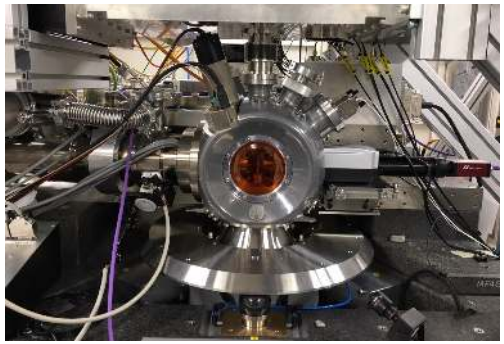
3 microfocus beamlines:



I04



I24

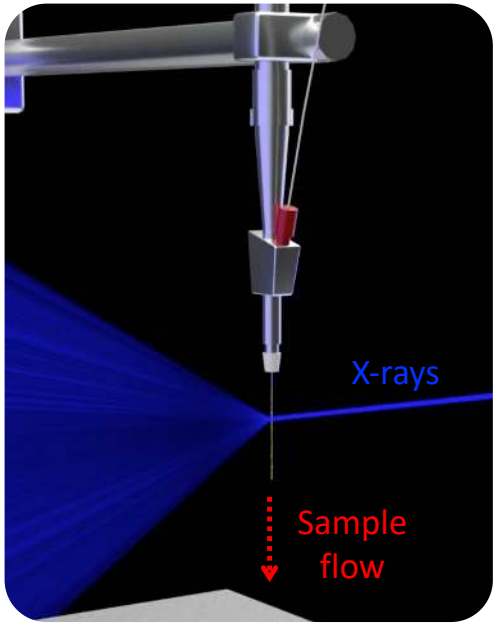


VMXm

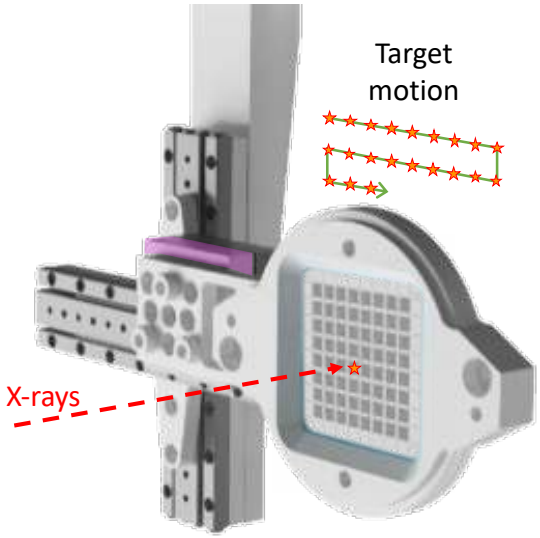
# Microfocus for time-resolved studies

## Serial crystallography (SSX and SFX)


Acoustic droplet ejection

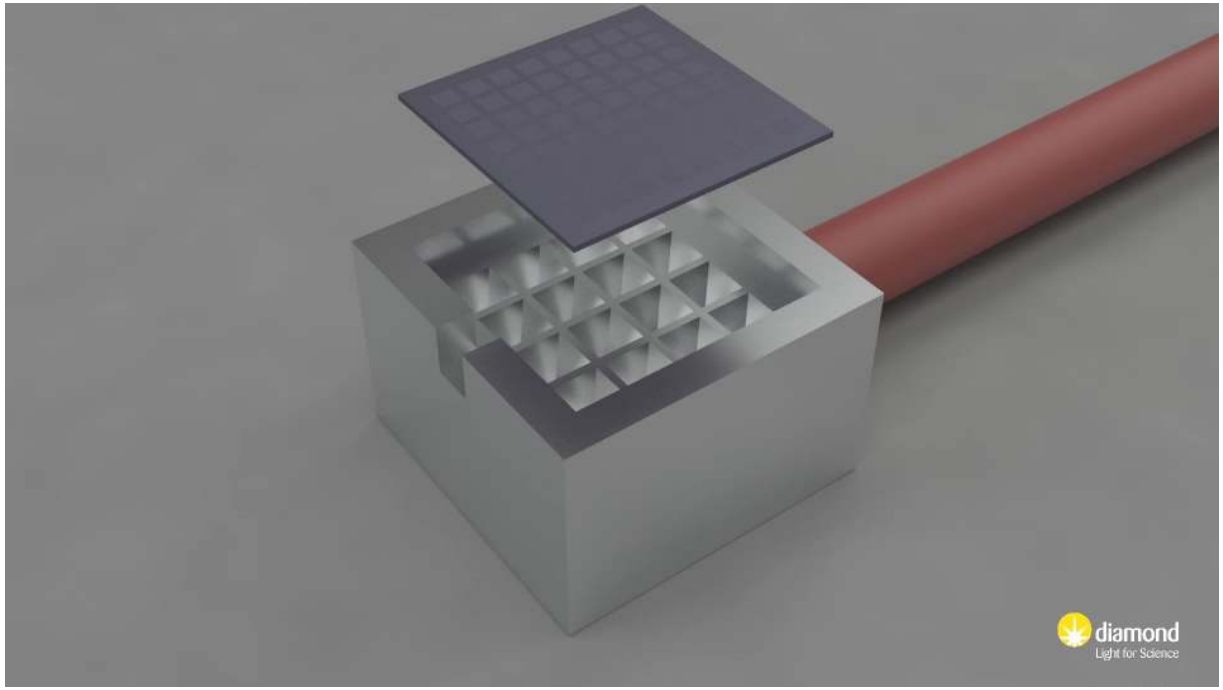



Liquid jets/ LCP extruders

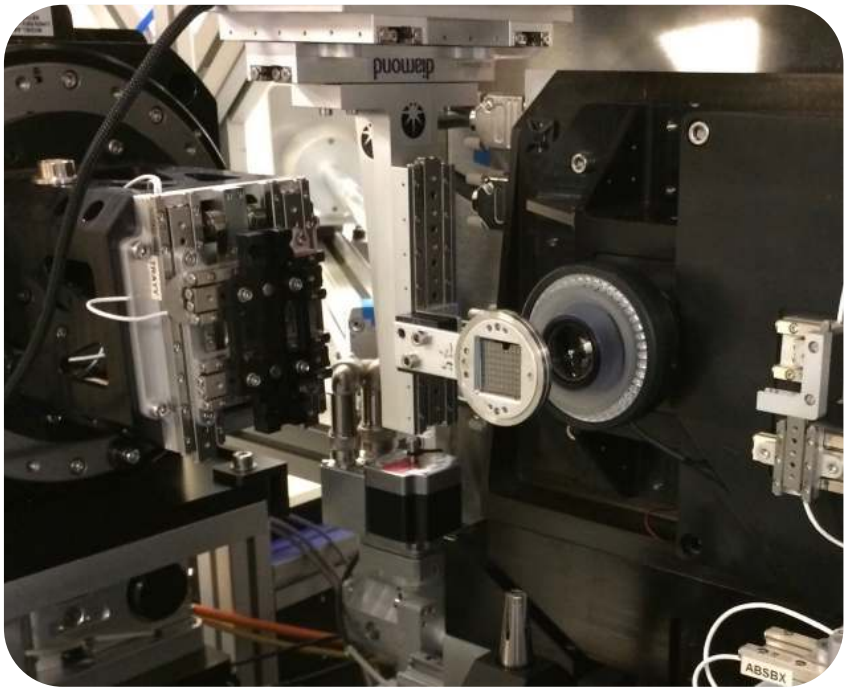



Fixed target

# Fixed target serial crystallography at I24

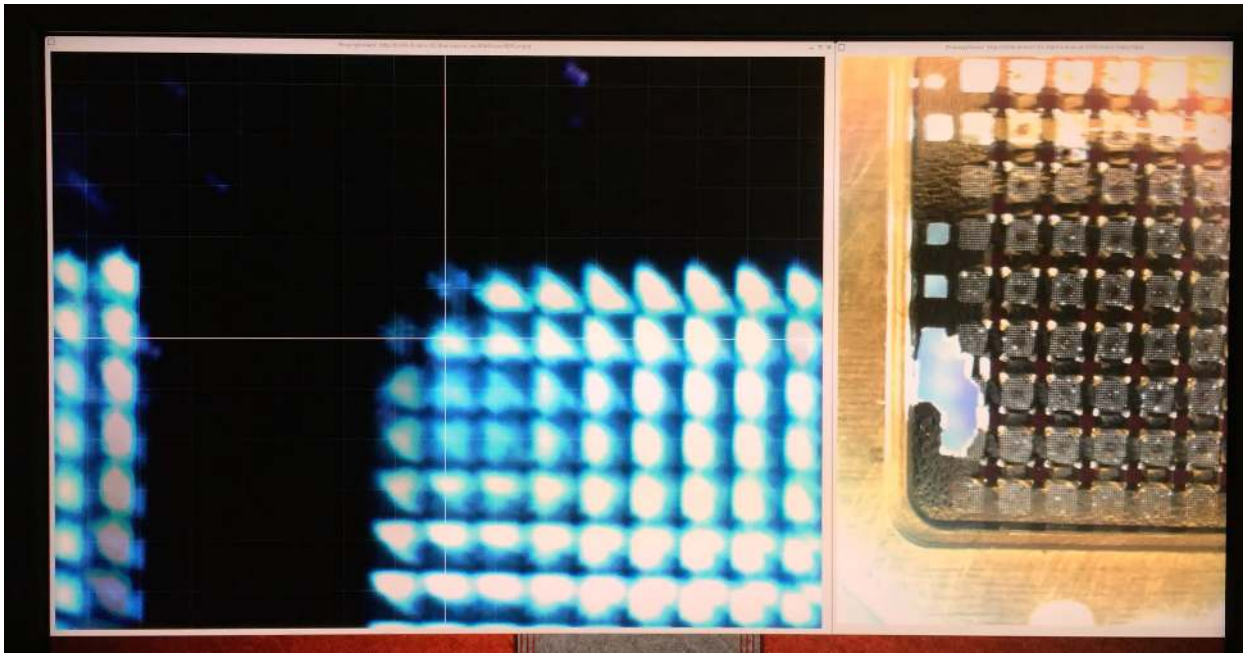



Siliicon nitride chips
25,600 positions/chip

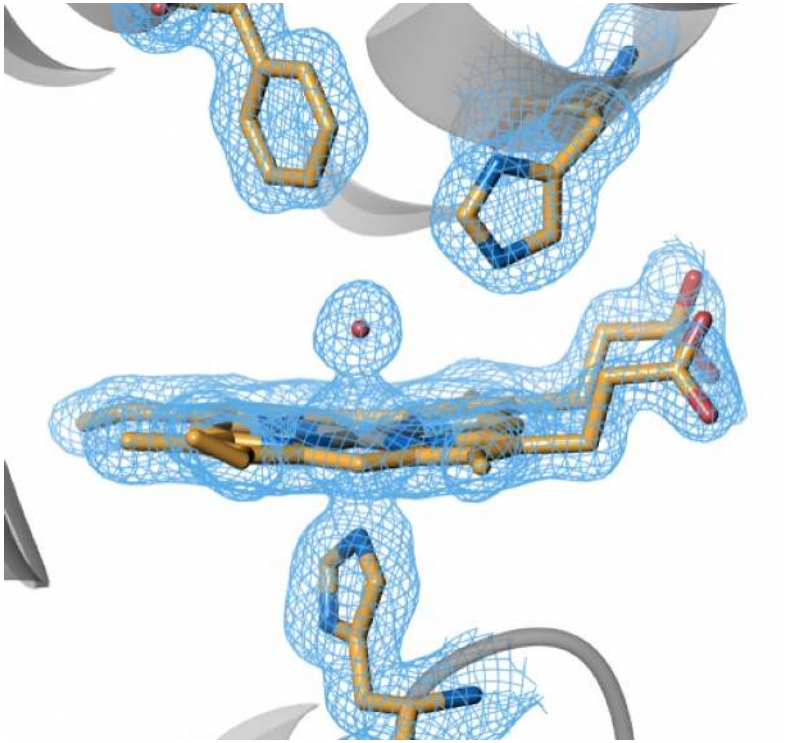



Bespoke mount in I24
Fast stages

# Fixed target serial crystallography at I24

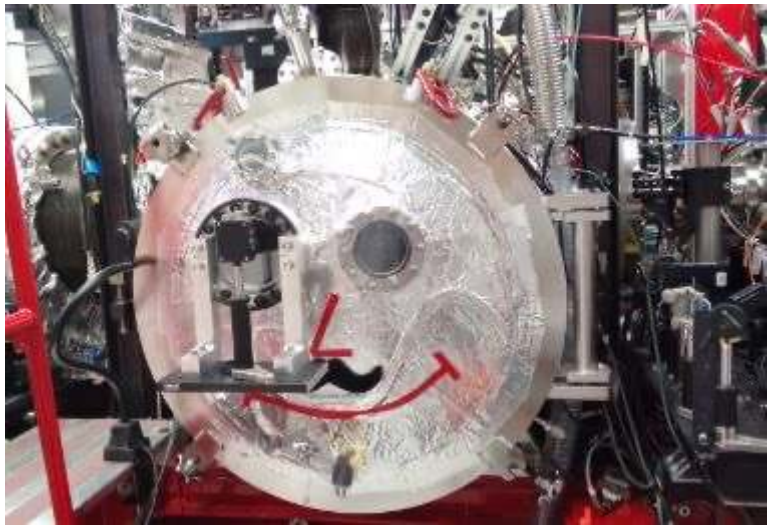


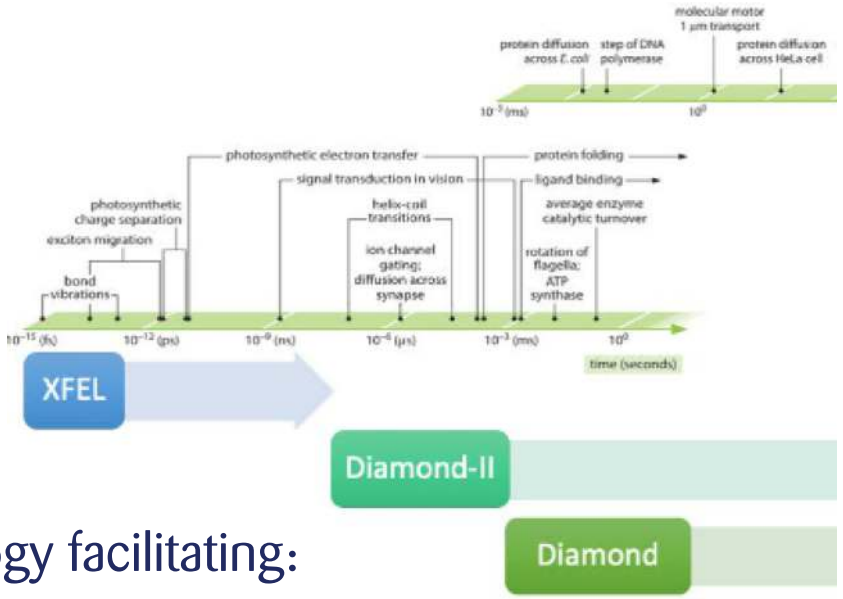
Full chip rastered in 6-20 mins
10-50 ms/frame (+ pause at each frame)
Typically ~25 chips/12 hours



1.5 Å structure
from 9000 crystals

# The UK-XFEL Hub

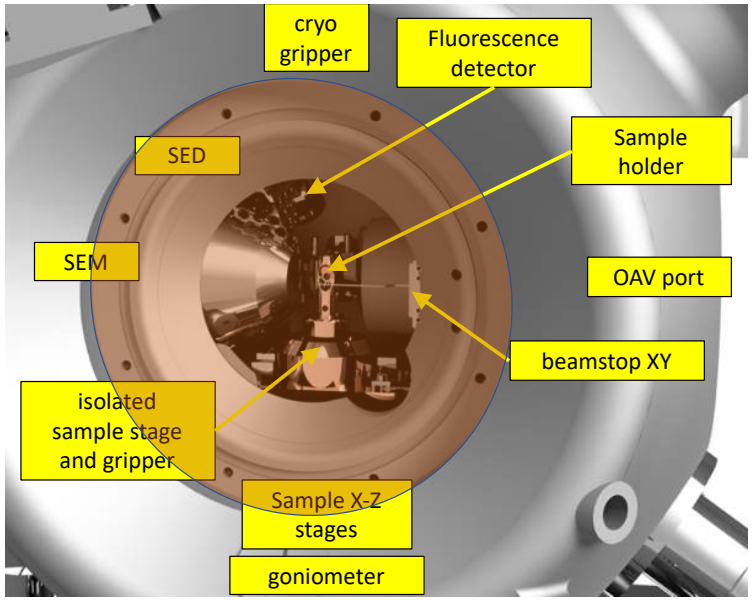




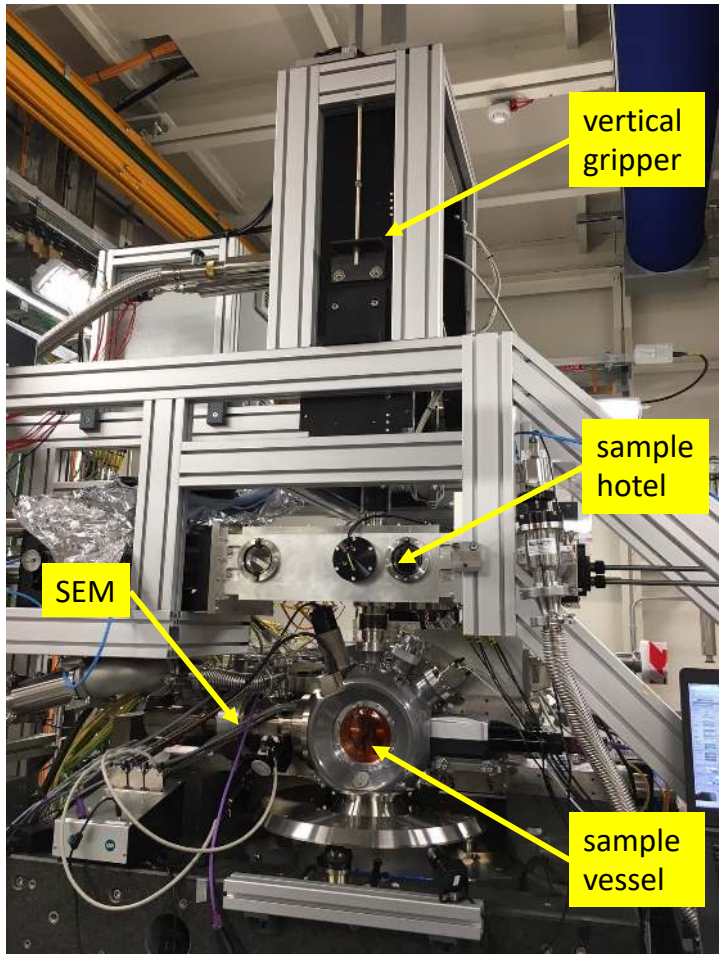

Tools and strategies to enable dynamic structural biology facilitating:

- (time-resolved) serial structural biology experiments via sample preparation, delivery, data collection, and processing
- the transfer of methods between XFEL, synchrotron, and/or cryo-EM sources
- access to, and data collection from complementary facilities (DLS, CLF, eBIC)

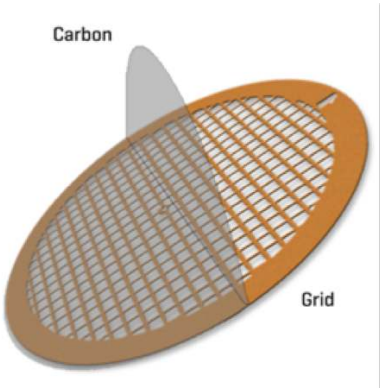
# Going even smaller: VMXm



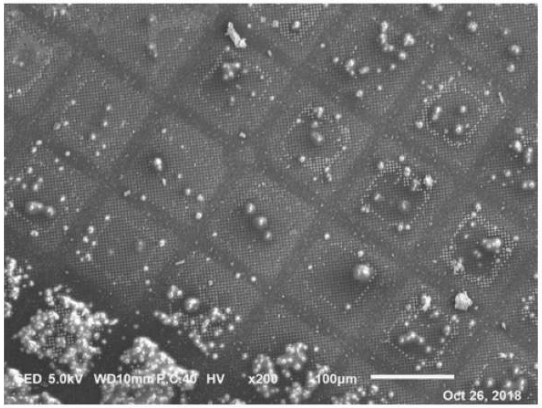

< 500 nm beamsize
Integrated SEM
In vacuum
Tunable 7-25 keV (2.0-0.57Å)



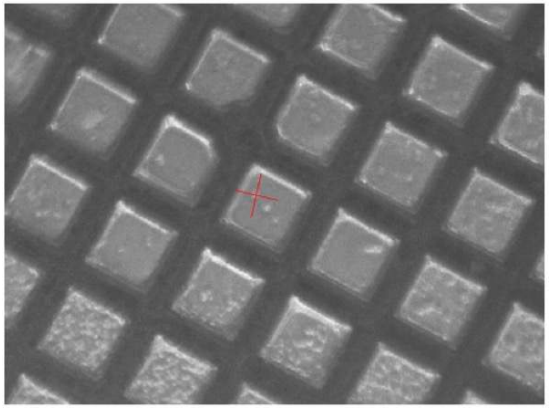
# Going even smaller: VMXm



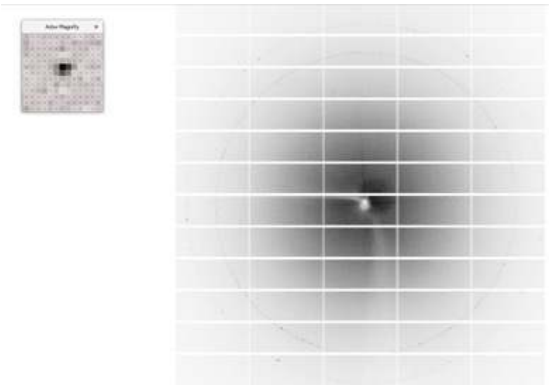

EM-grids as sample holders

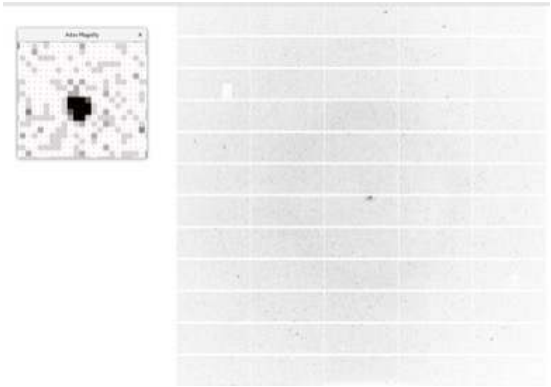



SEM



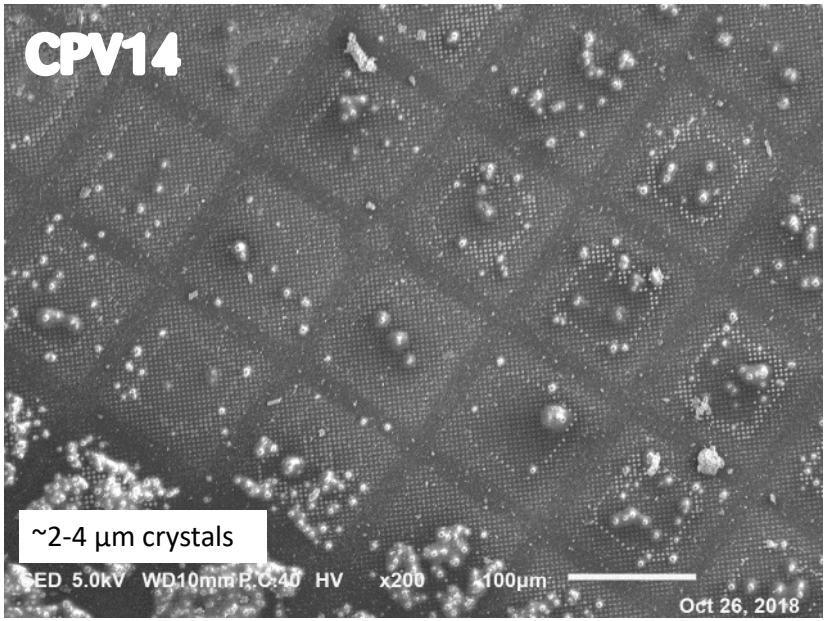
OAV

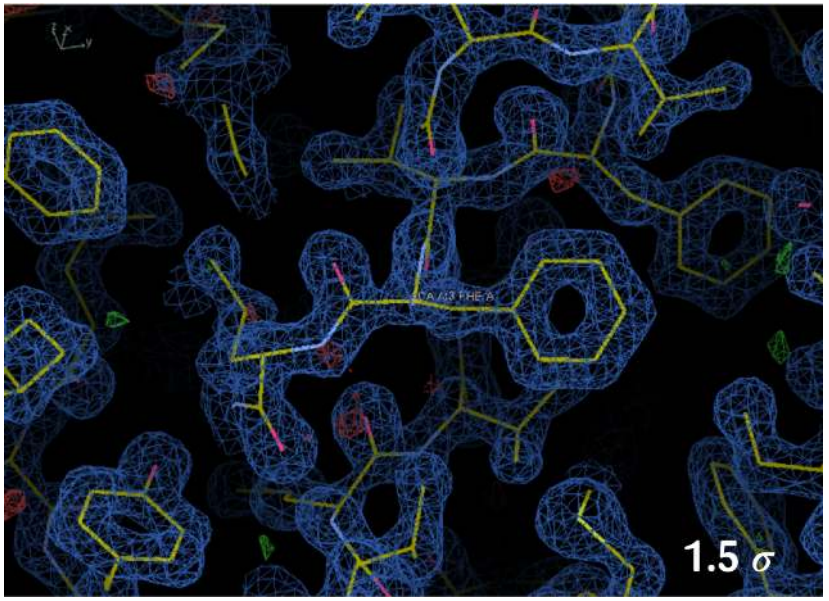




extremely low background

# VMXm: the first structures




27 crystals
1.5-2.0 deg/data per crystal

Rigid body+restrained refinement
Dials, CCP4, Refmac, Coot

Summary of merging statistics

| | Overall | Low resolution | High resolution |
|---|---|---|---|
| Resolution (Å) | 36.57 - 1.50 | 36.58 - 4.07 | 1.53 - 1.50 |
| Observations | 116543 | 7779 | 1406 |
| Unique reflections | 25790 | 1405 | 833 |
| Multiplicity | 4.5 | 5.5 | 1.7 |
| Completeness | 87.23% | 90.30% | 57.21% |
| Mean I/sigma(I) | 5.1 | 13.4 | 0.7 |
| Rmerge | 0.195 | 0.128 | 0.775 |
| Rmeas | 0.217 | 0.140 | 1.017 |
| Rpim | 0.088 | 0.053 | 0.646 |
| CC1/2 | 0.969 | 0.970 | 0.319 |

# The need for long wavelength

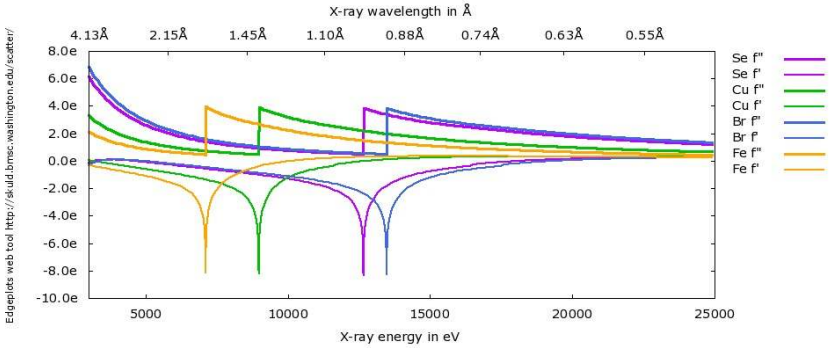
- Tuneable beamlines
- I23 dedicated
- Smart data collection strategies
- Expert systems

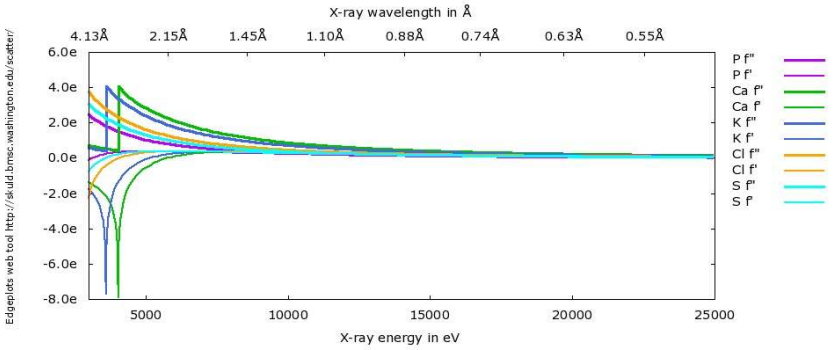
Long Wavelength

## Biologically relevant elements

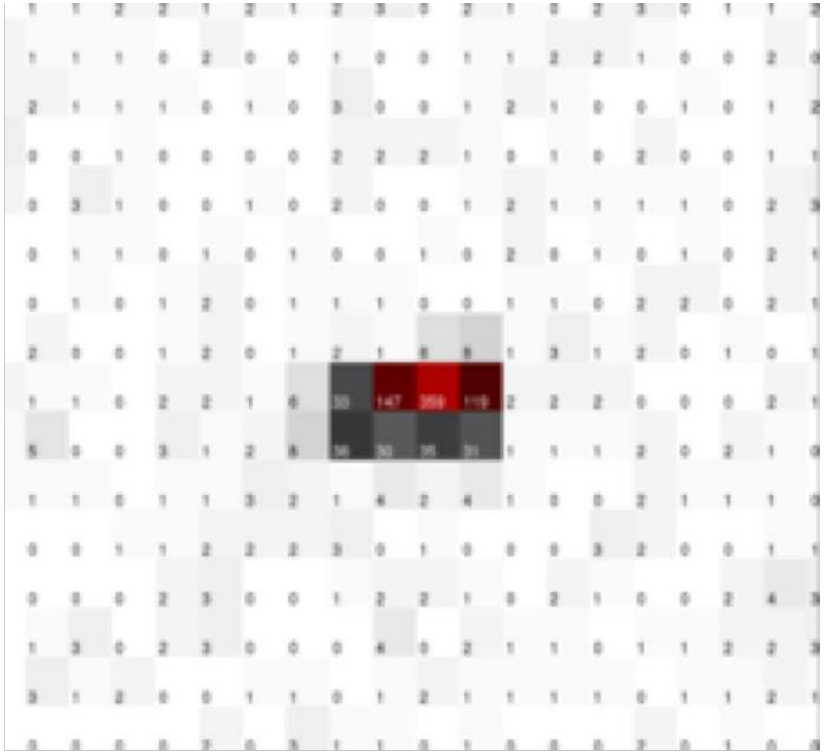
| | e (keV) | λ(Å) |
|---|---|---|
| P | 2.146 | 5.779 |
| S | 2.472 | 5.016 |
| Cl | 2.822 | 4.393 |
| K | 3.607 | 3.437 |
| Ca | 4.038 | 3.070 |

## Xray absorption edges outside the capabilities of standard beamlines



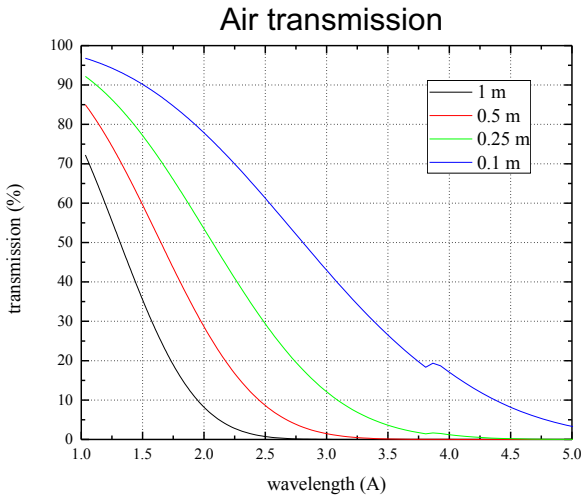
# Challenges and solutions: conceptual

Long wavelength increases the angle of diffraction
Need a curved detector to measure high resolution spots

# Challenges and solutions: technological

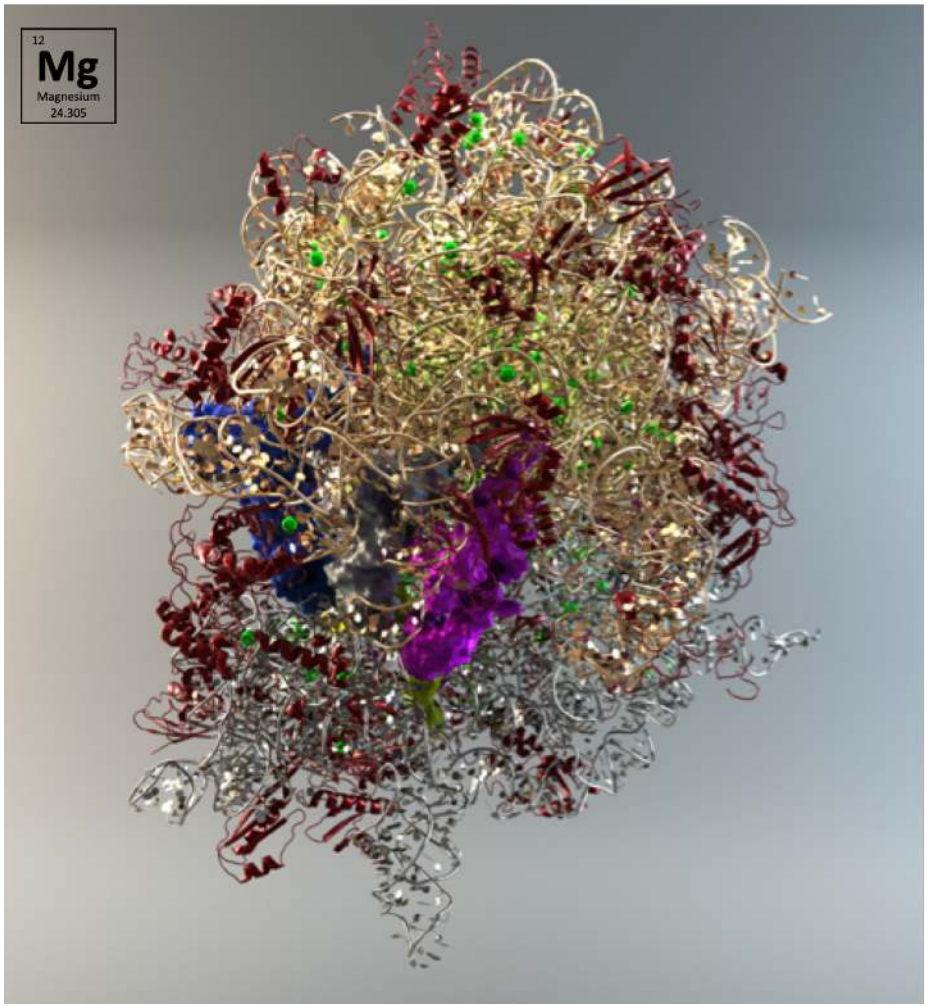
## Absorption by air of long wavelength X-rays

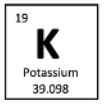



## In vacuum beamline

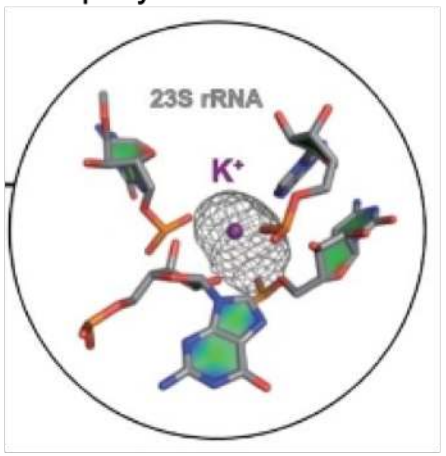

## Bespoke sample holders

# Results: ion identification

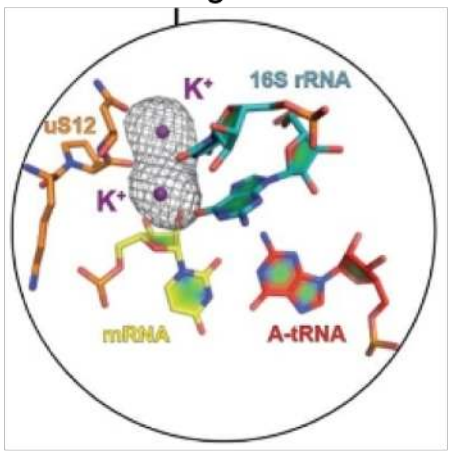


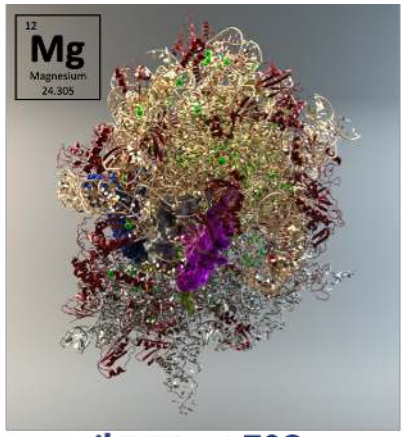

Peptidyl transferase center



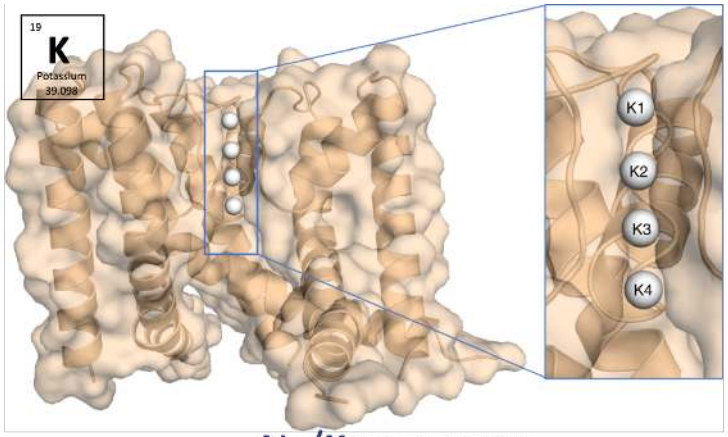

Decoding center



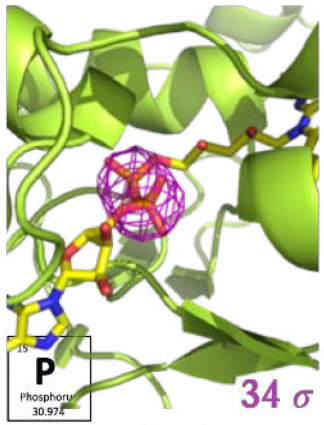
# Results: ion identification



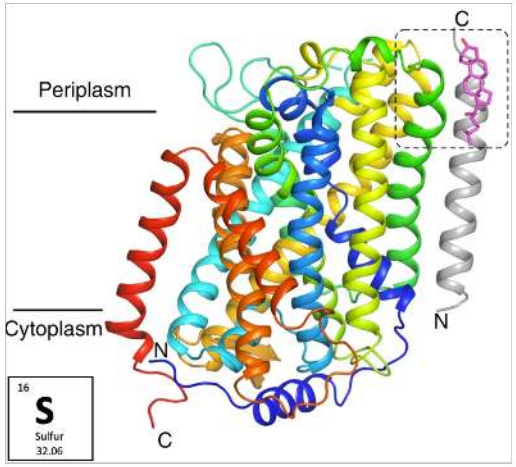

ribosome 70S



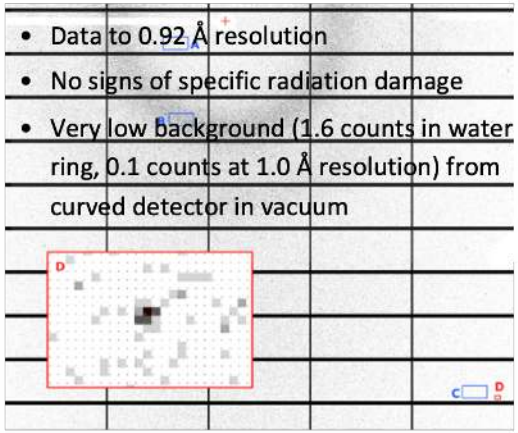

Na/K transporter



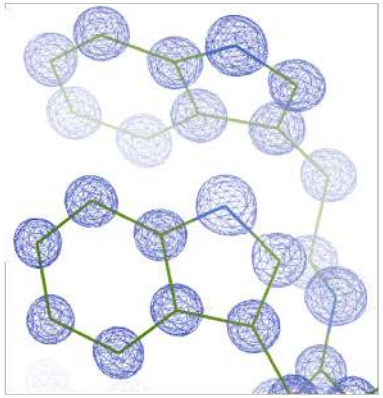

FAD-binding



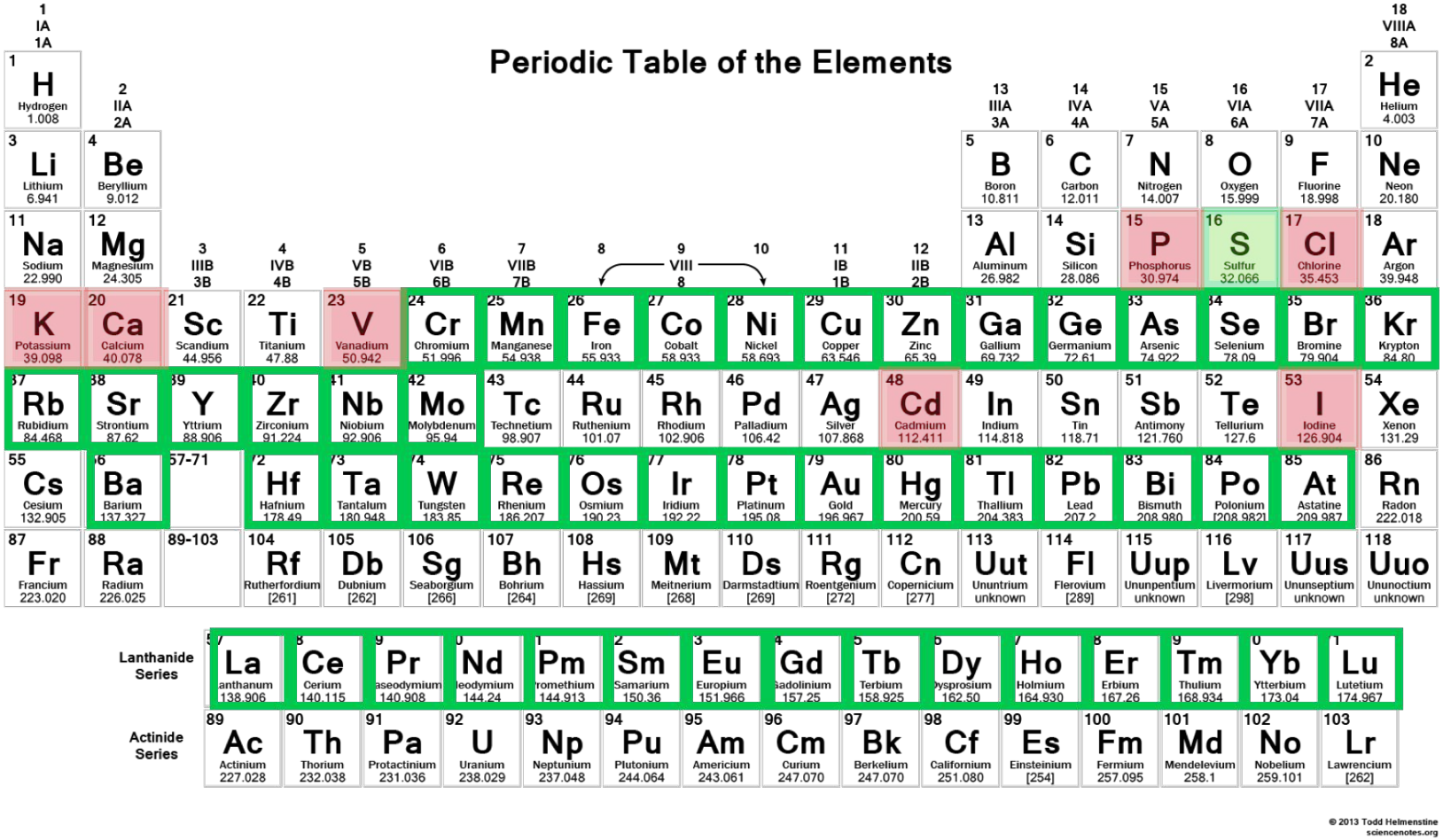

SLC7 transporter




PETase



# Reaching longer wavelengths: I23



# Increasing throughput

BART sample changer:
37 unipucks
592 samples
Exchange rate approx. 20 sec/sample

EIGER2 XE 16M
Pixel size: 75x75 $\mu m^2$
18,093,576 pixel
560 frames/sec (3600 frames in <7 sec)

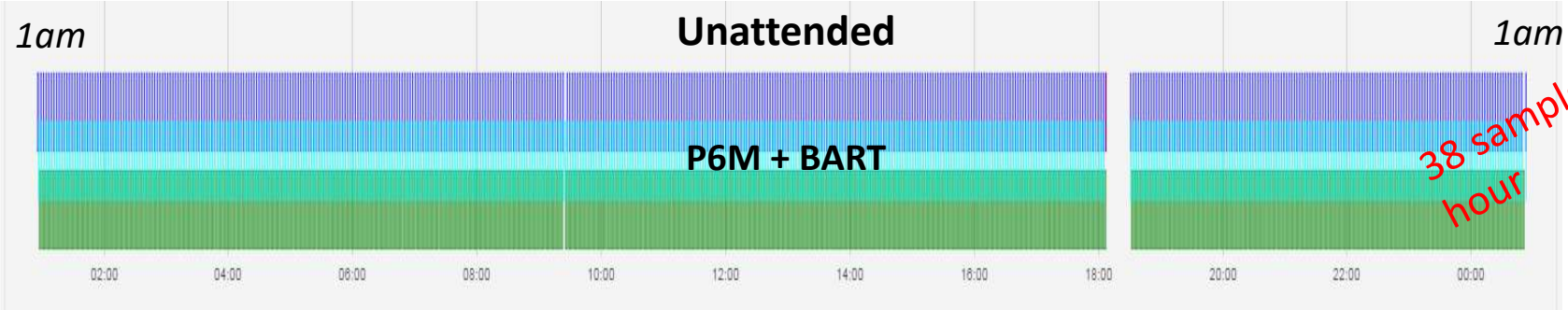
# Logistics


Weland/Randex automated store

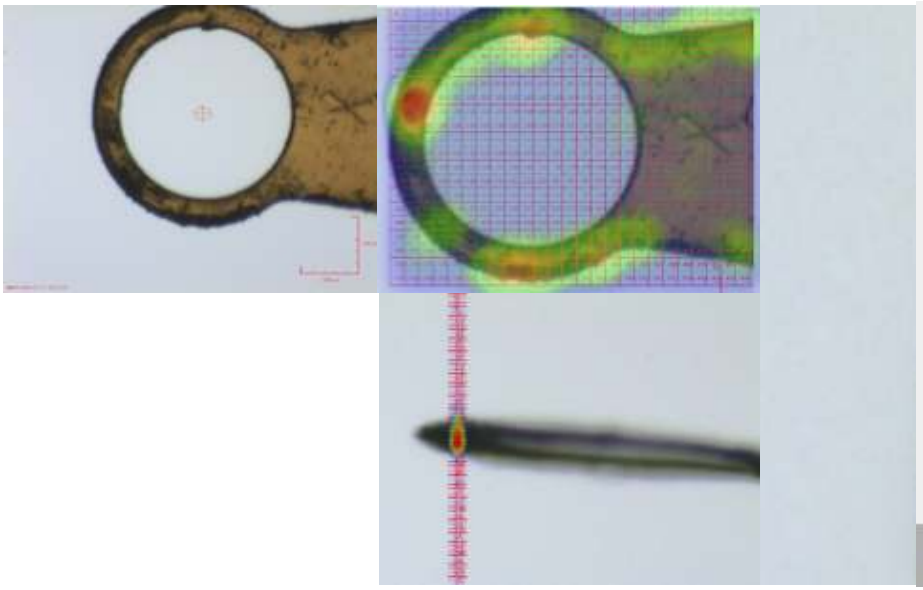
high capacity (up to 272 dewars)
barcode tracking (dewars and pucks)
logistics integrated in Synchweb/ISPyB

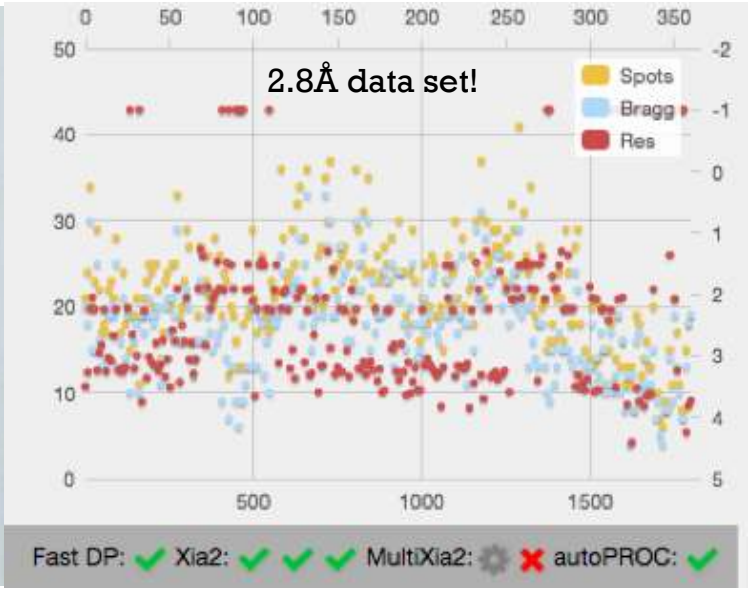

# Unattended data collection

Optimized for speed

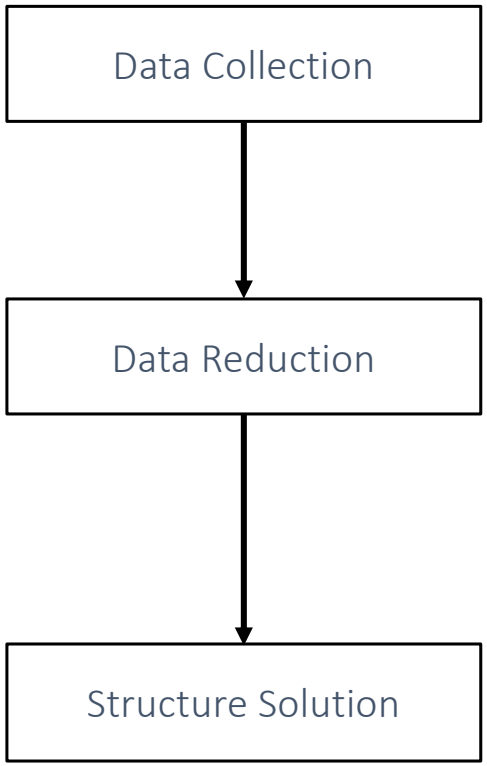


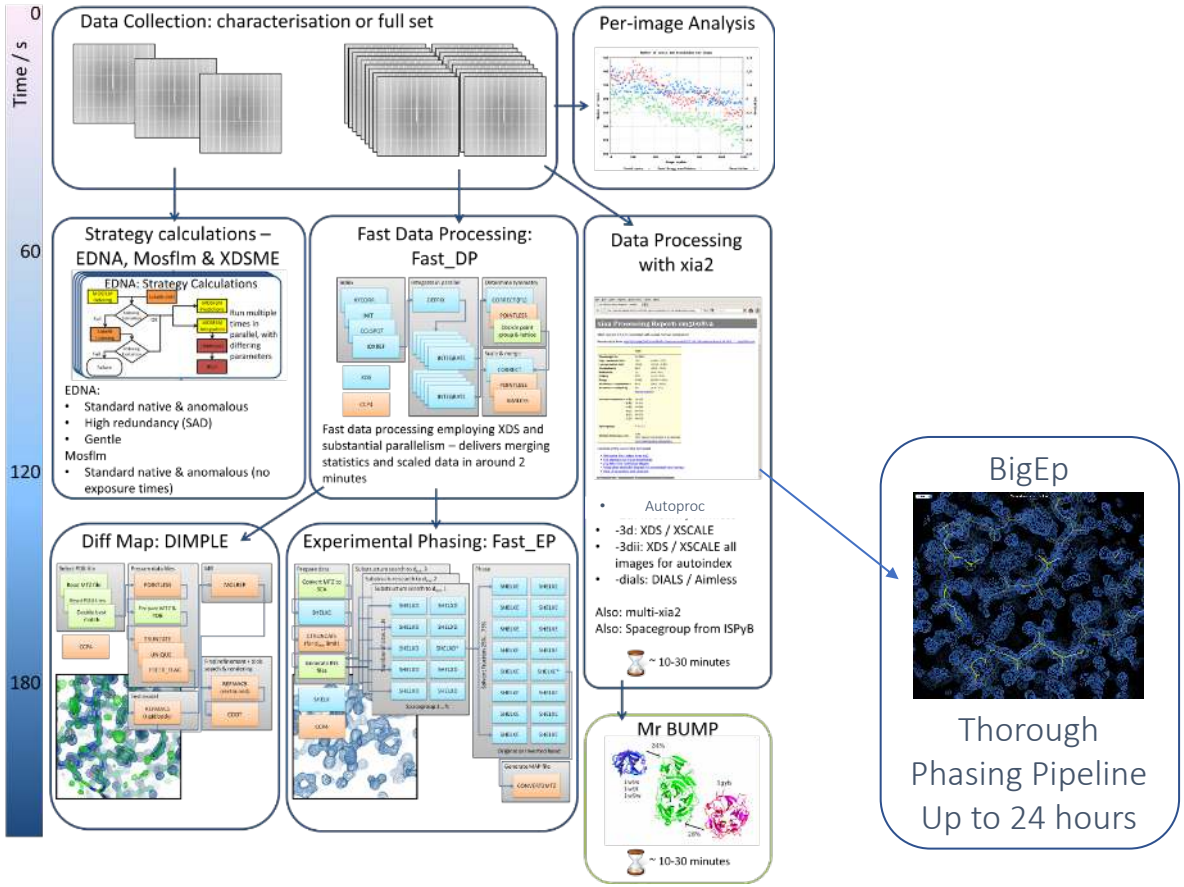

Xray-centering and strategies





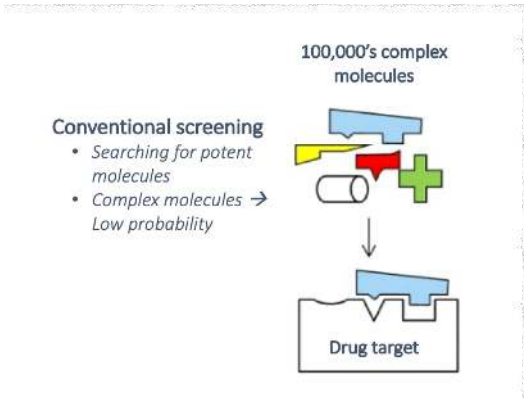
# Automated processing pipelines

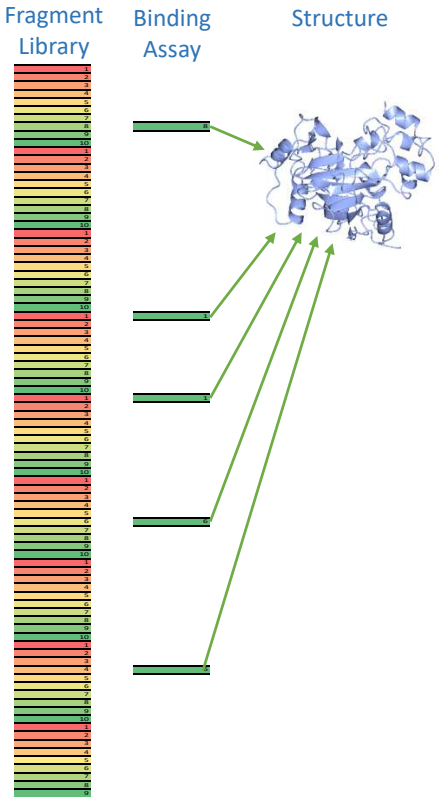


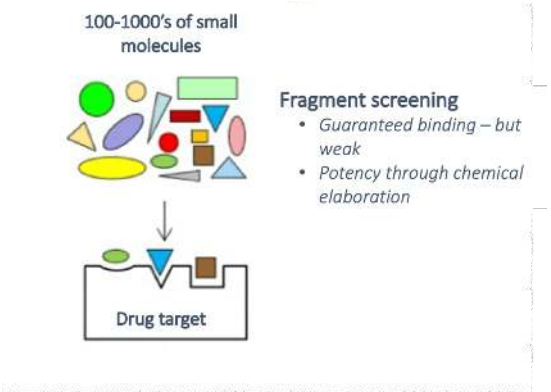
# High throughput fragment screening

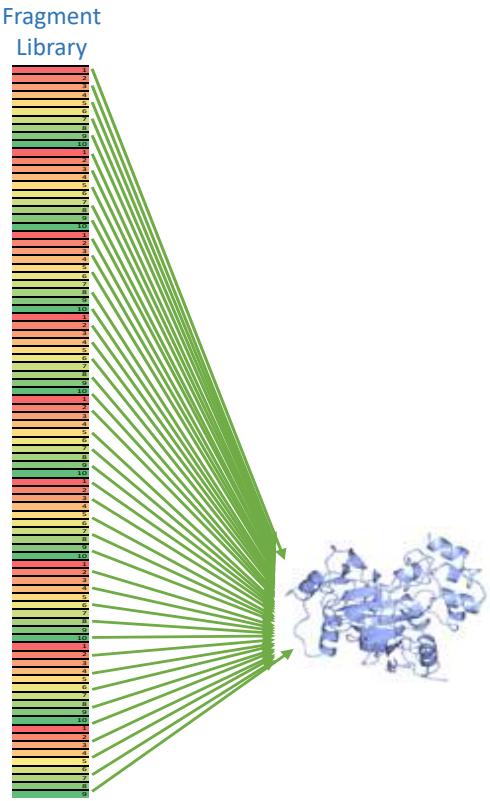
## Conventional screening
Preselection of best compounds



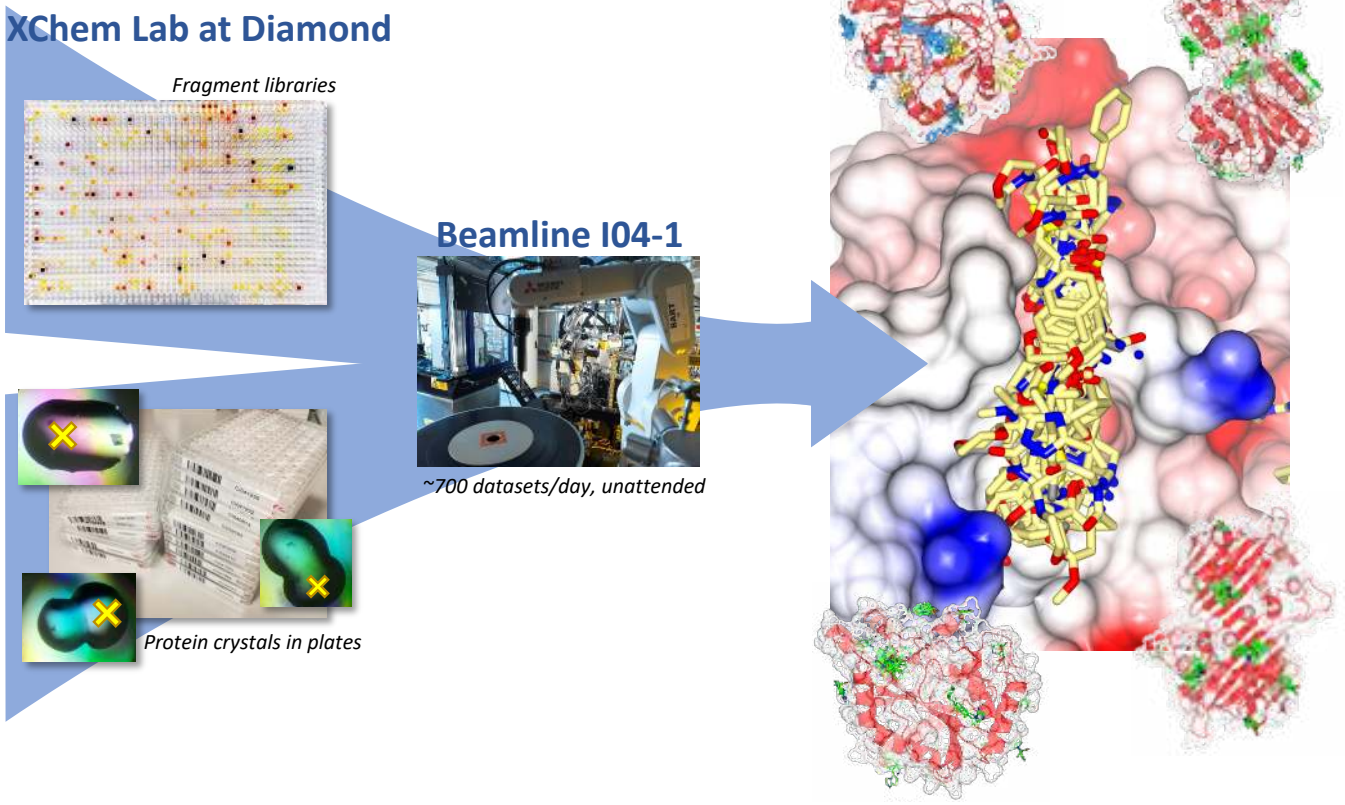




## Fragment-based approach
Screen by crystal



# A fragment-screening pipeline




1 crystal – 1 fragment

Soak/harvest/collect data
(up to 1000 crystals/week)

Routine since 2016

March 2020 COVID-19 targets

# Access to Diamond

## Non-proprietary (academic)

Peer-review access based on quality of proposal
- Free access
- Publishable experiments

Access modes:
- **BAG** (2 year programme)
- **Single application** (6 months)
- **Rapid access** (apply anytime for one-off experiments)

MX beamlines, XChem

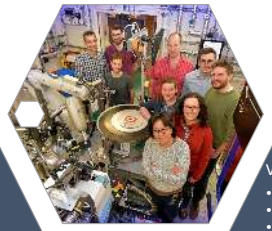
## Proprietary (industry and academia)

Organized via Industrial Liaison Office
- No need to publish
- Allocated slots throughout the schedule
- 10% of allocated beamtime

Services provided:
- Direct and remote access
- Unattended data collection
- Mail-in data collection and analysis service
- XChem fragment screening

# The team

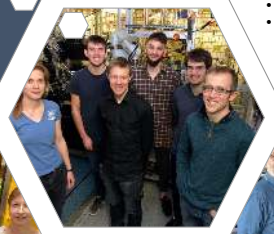
**Science group Leader**
Dave Hall



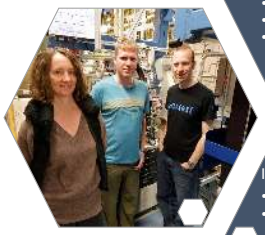
I04-1 / XChem
- **Frank von Delft**
- Jose Brandao-Neto
- Louise Dunnett
- Daren Fearon
- Charlie Tomlinson
- Warren Thompson

VMXi
- **Mike Hough**
- Juan Sanchez-Weatherby
- James Sandy
- Halina Mikolajek

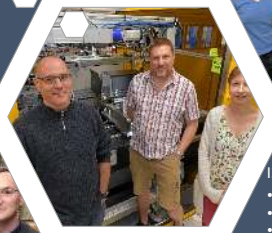
I24
- **Robin Owen**
- Danny Axford
- Selina Storm
- Sam Horrell

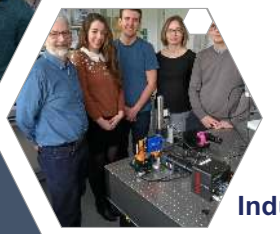


XFEL-Hub
- **Allen Orville**
- Pierre Aller
- Tiankun Zhou
- Jos Kamps
- Robert Bosman

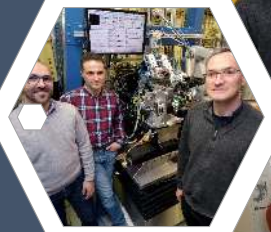


I04
- **Ralf Flaig**
- David Aragao
- Marco Mazzorana
- Pierpaolo Romano





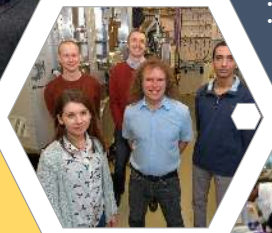
I03
- Neil Paterson
- Mark Williams



I23
- **Armin Wagner**
- Kamel el Omari
- Vitaliy Mykhaylyk
- Ramona Duman
- Christian Orr

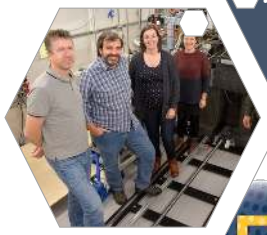
**User Support**
Felicity Bertram
Marco Mazzorana
Elliot Nelson

VMXm
- **Gwyndaf Evans**
- Jose Trincao
- Anna Warren
- Adam Crawshaw



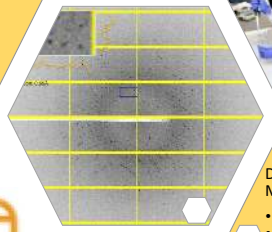
Membrane Protein Lab
- **Andrew Quigley**
- James Birch
- Harish Cheruvara

**Industrial Liaison**
Alexandre Dias
Jitka Waterman
Ailsa Powell
Christofer Bjorkelid



DIALS
- James Parkhurst
- Nicholas Devenish
- James Beilsten-Edmands
- Melanie Vollmar
- Luis Fuentes-Montero
- Graeme Winter
- David Waterman

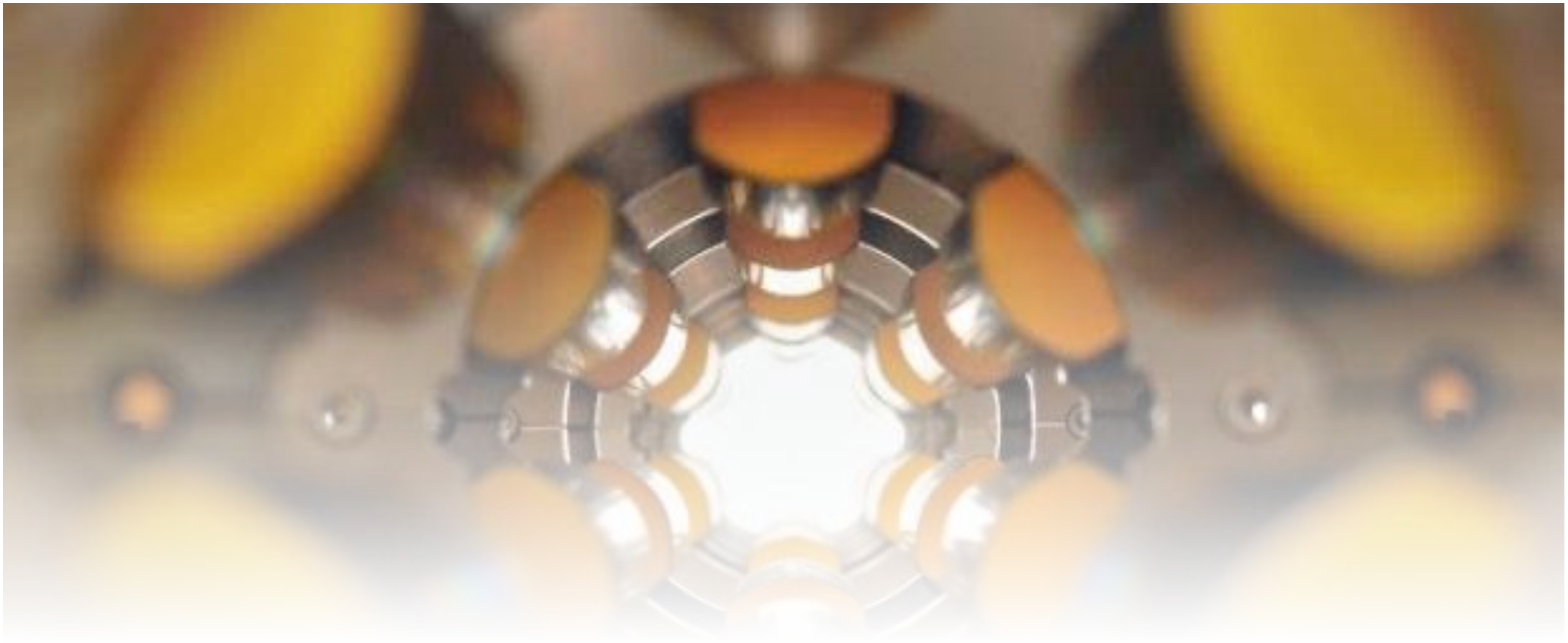
Sample preparation
- Crystallisation
- XFEL Hub
- XChem
- SSX



Data Analysis & Management
- Graeme Winter
- Richard Gildea
- Irakli Sikharulidze
- James Hall
- Karl Levik
- Stuart Fisher

**Electrical and Mechanical Technicians and Engineers**

**Beamline controls**



diamond

# bright ideas start here