




Multi-crystal approaches

Richard Gildea
DLS-CCP4 workshop
2022



Sooner or later, everything old is
new again

Stephen King, The Colorado Kid

What has been will be again,
what has been done will be done
again; there is nothing new under
the sun

Ecclesiastes 1:9

STRUCTURE OF MYOGLOBIN

A THREE-DIMENSIONAL FOURIER SYNTHESIS AT 2 Å. RESOLUTION

By DRs. J. C. KENDREW, R. E. DICKERSON, B. E. STRANDBERG, R. G. HART
and D. R. DAVIES*

Medical Research Council Unit for Molecular Biology, Cavendish Laboratory, Cambridge

AND

D. C. PHILLIPS and V. C. SHORE

Davy Faraday Laboratory, The Royal Institution, London

The data for each derivative were recorded on twenty-two precession photographs; a separate crystal had to be used for each photograph to keep radiation damage within acceptable limits. The results from the different photographs were scaled together on the computer, the best set of scaling factors being determined by solving an appropriate 22×22 matrix⁵. The degree of isomorphism of each derivative was tested, and found adequate, by means

STRUCTURE OF MYOGLOBIN

A THREE-DIMENSIONAL FOURIER SYNTHESIS AT 2 Å. RESOLUTION

By Drs. J. C. KENDREW, R. E. DICKERSON, B. E. STRANDBERG, R. G. HART
and D. R. DAVIES*

Medical Research Council Unit for Molecular Biology, Cavendish Laboratory, Cambridge

AND

D. C. PHILLIPS and V. C. SHORE

Davy Faraday Laboratory, The Royal Institution, London



simultaneously. Whereas myoglobin crystals give 400 reflexions having spacings greater than 6 Å., the number of reflexions with spacings greater than 2 Å. is 9,600, each of which has to be measured not only for the unsubstituted protein but also for each of the derivatives. The very much greater number of data posed many problems, both in recording intensities and in computation, and in this stage we relied much more heavily than before on the use of a high-speed computer; it was fortunate that about the time the work began the *Edsac* Mark I computer used previously was superseded by the very much faster and more powerful Mark II.

Structure of a human common cold virus and functional relationship to other picornaviruses

Michael G. Rossmann*, Edward Arnold*, John W. Erickson*†, Elizabeth A. Frankenger*†, James P. Griffith*, Hans-Jürgen Hecht*†, John E. Johnson*, Greg Kamer*, Ming Luo*, Anne G. Mosser†, Roland R. Rueckert†, Barbara Sherry† & Gerrit Vriend*

* Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA

† Biophysics Lab, University of Wisconsin, 1525 Linden Drive, Madison, Wisconsin 53706, USA

The data used for the results given here were collected at the Cornell High Energy Synchrotron Source (CHESS). A new crystal was used for every exposure. A total of 83, 0.3°-oscillation film packs were eventually included in the native data (Table 1). Surveys for suitable isomorphous heavy-atom derivatives eventually produced two related compounds, 1 mM KAu(CN)₂ and 5 mM KAu(CN)₂. Full high-resolution data were collected of the former, but only partial low resolution (0.6 to 0.8° oscillation angles) of the latter.

Table 1 Data used in structure determination

Resolution range (Å)	No. of unique observations ($F^2 > 1\sigma(F^2)$) for each data set and percentage of possible total					
			1 mM		5 mM	
	Native No.	%	KAu(CN) ₂ No.	%	KAu(CN) ₂ No.	%
∞–30	358	63	331	57	231	40
30–15	3,445	86	2,847	70	1,850	46
15–10	9,109	84	7,757	70	5,029	46
10–7.5	17,532	83	14,641	68	9,486	45
7.5–5.0	70,043	81	57,577	65	34,046	39
5.0–3.5	184,864	78	138,477	58		
3.5–3.0	133,101	63	83,852	39		
3.0–2.75	60,629	36	25,271	15		
2.75–2.6	26,834	11	8,175	3		
Total	509,915	58	338,928	39	50,641	41
No. of film packs	83		48		11	
R-factor (%)	11.0		12.9		12.0	

The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution

Ravindra Acharya*, Elizabeth Fry*, David Stuart^{*‡}, Graham Fox[†], David Rowlands[†] & Fred Brown[†]

* Laboratory of Molecular Biophysics, Oxford, OX1 3QU, UK

† Department of Virology, Wellcome Biotech, Beckenham, Kent, BR3 3BS, UK

Structure determination

We previously described¹⁶ the growth, purification, crystallization and preliminary X-ray diffraction analysis of a virus of serotype 0₁ strain BFS 1860. The crystals belong to space group *I*23, $a = 345$ Å, with one-twelfth of the virus particle in the crystallographic asymmetric unit (5-fold non-crystallographic redundancy). The small size of the crystals (average dimensions for those used here were $0.12 \times 0.12 \times 0.06$ mm³), coupled with the large number of observable reflections, led to very weak diffraction. In common with many other picornavirus crystals they were also extremely radiation-sensitive, enabling only a single small angle (0.4° – 0.5°) oscillation photograph to be recorded from each crystal, using the SERC Synchrotron Radiation Source (Daresbury, UK). Data collection was made feasible by the characteristics of the radiation: (i) the extreme parallelism of the beam; (ii) its great intensity; and (iii) the relatively short wavelength (close to 0.9 Å), all of which contributed to an increase in the yield of data from each crystal. Many crystals were examined using the so-called American method¹⁷ resulting in 135 usable filmpacks of which 106 were analysed (Table 1).

Table 1 Data used in structure determination

Resolution range (Å)	No. unique observed reflections	Percentage of possible observations
10,000–25	43	18.86
25–15	888	92.50
15–10	2,800	95.96
10–7.5	4,758	95.79
7.5–5.0	20,274	96.05
5.0–3.5	51,750	95.39
3.5–3.0	43,101	85.45
3.0–2.9	5,488	39.81
2.9–2.8	1,836	10.73
Total (to 2.9)	129,100	86.98

R-factor (%), 13.9 (all data included). The synchrotron parameters were: Wiggler field strength, 5 T; current, 150–280 mA; energy, 2 GeV; wavelength, 0.85–1.0 Å. Crystal to film distance, 170–190 mm; collimator diameter, 300 µm; oscillation range 0.4–0.5°. Number of crystals examined, 500; number of usable filmpacks (1 filmpack per crystal), 135; number processed, 106.

Methods Used in the Structure Determination of Foot-and-Mouth Disease Virus

BY ELIZABETH FRY, RAVINDRA ACHARYA* AND DAVID STUART

Laboratory of Molecular Biophysics, Rex Richards Building, South Parks Road, Oxford OX1 3QU, England

(Received 14 March 1992; accepted 13 July 1992)

Data were recorded on small-angle (0.4°) rotation photographs (one per crystal) using the Enraf-Nonius oscillation camera. The crystals diffracted to at least 2.9 \AA for a typical exposure time of 15 min. Most data were recorded at the fixed temperature of 288 K. To test the effect of lower temperature, a crystal was cooled to 277 K; the diffraction pattern clearly indicated disordering of the crystal lattice, which was, however, reversed by warming the crystal to 288 K. This project was one of the first to employ exclusively short wavelengths $< 1 \text{ \AA}$. Comparative tests involving the collection of data using $\lambda = 1.488 \text{ \AA}$ (station 7.2) and approximately $\lambda = 0.90 \text{ \AA}$ (station 9.6) showed considerable improvement not only in the signal-to-noise ratio for comparable crystals but also in the crystal lifetime.

Once mounted in 0.7 mm quartz capillaries (disease security regulations required the use of quartz), most crystals preferentially aligned such that a $[\bar{1}0\bar{1}]$ zone axis was perpendicular to the capillary wall. With this axis along the X-ray beam, there was no guarantee that any given axis was aligned with the rotation axis. Optical alignment was used to prealign the crystals to some extent but their fragility prevented reorientation in the quartz tube. The neglect of setting operations is common with virus crystals prone to radiation damage (Rossmann *et al.*, 1985; Fry, Logan & Stuart, 1992).

Methods Used in the Structure Determination of Foot-and-Mouth Disease Virus

BY ELIZABETH FRY, RAVINDRA ACHARYA* AND DAVID STUART

Laboratory of Molecular Biophysics, Rex Richards Building, South Parks Road, Oxford OX1 3QU, England

(Received 14 March 1992; accepted 13 July 1992)

Data used in the structure determination were initially processed to 4.5 Å resolution to enable refinement of the crystal-orientation and rocking-curve parameters before going to higher resolution and to provide a data set to resolve the ambiguity of the orientation of the particle in the cubic unit cell. This ambiguity arises because icosahedral particles do not possess fourfold symmetry. Thus, there are two different ways of placing an icosahedron on a specific threefold axis, which are related by a 90° rotation about a particle twofold axis. If Friedel's law holds, we may bring all data to a common orientation by exchanging the *h* and *k* indices as required. A subset of film packs was accumulated showing good agreement without any indices swapped. Subsequent packs were tested for agreement against this reference set using a convolution-search method over a mis-indexing grid, calculating correlation coefficients between the intensities of reflections from the reference set of data and those from the new pack, with both possible choices of indices.

The atomic structure of the bluetongue virus core

Jonathan M. Grimes^{*†}, J. Nicholas Burroughs[‡], Patrice Gouet^{*†}, Jonathan M. Diprose^{*}, Robyn Malby^{*}, Stephan Zióntara[§], Peter P. C. Mertens[‡] & David I. Stuart^{*||}

Data collection. The growth, purification and crystallization of cores of BTV 1(SA) were done as described previously^{5,13}. Data collection (Table 1) and structure determination will be described in greater detail elsewhere. The crystals (usually between 0.3 and 0.4 mm in the longest dimension) were pre-mounted in quartz capillary tubes at IAH, Pirbright. Data were collected on a Marresearch imaging plate (diameter 30 cm), to low resolution at PX station 9.6, SRS, Daresbury¹³, and subsequently to higher resolution at the undulator beamline ID2, ESRF, Grenoble. Data to a nominal resolution of 3.5 Å Bragg spacings were collected on ID2, by oscillating the crystal through 0.25° at ~8 °C. Only one exposure was possible per sample position, owing to radiation damage, and consequently well over 1,000 crystals were examined to obtain these data. The beam cross-section was no more than 0.1 × 0.1 mm² for ESRF data collection. The crystals (space group $P2_12_12$, unit cell dimensions $a = 795.6$ Å, $b = 821.8$ Å and $c = 753.3$ Å), contain half a particle in the asymmetric unit (30-fold non-crystallographic symmetry, n.c.s.).

resolution data before the crystal was destroyed. We have investigated methods for cooling the crystals; however, the usual protocols proved ineffective. We therefore worked with crystals cooled to approximately 277 K using an FTS device.

J. Synchrotron Rad. (1999). 6, 865–874

Bluetongue virus: the role of synchrotron radiation

Jonathan M. Diprose,^{a*} Jonathan M. Grimes,^a Patrice Gouet,^{a†} Robyn Malby,^{a‡} J. Nicholas Burroughs,^b Julien Lescar,^c Bjarne Rassmussen,^d Peter P. C. Mertens^b and David I. Stuart^{a,e}

Diffraction was first demonstrated on PX9.6 at the SRS, a second-generation source upgraded some years ago by the installation of a high-brightness lattice, although the need for higher fluxes for this project has relegated this station largely to a testing role. Data collection on DW21B and DW32 at LURE (Orsay, France), an unreconstructed source with lower flux density, demonstrated the crucial importance of brilliance and provided no advantage over the (much closer) SRS. The high-brilliance beam available on ID2B from the third-generation ESRF allows an image to be recorded in a few seconds, so that over 600 crystals could be examined in a two-day experiment. Other stations, in particular those belonging to the ID14 family at the ESRF, may provide a comparable service in the future.

using this setting (Fig. 4). Despite crystals that often only lasted for a single exposure, we could maintain a collection rate of 2–3 min image⁻¹ by pre-mounting crystals on goniometer heads and carefully pre-centring using an optical goniometer (we endeavoured to use random crystal orientations so as to optimize data completeness). Scan and

Grimes, J., Burroughs, J., Gouet, P. et al. (1998) *Nature* 395, 470–478

Diprose, J. M., Grimes, J. M., Gouet, P., et al. (1999) *J. Synchrotron Rad.* 6, 865–874



Macromolecular Cryocrystallography

ELSPETH F. GARMAN^a AND THOMAS R. SCHNEIDER^{bc}

^aLaboratory of Molecular Biophysics, University of Oxford, Oxford OX1 3QU, England, ^bEuropean Molecular Biology Laboratory (EMBL), c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany, and ^cMax-Planck-Institute for Molecular Physiology, Rheinlanddamm 201, 44139 Dortmund, Germany. E-mail: elspeth@biop.ox.ac.uk

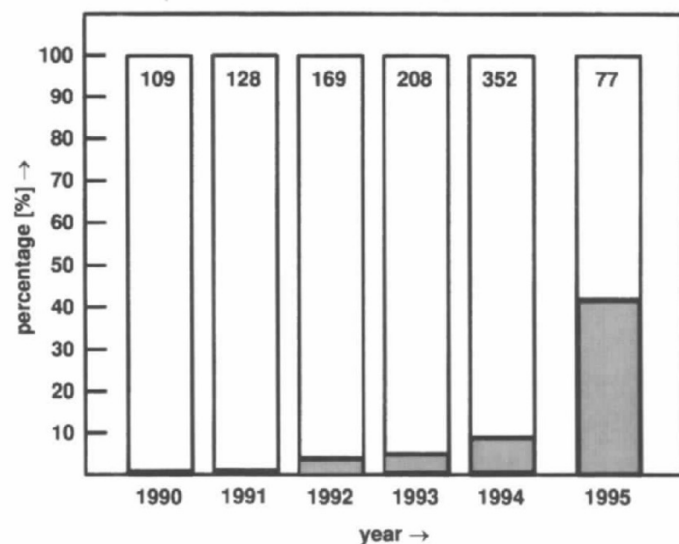
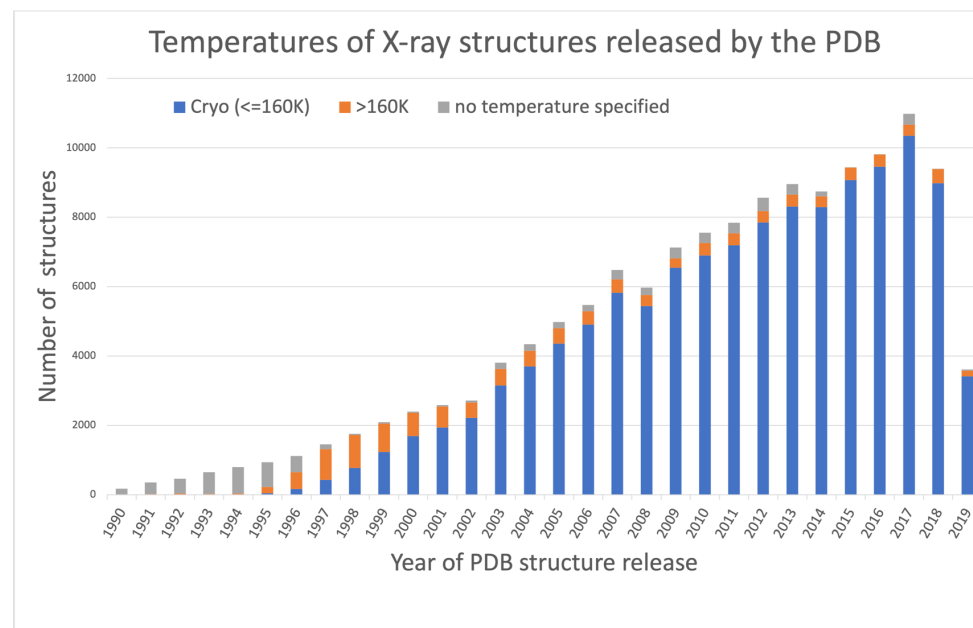


Fig. 1. The percentage of crystal structures determined with cryogenic methods versus year. For the years 1990 to 1994 these data were extracted from *Macromolecular Structures* (Hendrickson & Wüthrich, 1991, 1992, 1993, 1994, 1995). For the year 1995, all articles on macromolecular crystal structures as published in the journals *Cell*, *Nature* and *Science* were analysed. In cases where the temperature during the experiment was not reported, room-temperature was assumed. The total numbers for each year are shown at the top of each bar.

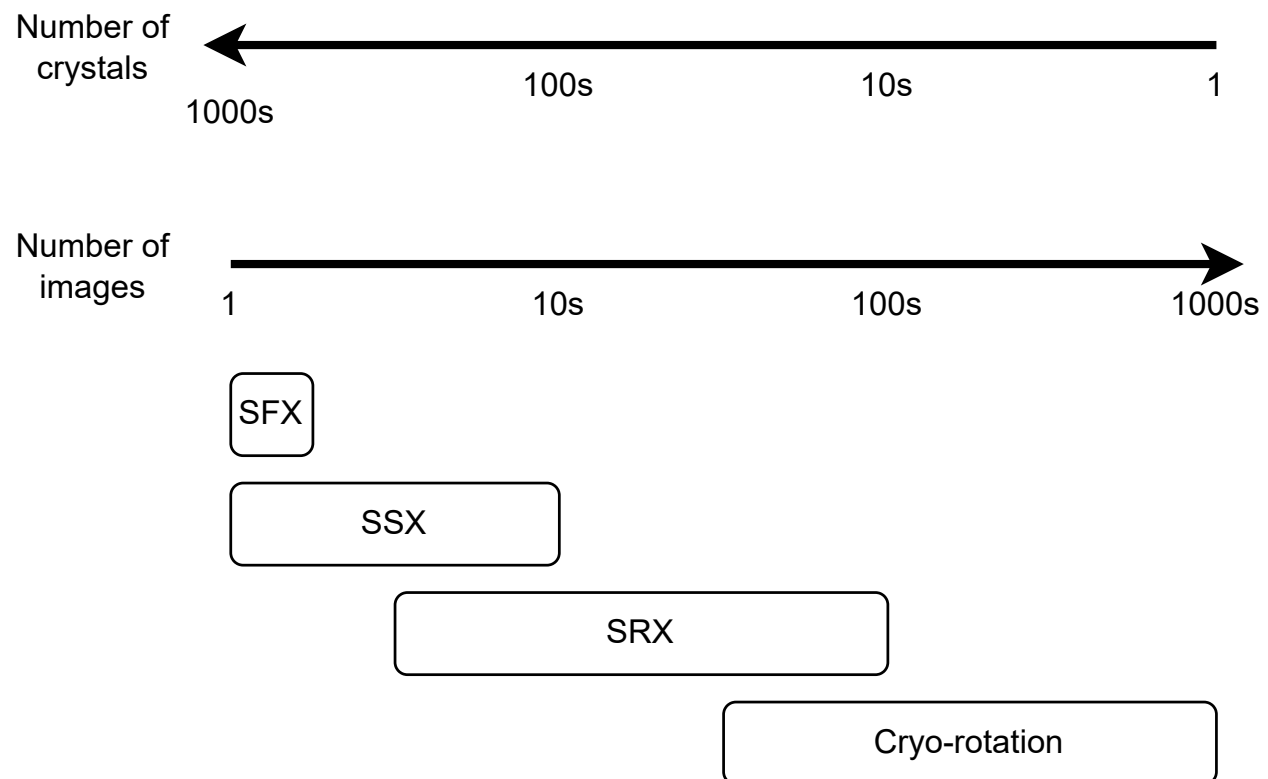
For data collected at low temperatures, the dramatic reduction in radiation damage allows complete data sets to be collected from a single crystal in many cases.





Why multi-crystal?

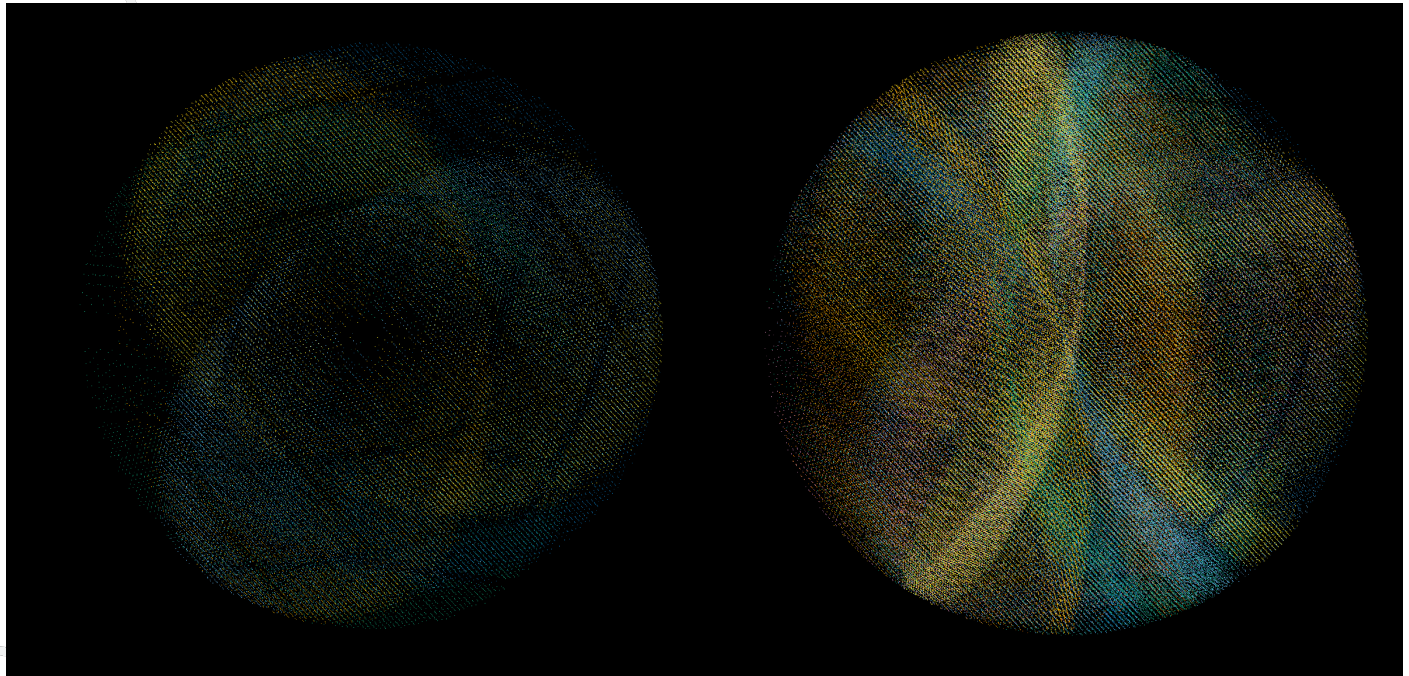
Multi-crystal approaches



Data completeness

Only possible to collect incomplete data from a single crystal due to radiation damage: obtain a complete dataset by combining data from multiple crystals

- small crystals
- room-temperature data collection



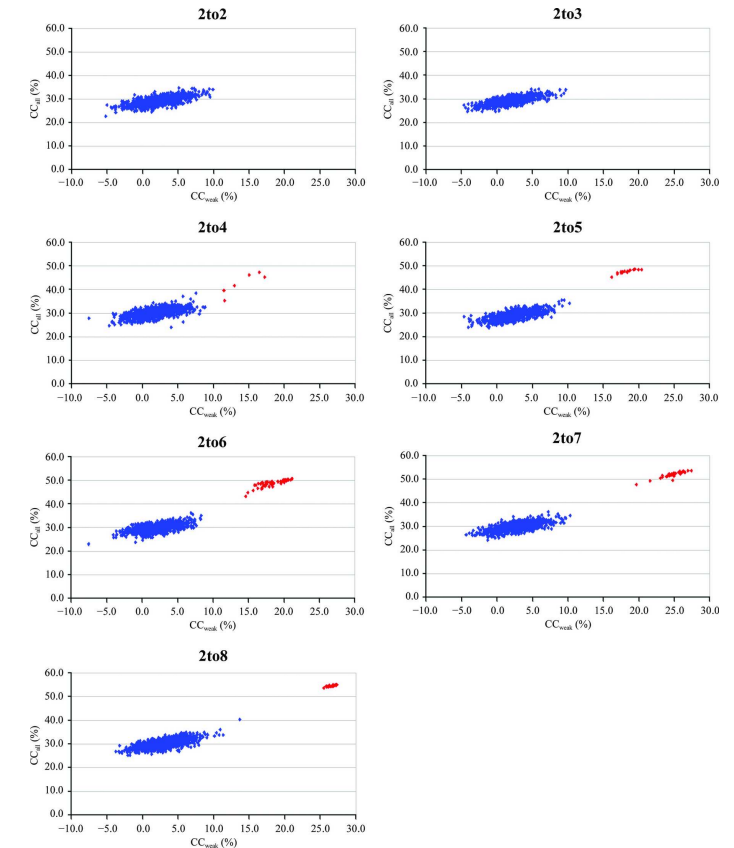
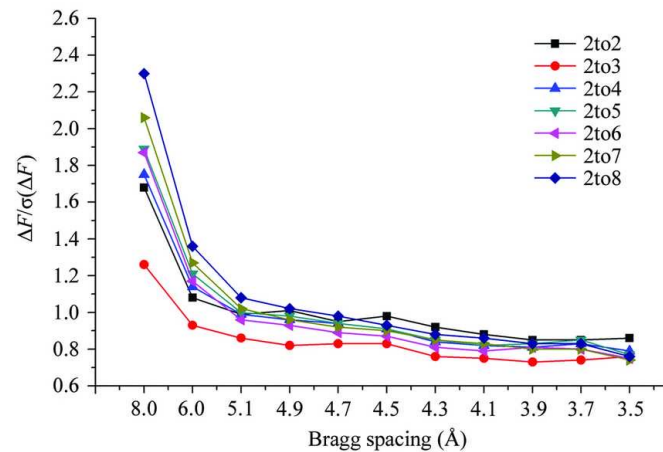
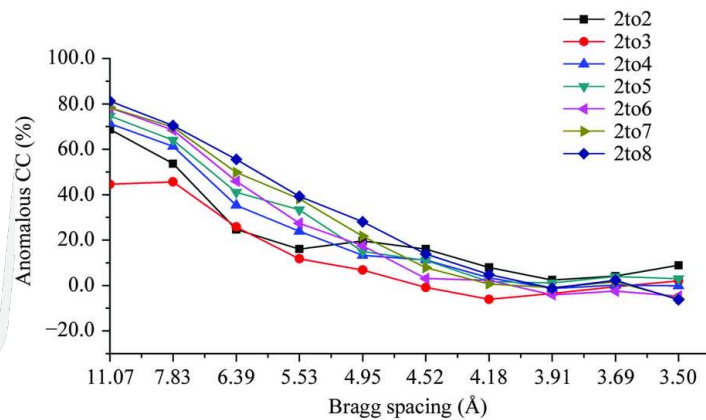
Increased signal-to-noise for weak data

- Given only random errors, the standard error σ in a measurement is reduced by \sqrt{n} if the measurement is repeated n times
- Beware R_{merge} and R_{meas} !
- Use $\text{CC}_{1/2}$ and $\langle I/\sigma \rangle_{\text{mrgd}}$

Dataset	<i>Big</i>	<i>Tiny</i>	<i>T100</i>	<i>Big+T100</i>	<i>Big2</i>
Multiplicity	2	2	200	202	4
$\langle I/\sigma \rangle_{\text{ind}}$	2.0	0.2	0.2	0.22	2.0
R_{merge}	28%	280%	399%	395%	35%
R_{meas}	40%	400%	400%	396%	40%
R_{pim}	28%	280%	28%	28%	20%
$\langle I/\sigma \rangle_{\text{mrgd}}$	2.8	0.28	2.8	4.0	4.0
$\text{CC}_{1/2}$	0.66	0.04	0.66	0.80	0.80

Increased signal-to-noise for weak data

- Improve weak anomalous signals for phasing



Room temperature

- Cryo-cooling may hide biologically significant structural features
- Virus crystals often suffer on cryo-cooling
- *In situ* data collection
- Radiation damage occurs at an absorbed dose one to two orders of magnitude lower than at cryogenic temperatures



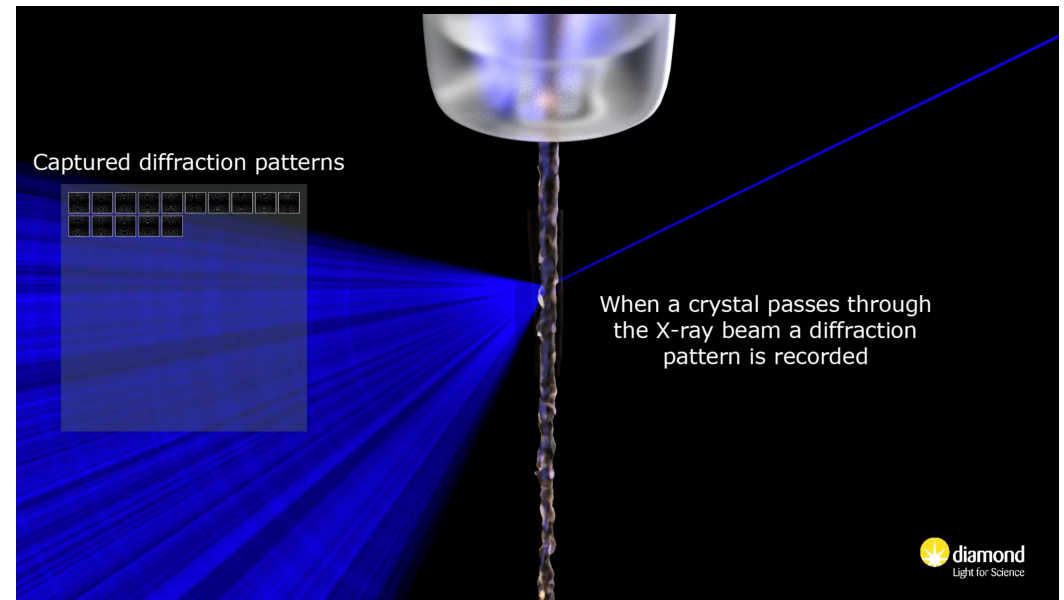
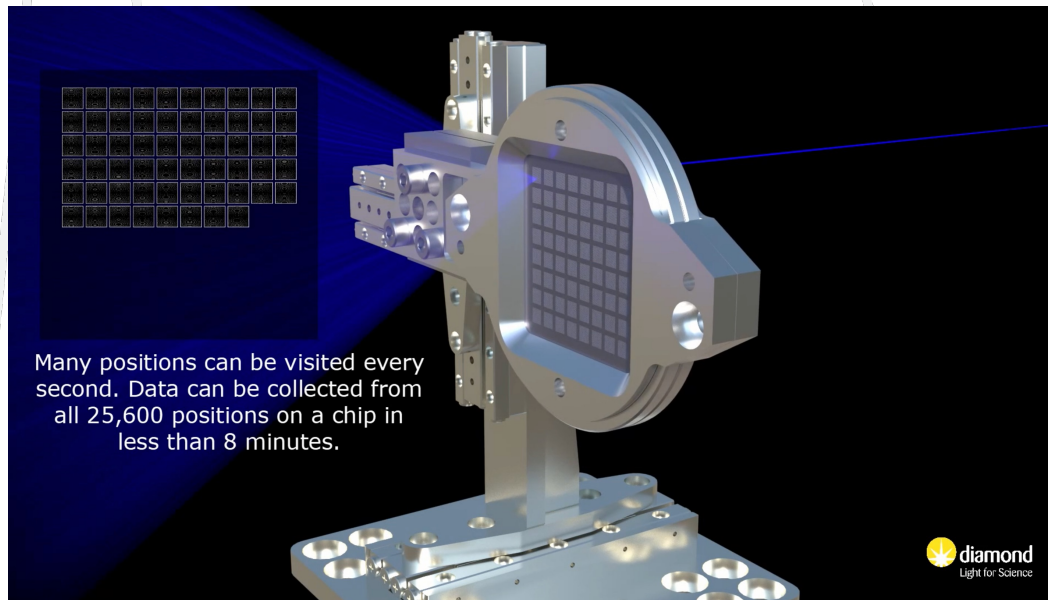
Fraser, J., Clarkson, M., Degnan, S. et al. (2009) Nature 462, 669–673

Sanchez-Weatherby, J., Sandy, J., Mikolajek, H. et al. (2019) J. Synchrotron Rad. 26, 291–301

Nave, C. & Garman, E. F. (2005) J. Synchrotron Rad. 12, 257–260

Time-resolved crystallography

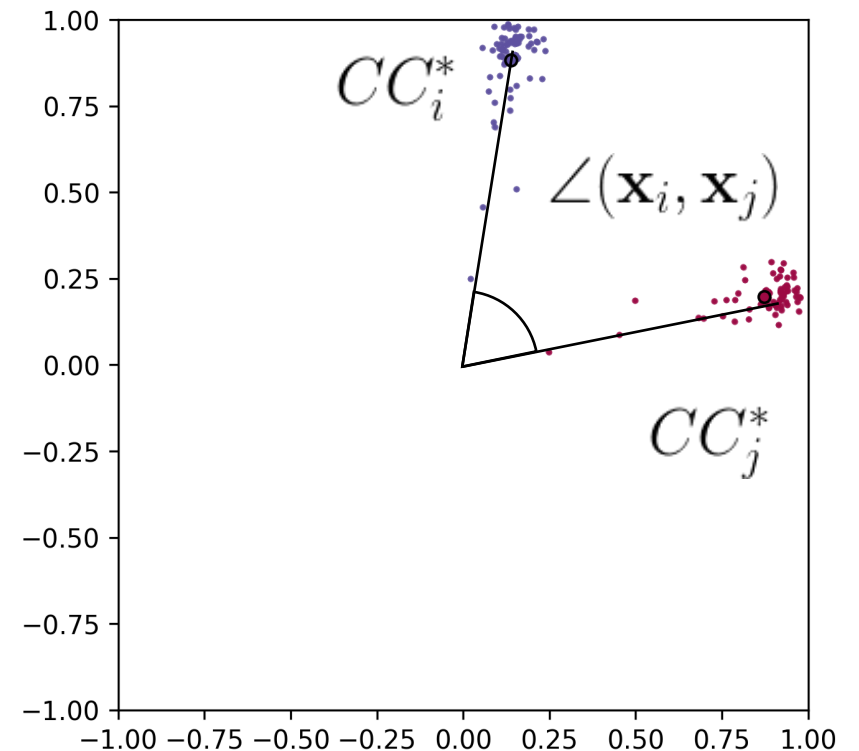
- Study dynamics of light- or chemical-induced reactions using serial femtosecond (SFX) or synchrotron (SSX) crystallography



Challenges

Symmetry determination

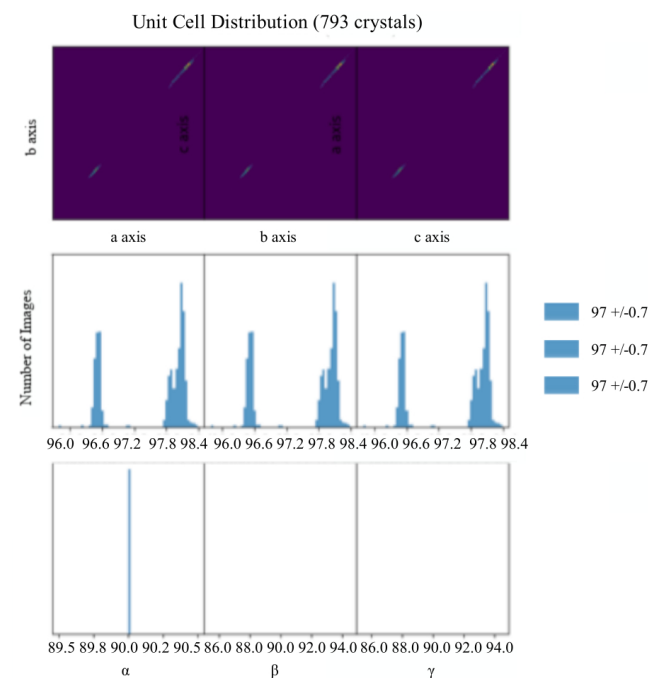
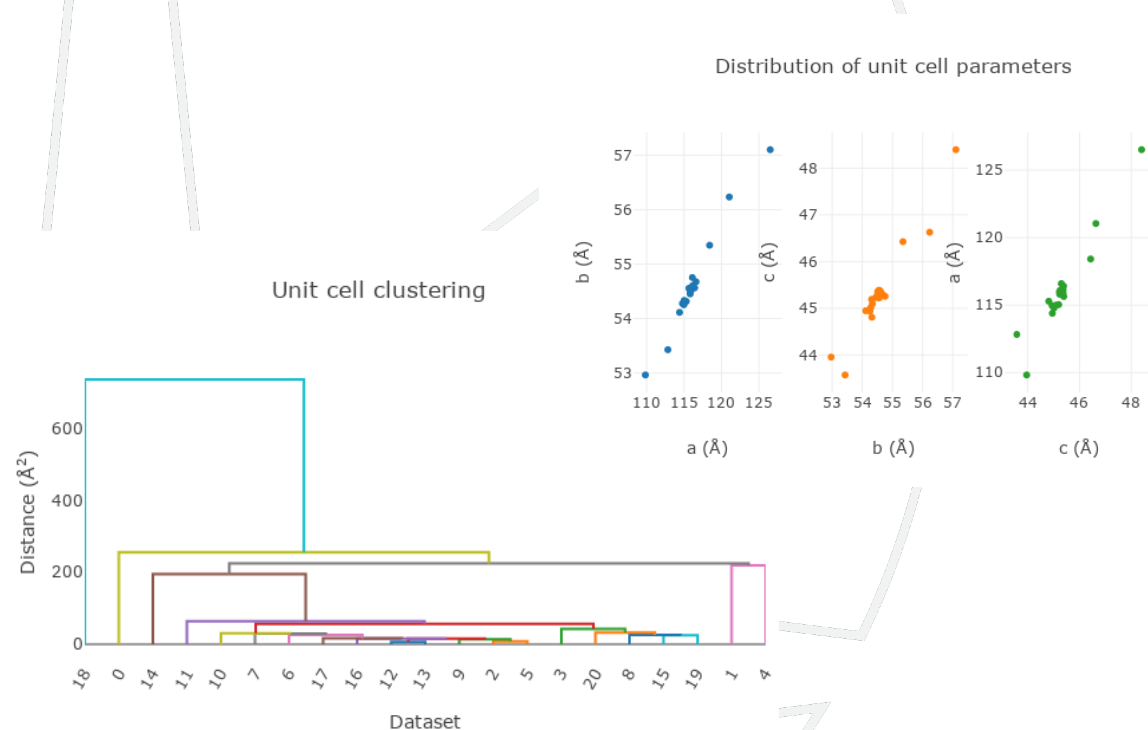
- Identification of consensus symmetry from narrow wedges or stills can be challenging
- Complicated by presence of potential indexing ambiguity
- New algorithms have been developed to help in symmetry determination from narrow wedges and stills



Brehm, W. & Diederichs, K. (2014). *Acta Cryst.* D70, 101–109
Gildea, R. J. & Winter, G. (2018). *Acta Cryst.* D74, 405–410

Non-isomorphism

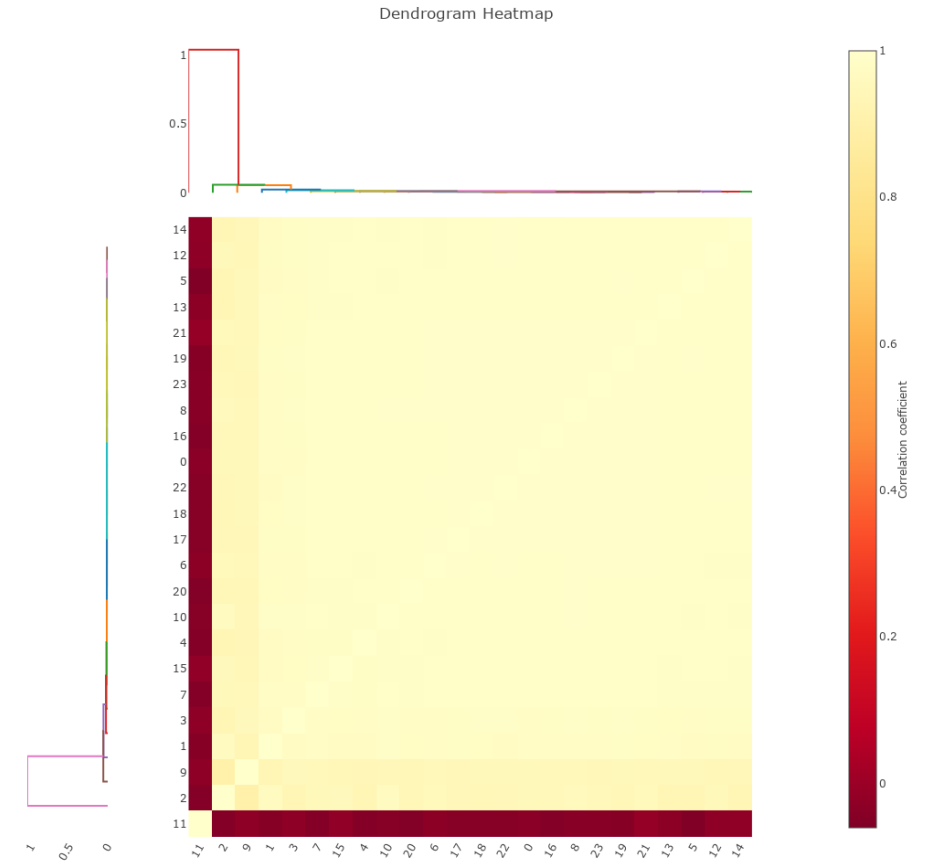
- Inclusion of non-isomorphous crystals may degrade the final data set
- Unit cell clustering may help identify outliers or different populations



Non-isomorphism

Clustering on pairwise correlation coefficients may help identify outliers

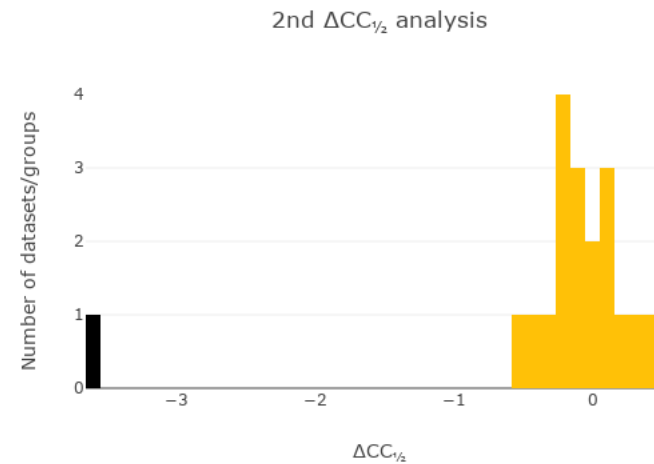
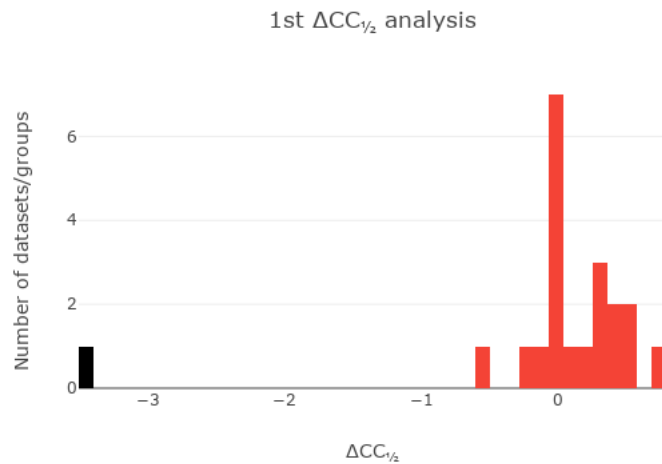
$$r_{i,j} = \frac{\sum_h [I_i(h) - \bar{I}_i] [I_j(h) - \bar{I}_j]}{\left\{ \sum_h [I_i(h) - \bar{I}_i]^2 \sum_h [I_j(h) - \bar{I}_j]^2 \right\}^{1/2}}$$



Giordano, R., Leal, R. M. F., Bourenkov, G. P. et al. (2012). Acta Cryst. D68, 649–658
Santoni, G., Zander, U., Mueller-Dieckmann, C. et al. (2017). J. Appl. Cryst. 50, 1844–1851

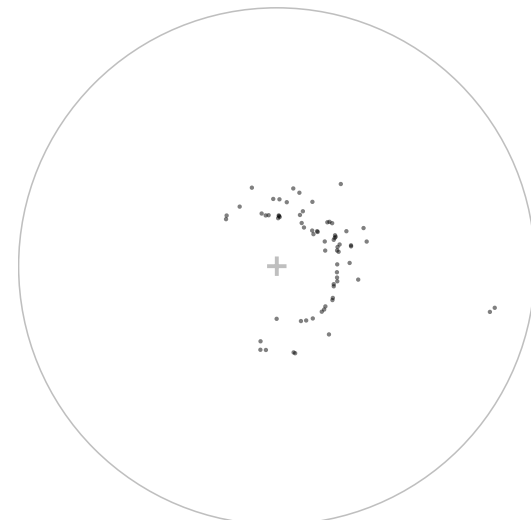
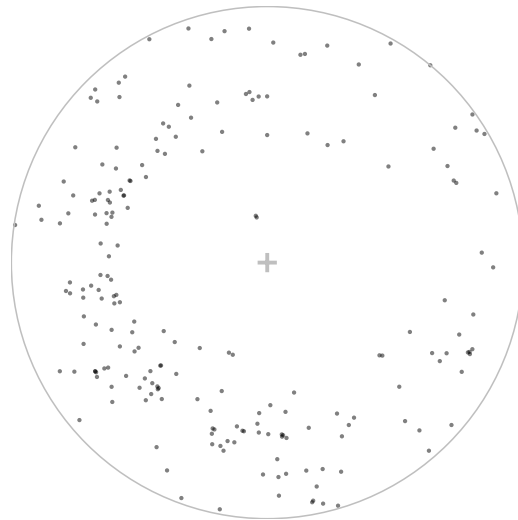
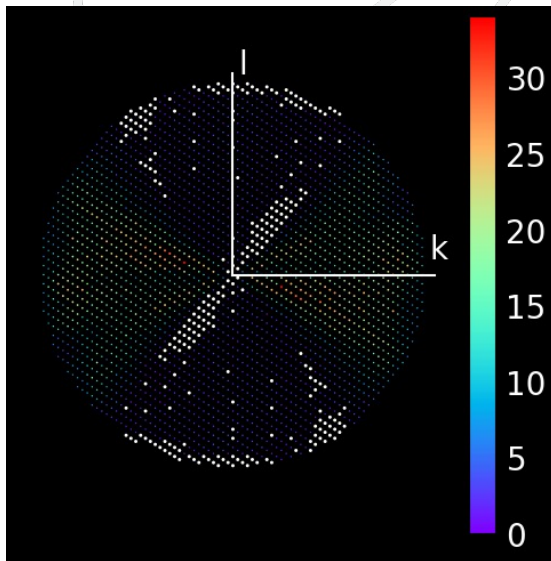
Non-isomorphism

- $\Delta CC_{1/2}$ method omits each data sets in turn and calculates the resulting improvement (or deterioration) in $CC_{1/2}$



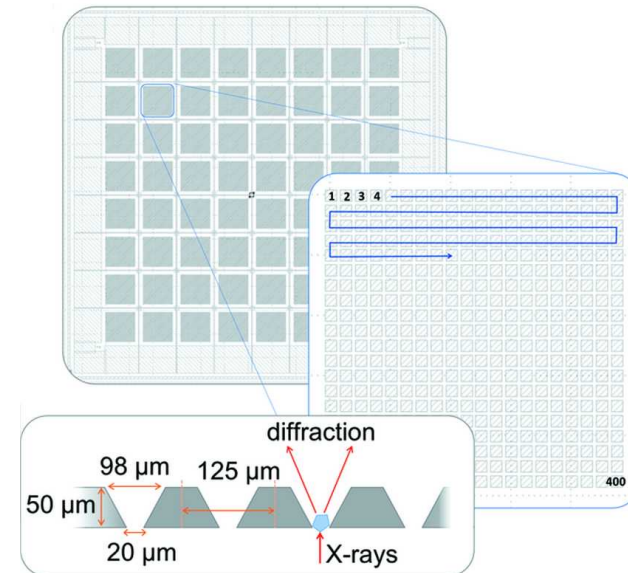
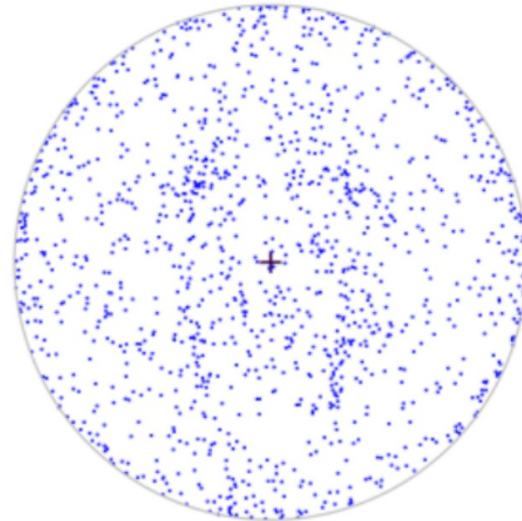
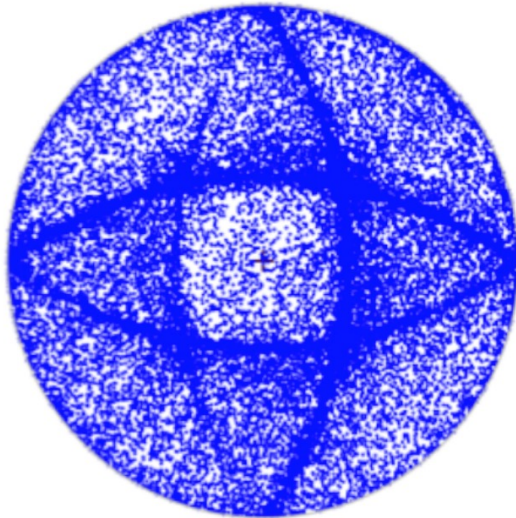
Preferential orientation

- Crystal symmetry and morphology combined with data collection conditions may lead to preferential crystal orientation
- May result in under-sampled regions of reciprocal space
- Check stereographic projection and multiplicity plots



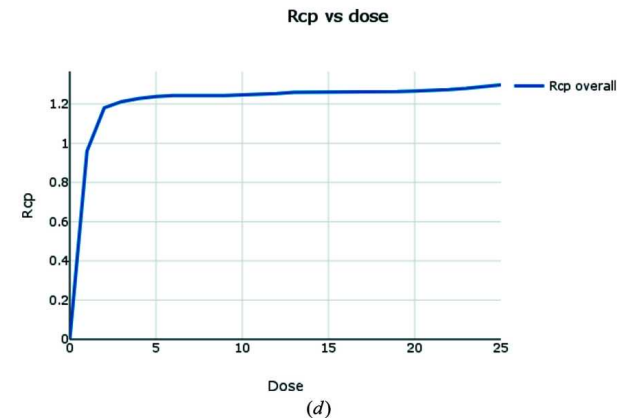
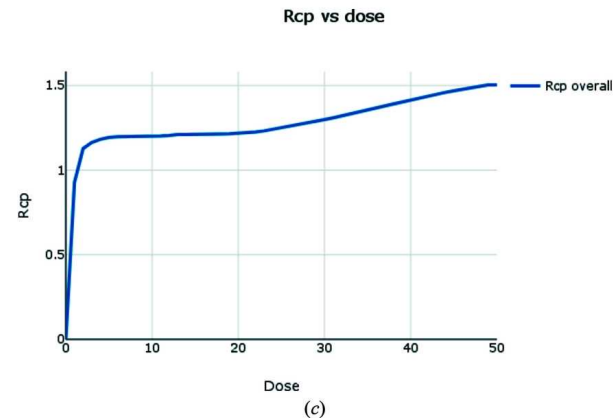
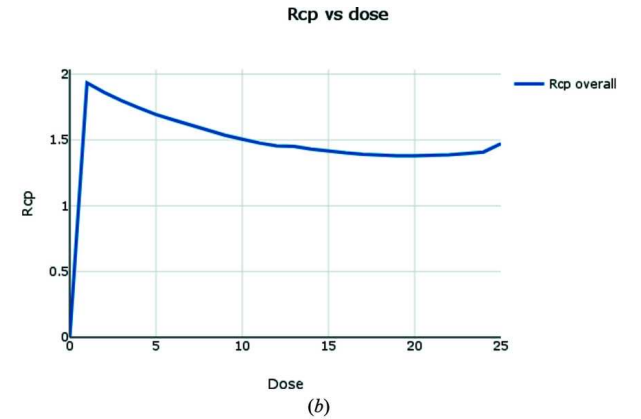
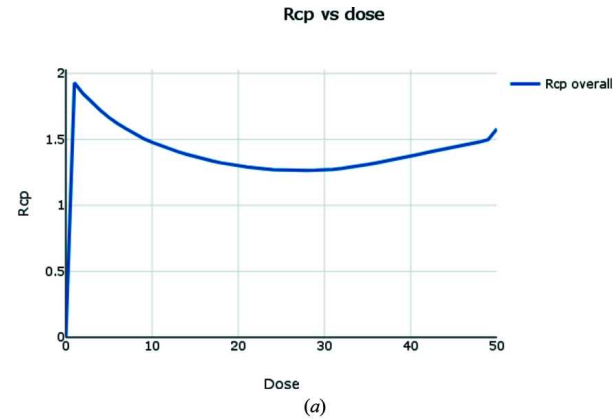
Preferential orientation

- Crystal symmetry and morphology combined with data collection conditions may lead to preferential crystal orientation
- May result in under-sampled regions of reciprocal space
- Check stereographic projection and multiplicity plots



Radiation damage

R_{cp} statistic accumulates the pairwise measured intensity differences as a function of dose, providing an assessment of radiation damage



Winter, G. (2009). PhD thesis

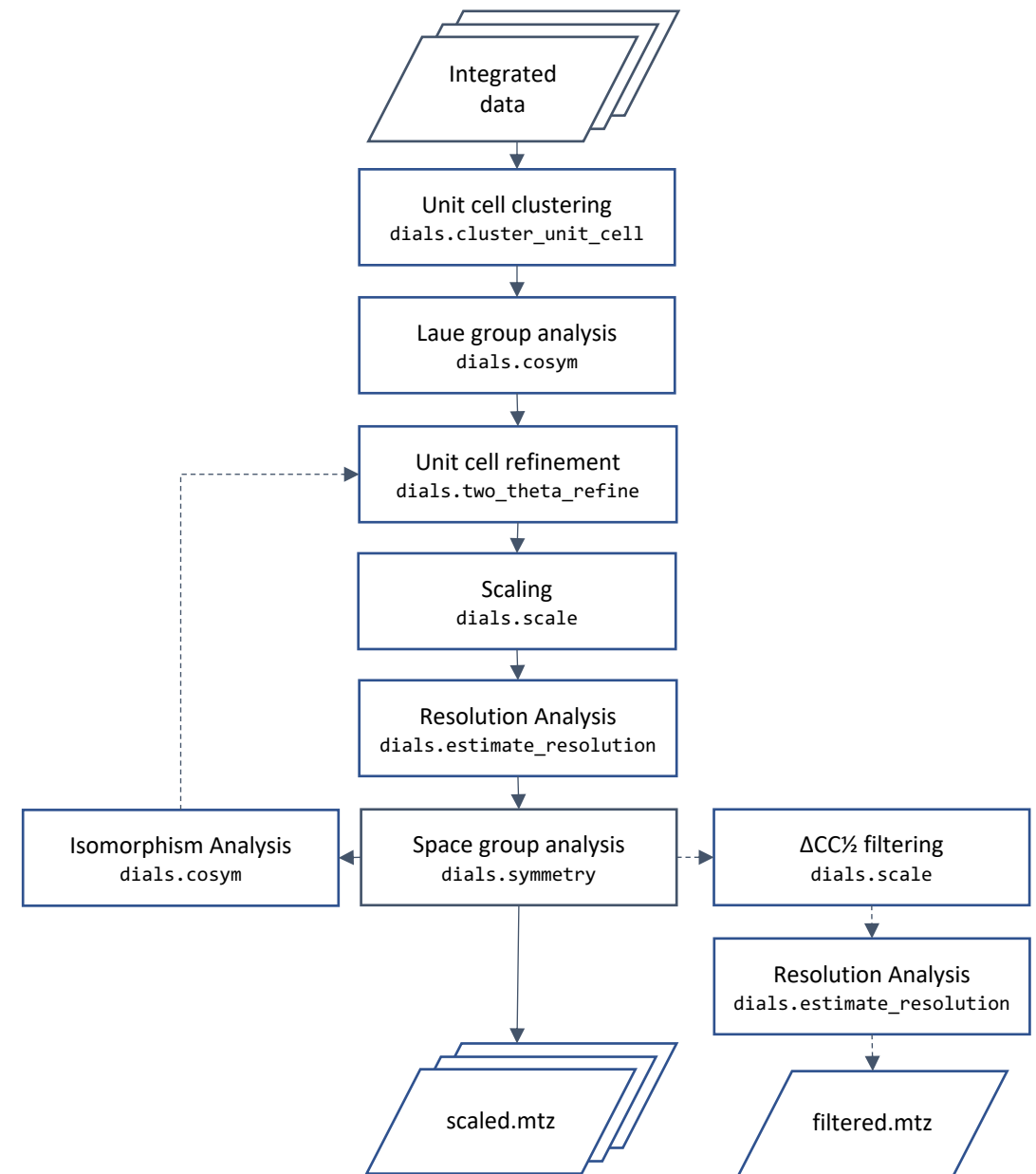
Winter, G., Gildea, R. J., Paterson, N. et al. (2019). Acta Cryst. D75, 242–261



xia2.multiplex

xia2.multiplex

- Automates typical multi-crystal workflows
- Run automatically at Diamond for data collections on the same sample
- Included with DIALS
- Coming soon to CCP4 i2

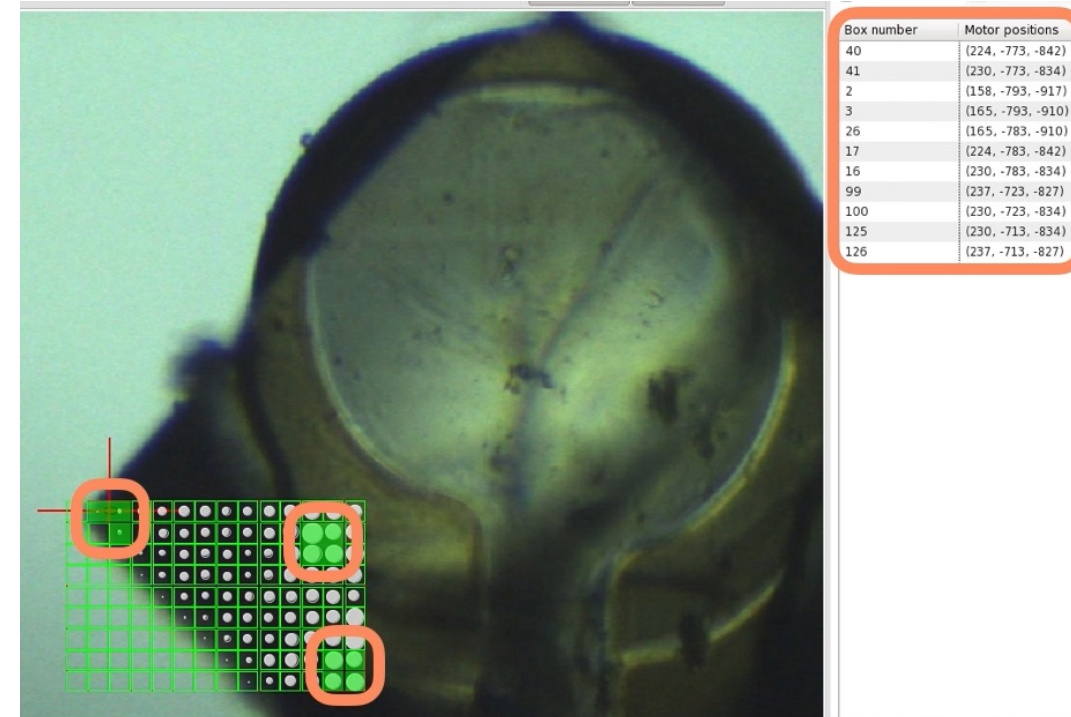




Multi-crystal data collections at Diamond

Multi-crystal data collections on I24

- Microfocus beam: 5 x 5 μm - 50 x 50 μm
- Wide energy range: 7.8 -24 keV (0.7 - 2.0 Å)
- CdTe Eiger2 9M: ideal for high energy
- Pilatus3 6M: suitable for medium energy
- Fast gridscans: queue data collection from grid scan results
- Serial crystallography: fixed target, thin film and LCP extruder

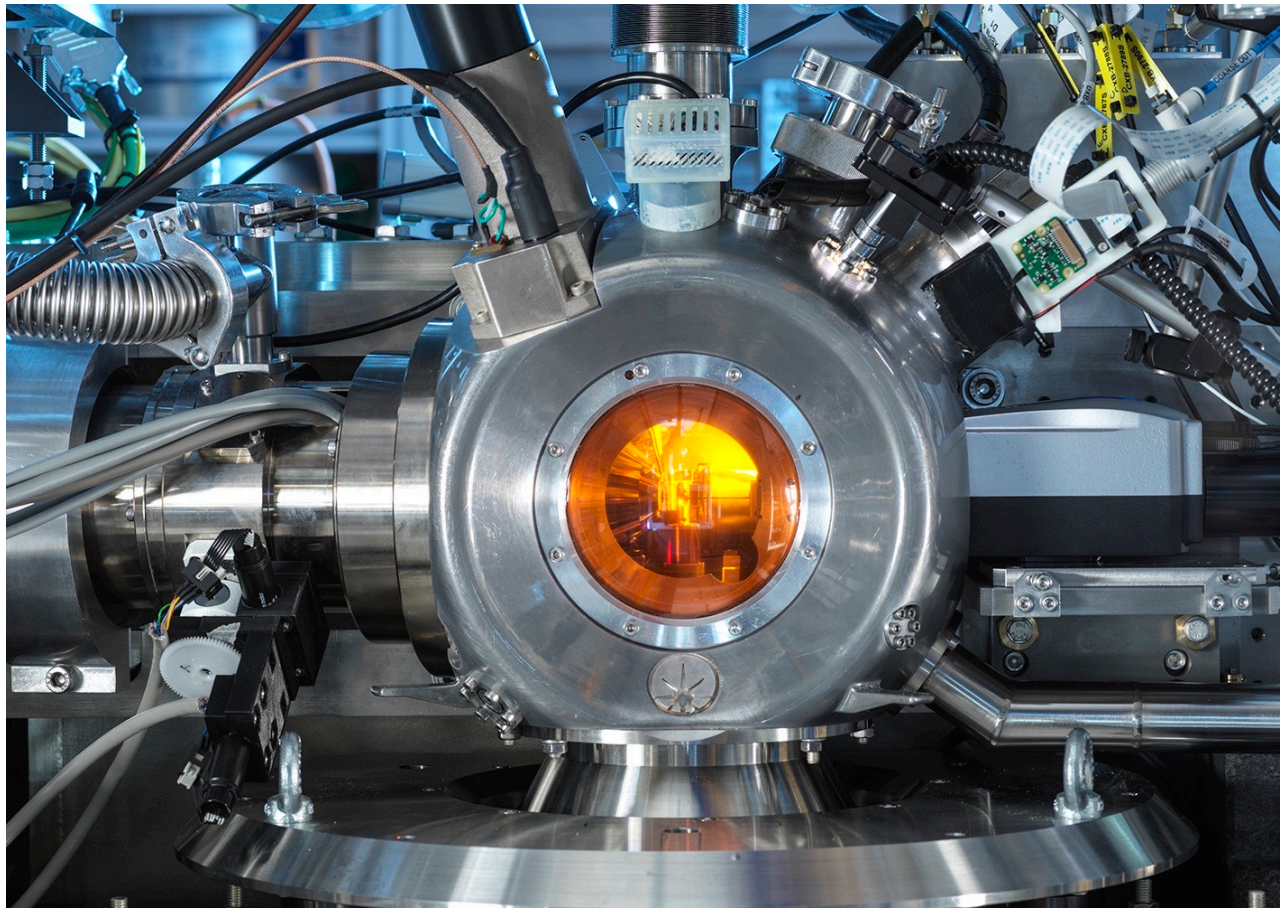


Multi-crystal data collections on VMXi

- *in-situ* beamline dedicated to room temperature data collection from macromolecular crystals
- Entirely automated facility
- Oscillation experiments (up to 60 degrees of data / crystal) for crystals >20 μm with automatic data processing and merging with results accessed via ISPyB
- Grid scans over a user-defined area to assess diffraction quality
- Provides feedback on the diffraction, crystal quality and, in many cases, unit-cell parameters and space group, even in the case of micro-crystals (2-5 μm), allowing effective optimisation of crystallisation
- Structural data is collected at near physiological conditions

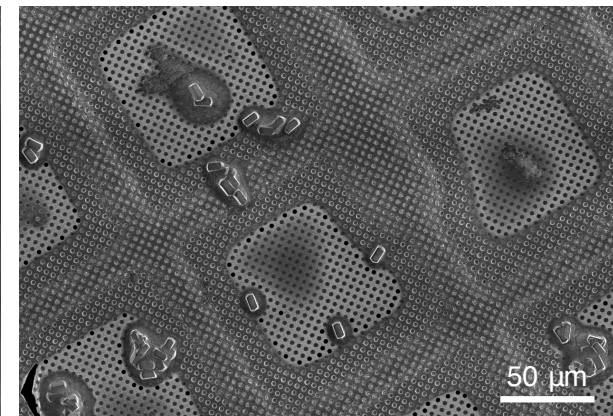
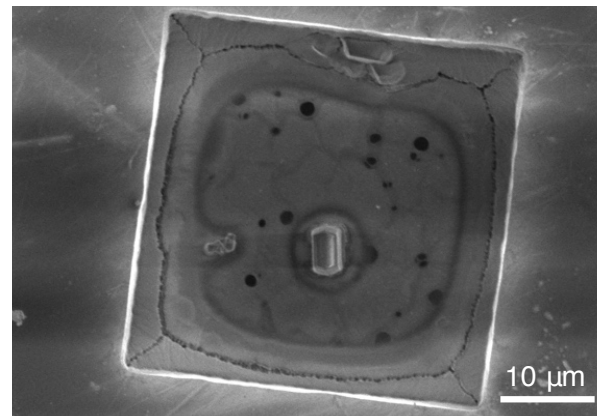
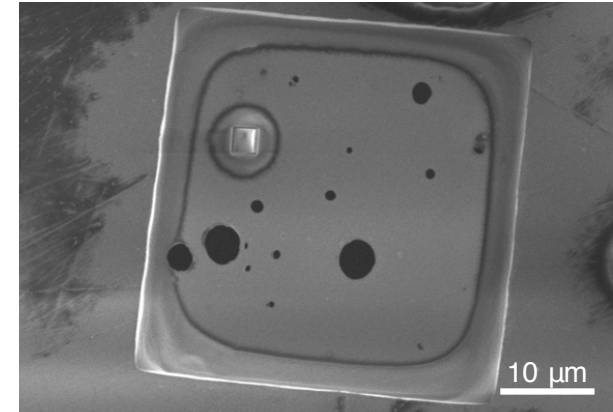
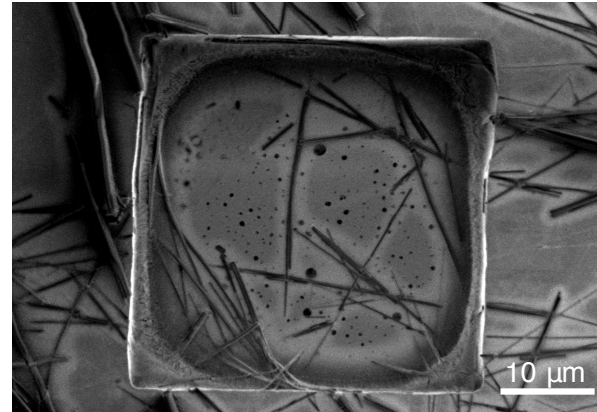
Multi-crystal data collections on VMXm

- Aim: single crystal rotation data collection from micron sized crystals.
- Micro/nano-focus MX beamline:
 - Focused beam (V x H): $0.4\ \mu\text{m} \times 1.3\ \mu\text{m}$
 - Defocused beam (V x H): $9\ \mu\text{m} \times 13\ \mu\text{m}$
- Standard data collection energy: 21.3 KeV
 - Eiger X 9M CdTe (High energy)
 - Pilatus 3 6M (Medium energy)
- In vacuum sample position – low background, high signal-to-noise data collection.
- Data collection at cryogenic temperatures.
- Integrated scanning electron microscope.



When to use VMXm

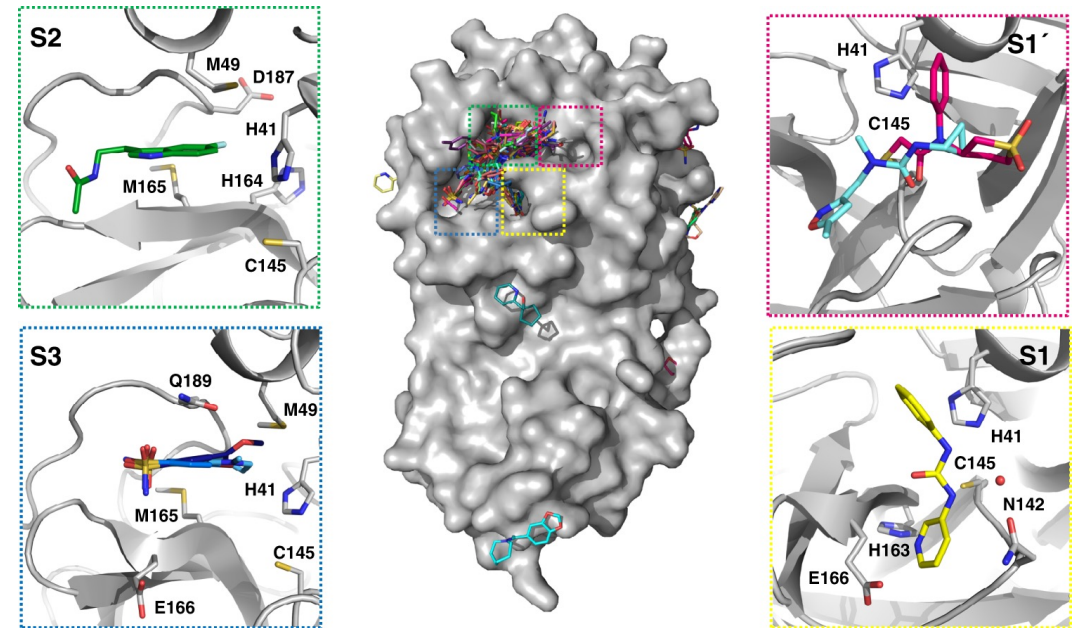
- Samples measuring less than 15 μm to 1 μm .
- Needles benefit from low signal-to-noise and micro focused X-ray beam.
- If you can see the microcrystals – VMXm can measure them!
- If you can't see the microcrystals – see electron diffraction.
- VMXm is sample efficient - large number of crystals on a single grid.
- A single sparse matrix crystallization droplet can provide enough material for 3 grids.



Examples

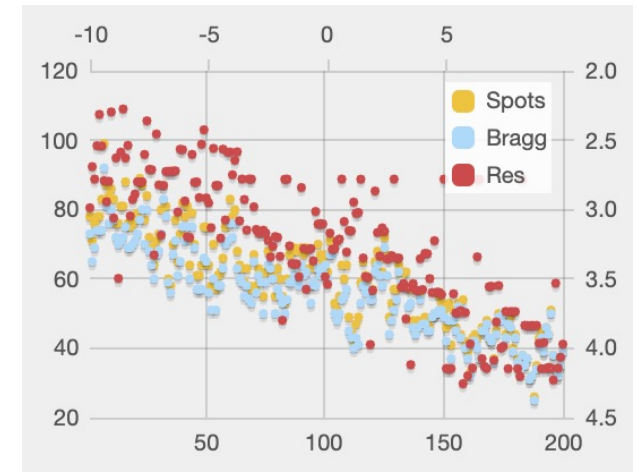
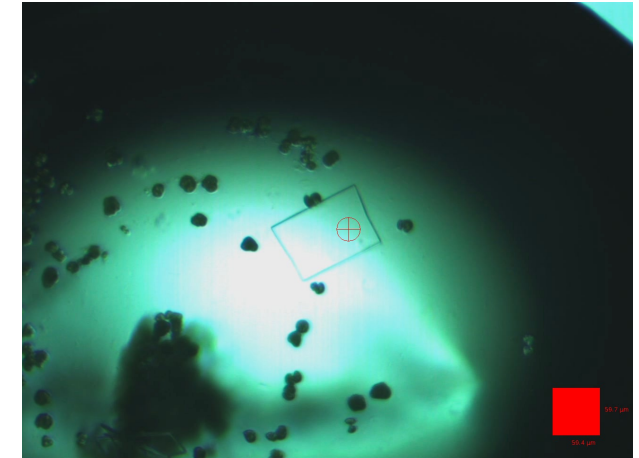
SARS-CoV-2 M^{pro}: ligand screening

- SARS-CoV-2 main protease
- Central role in viral replication
- Key antiviral drug target
- Conventional fragment-screening campaign performed on I04-1
- Over 1250 unique fragments, identifying 74 high-value fragment hits

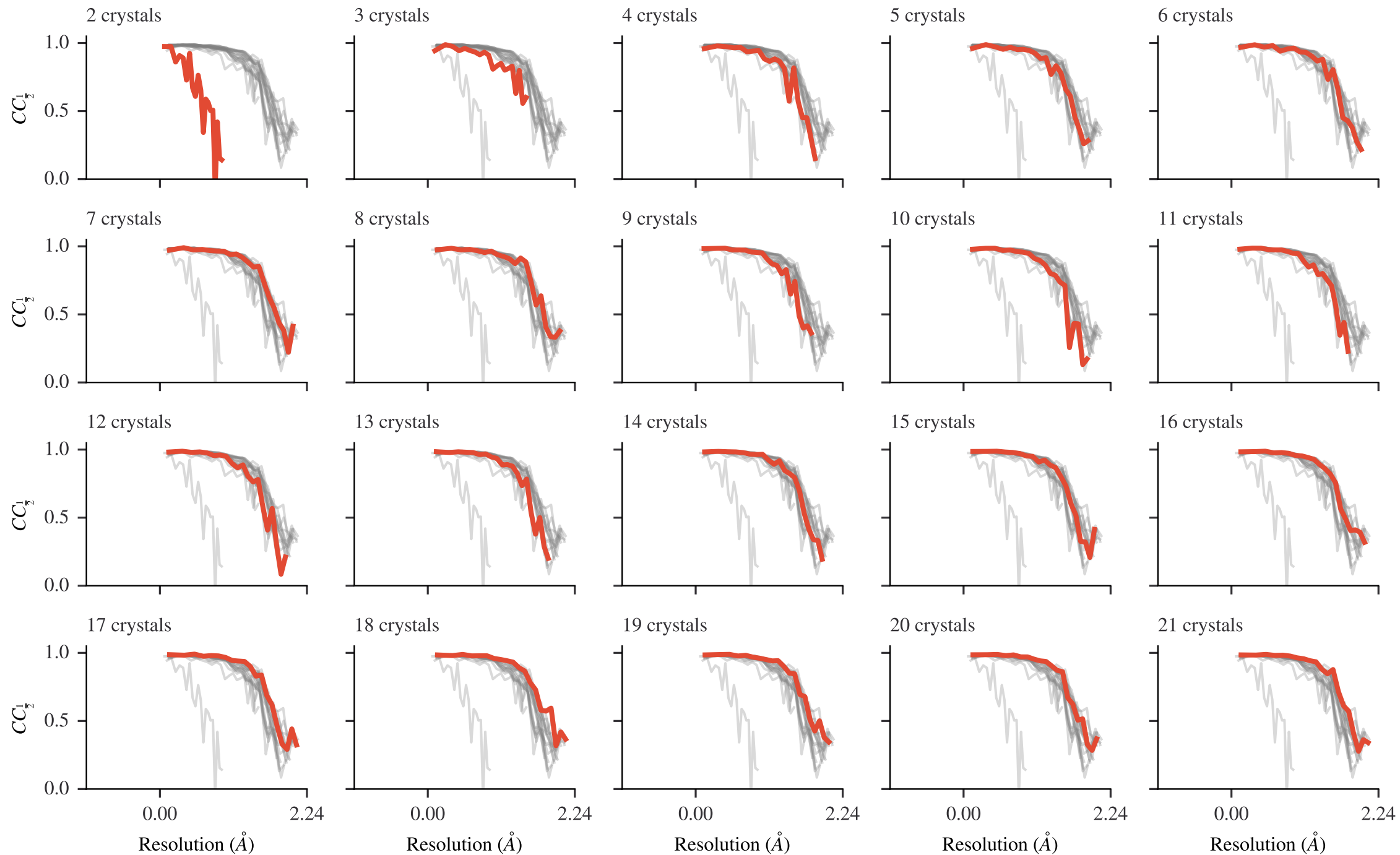


SARS-CoV-2 M^{pro}: *RT* ligand screening

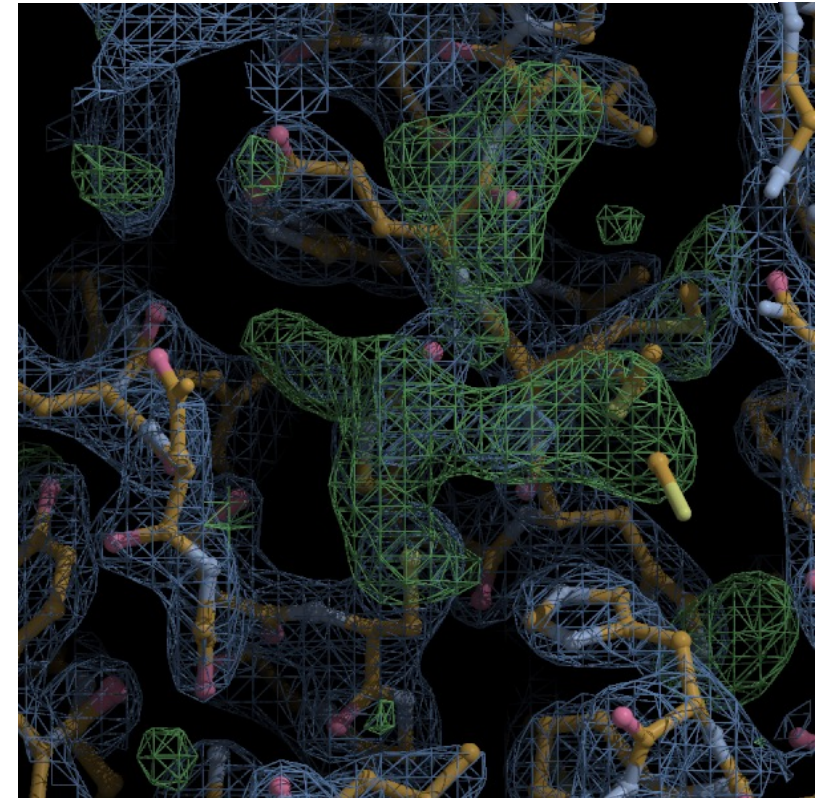
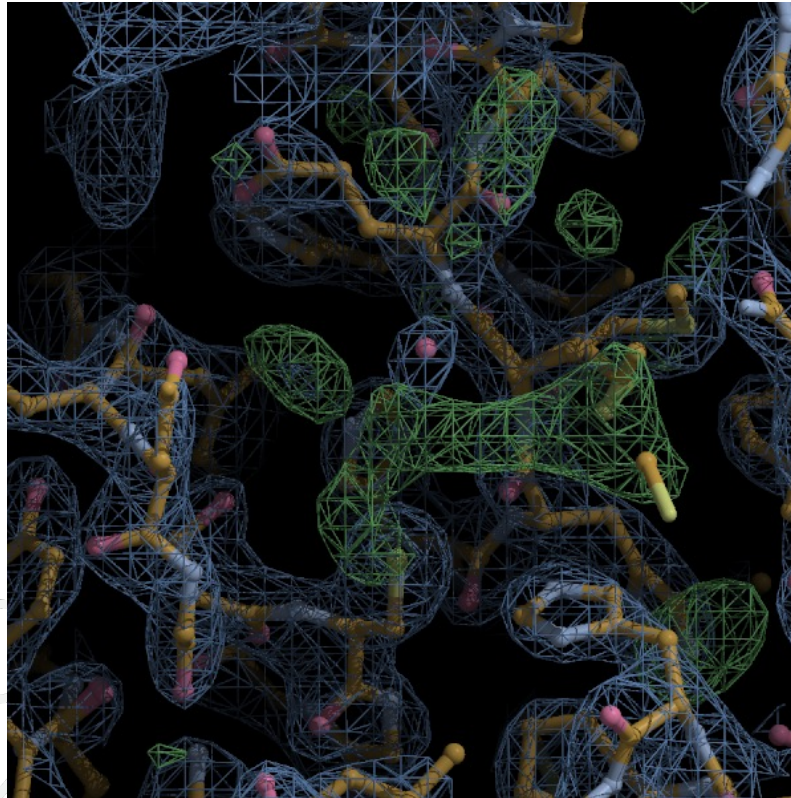
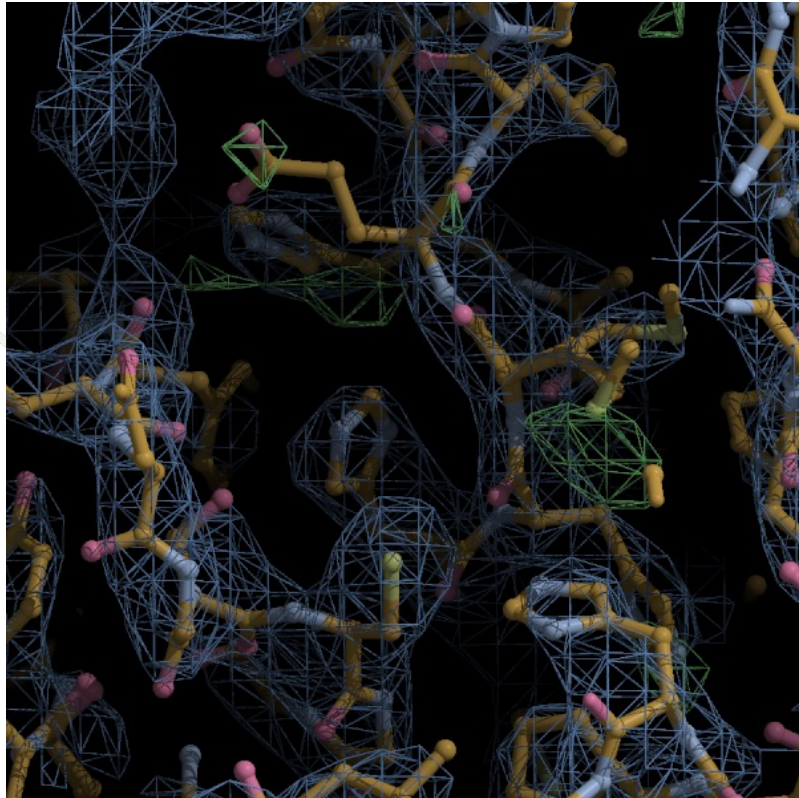
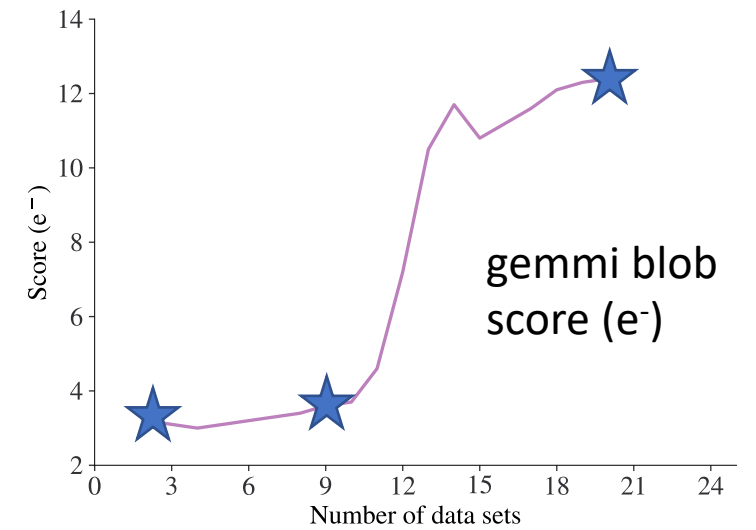
- Initial fragment screening performed at cryo-temperatures (100 K)
- Are room temperature structures identical?
- RT in situ data collections on known ligand hits performed on I24 and VMXi
- Preferred orientation (plate-like crystals): vary starting angle
- xia2.multiplex provided near real time feedback during the experiment



xia2.compare_merging_stats unmerged_1.mtz
unmerged_2.mtz [...] small_multiples=True

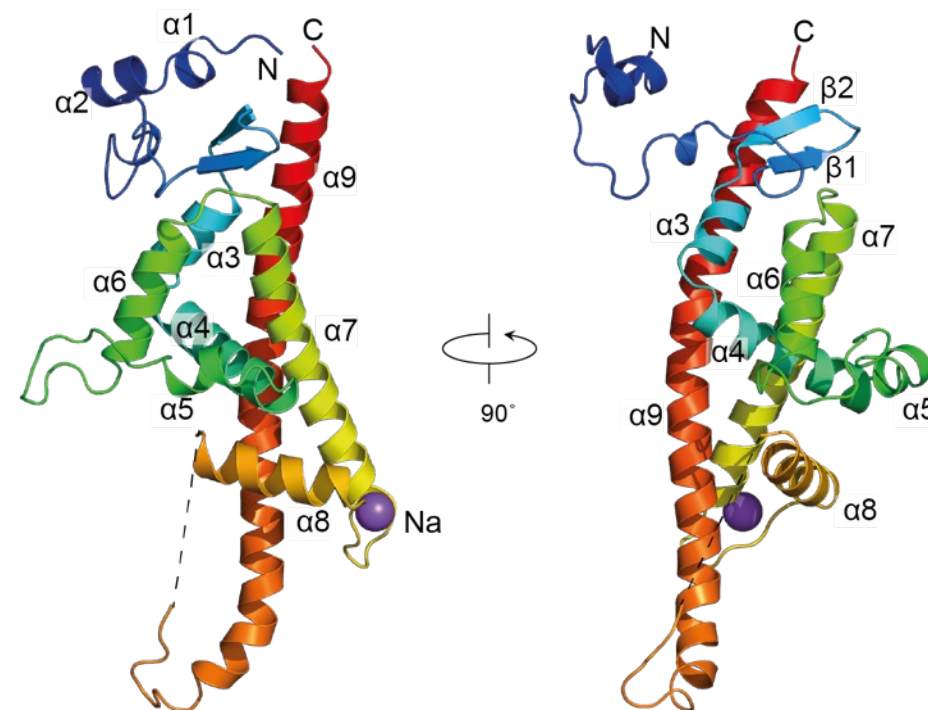
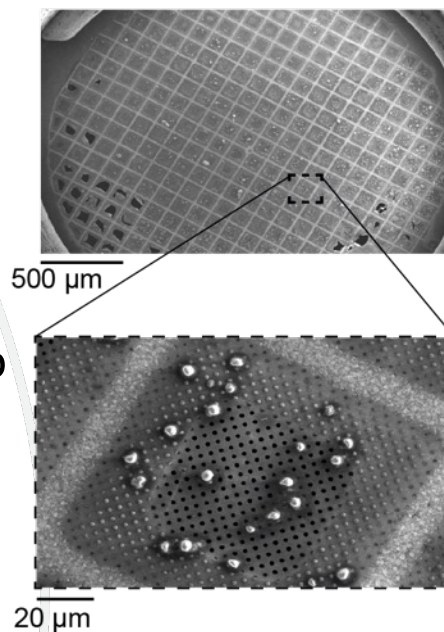


Automatic dimple maps



ToNV

- Polyhedra protein from a nudivirus
- Self-assembles into a dense lattice around new viral particles
- MR failed despite AlphaFold models
- 3 x Met engineered into WT clone to allow SeMet protein
- $P3_121$ $a=53.5 \text{ \AA}$, $c=105.2 \text{ \AA}$, $\gamma=120^\circ$
- Solvent content 21%
- Crystals $\sim 5 - 7 \text{ \mu m}$ (SeMet); $3 - 5 \text{ \mu m}$ (WT)



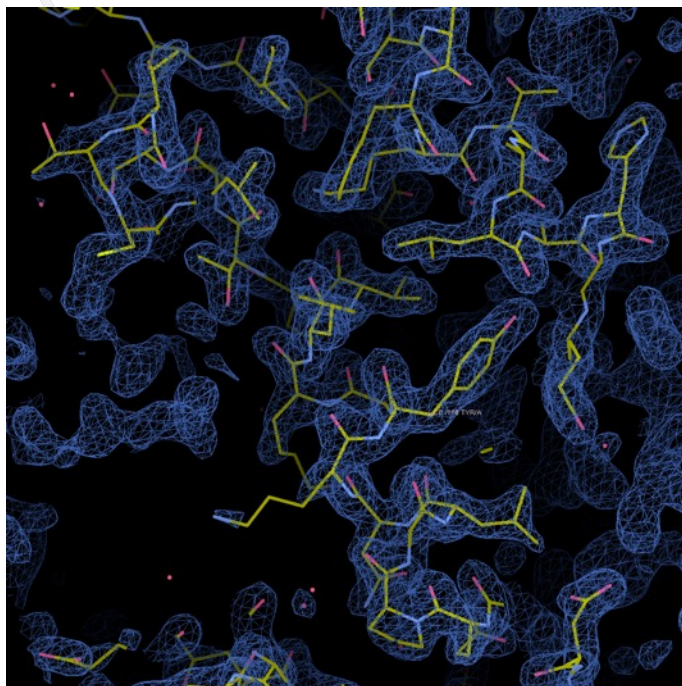
Jeremy Keown and Jonathan Grimes (STRUBI, Uni. of Oxford)

Adam Crawshaw, Richard Gildea, Jose Trincao (VMXm)

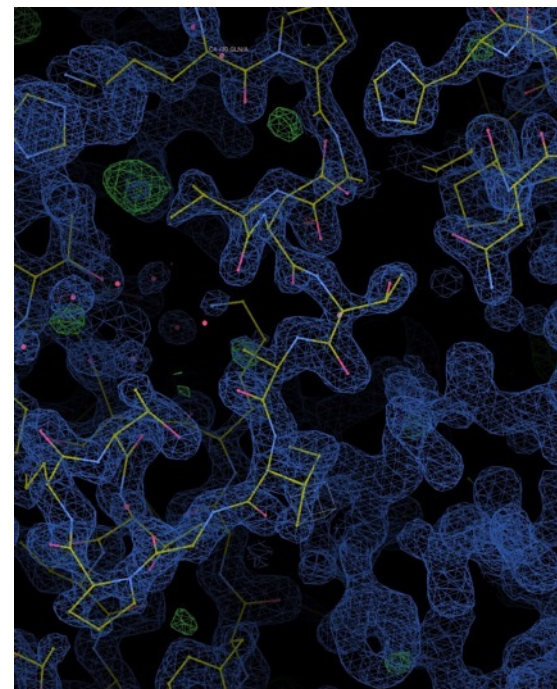
ToNV Se-SAD structure solution

Se-Met TonV on VMXm	
Detector	Eiger2 X CdTe 9M
Number of crystals	67
Energy keV	12.67 keV
Resolution	105.0-1.86 (1.90-1.86)
Unique Reflection	13841(244)
Completeness (%)	91.3(32.0)
R _{merge}	0.288(1.342)
I/ σ I	17.7(0.7)
CC1/2	0.971(0.133)
Beamsize	3.6 x 3.6 μ m

Se-Met to 1.86Å



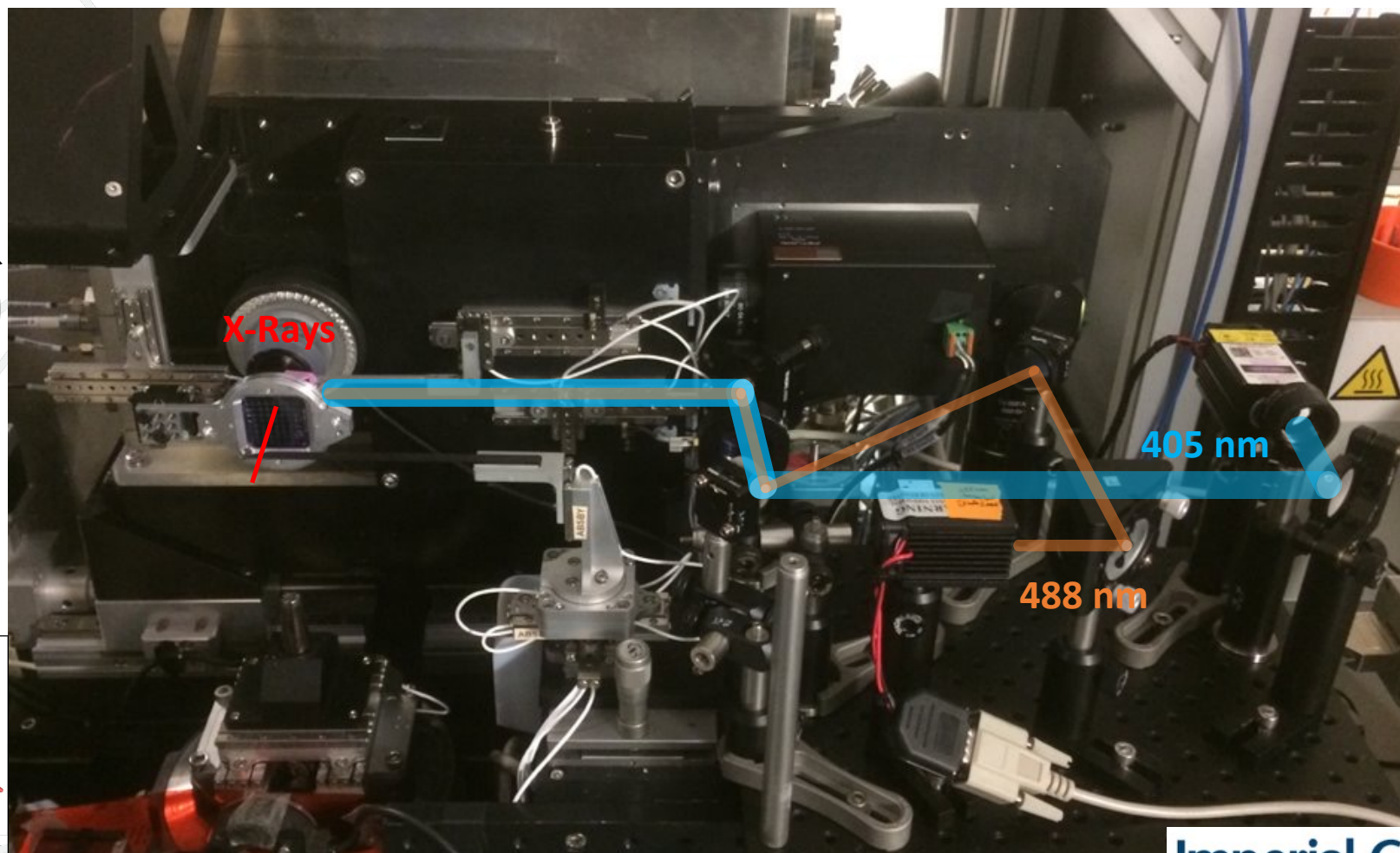
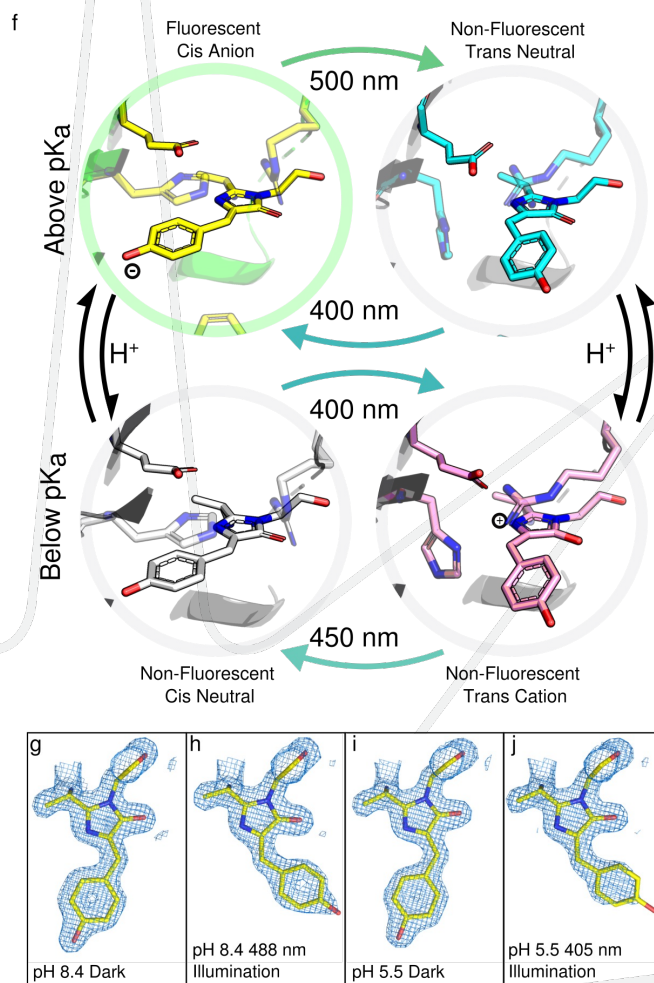
WT to 1.7Å



Analysis:

xia2-dials → xia2.multiplex →
Fast_ep/Crank2

I24 SSX: Van Thor Group – Imperial College London



Baxter, J. M., Hutchison, C. D., Maghlaoui, K. et al. (2022) J. Phys. Chem. B, 126, 45, 9288–9296

Acknowledgements

- Pierre Aller
- Danny Axford
- James Beilsten-Edmands
- Adam Crawshaw
- Nick Devenish
- Jonathan Grimes
- Sam Horrell
- Jeremy Keown
- David Owen
- Robin Owen
- Juan Sanchez-Wetherby
- James Sandy
- Selina Storm
- Jose Trincao
- Martin Walsh
- David Waterman
- Graeme Winter