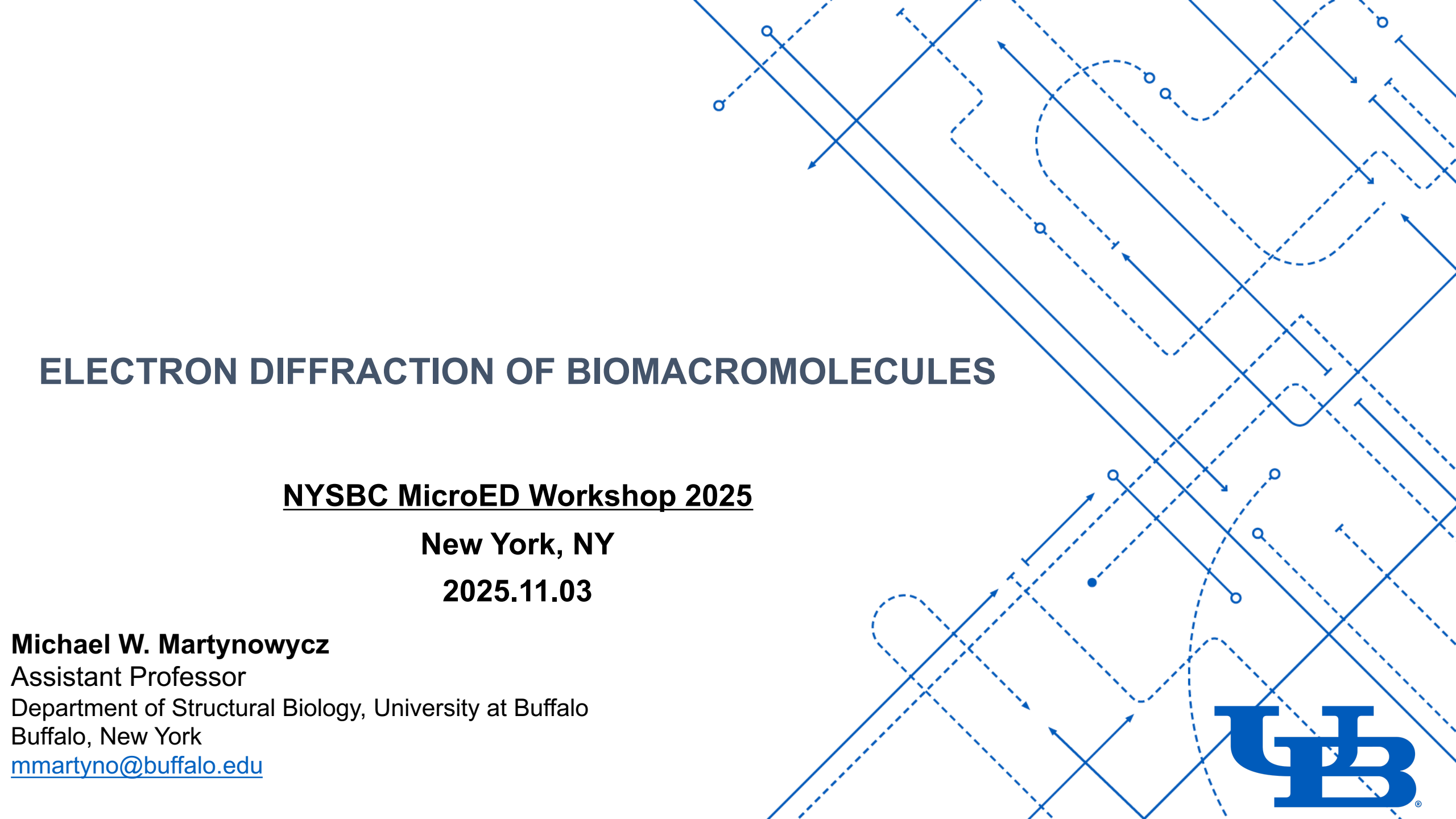


An abstract graphic in the top right corner of the slide. It features a network of blue lines, some solid and some dashed, with arrows indicating direction. Small blue circles are scattered along these lines. The lines intersect and curve, creating a complex, geometric pattern.

ELECTRON DIFFRACTION OF BIOMACROMOLECULES

NYSBC MicroED Workshop 2025

New York, NY

2025.11.03

Michael W. Martynowycz

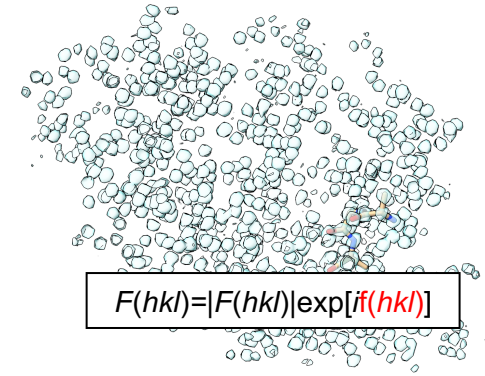
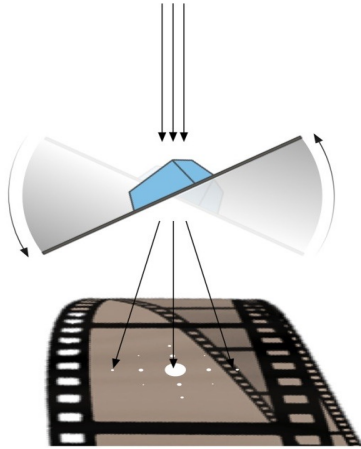
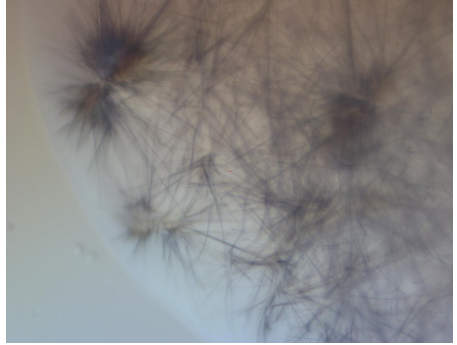
Assistant Professor

Department of Structural Biology, University at Buffalo
Buffalo, New York

mmartyno@buffalo.edu



Macromolecular structure determination in MicroED



Crystallization

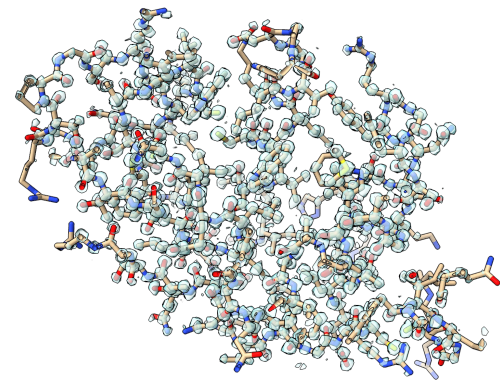
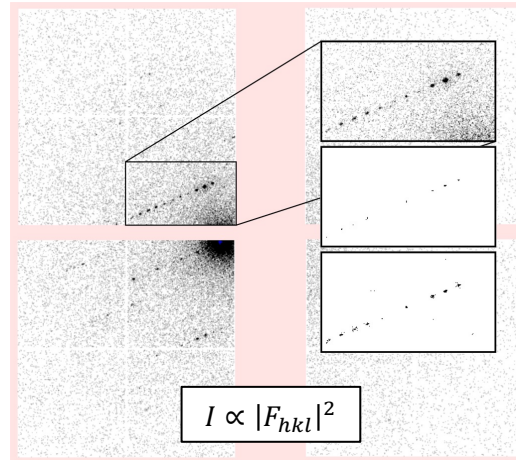
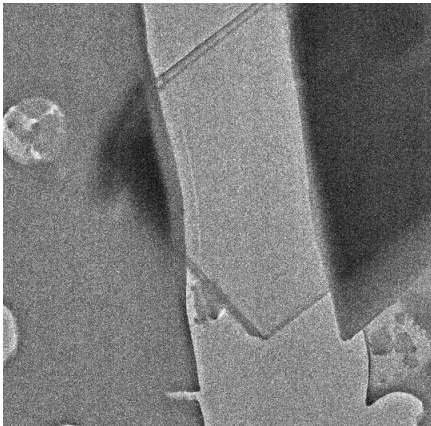
FIB milling

Data collection

Integration

Phasing

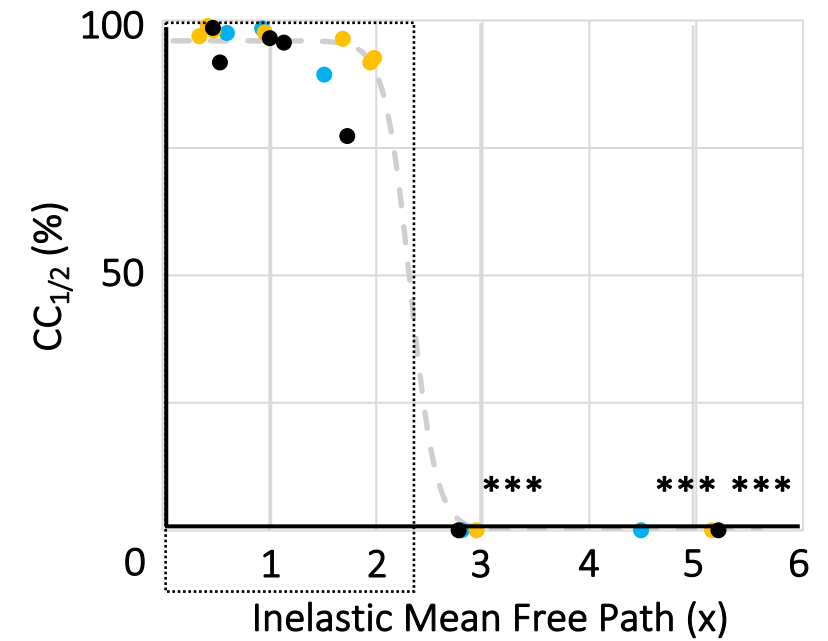
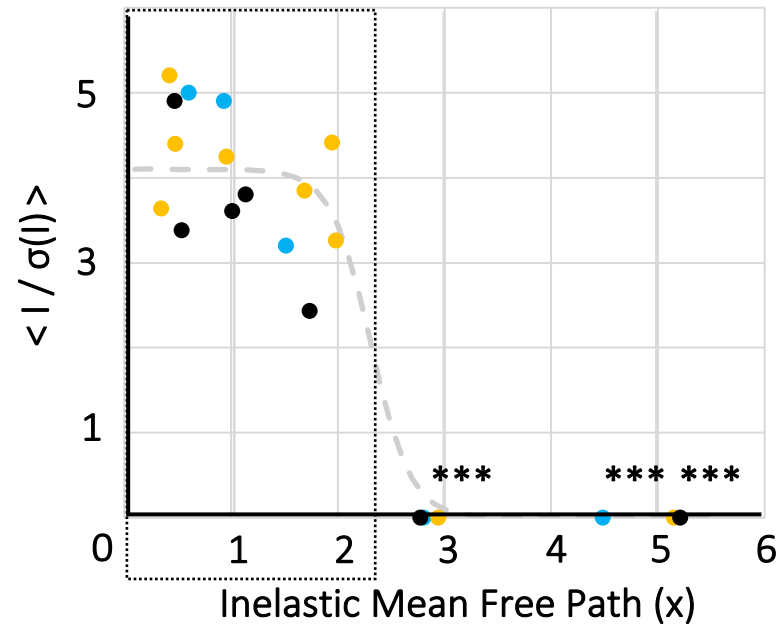
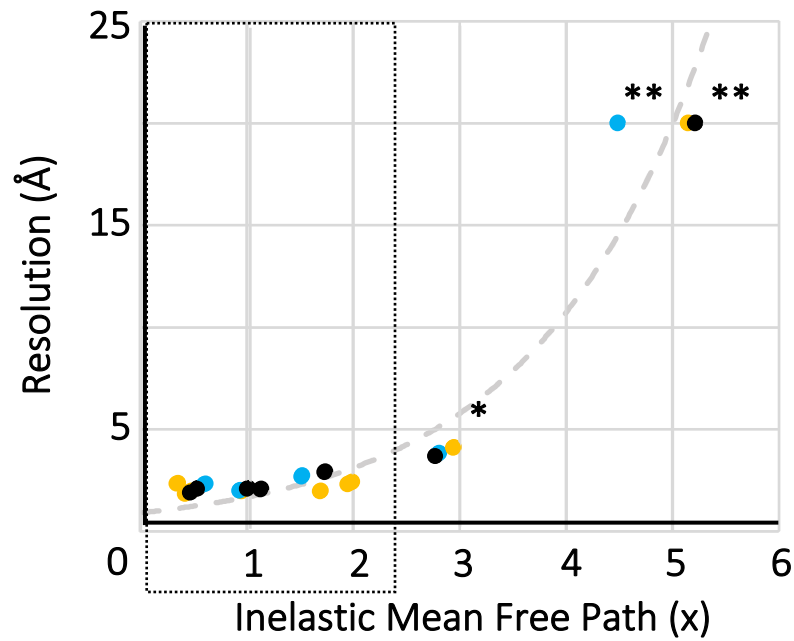
Model building
and refinement



$$I \propto |F_{hkl}|^2$$

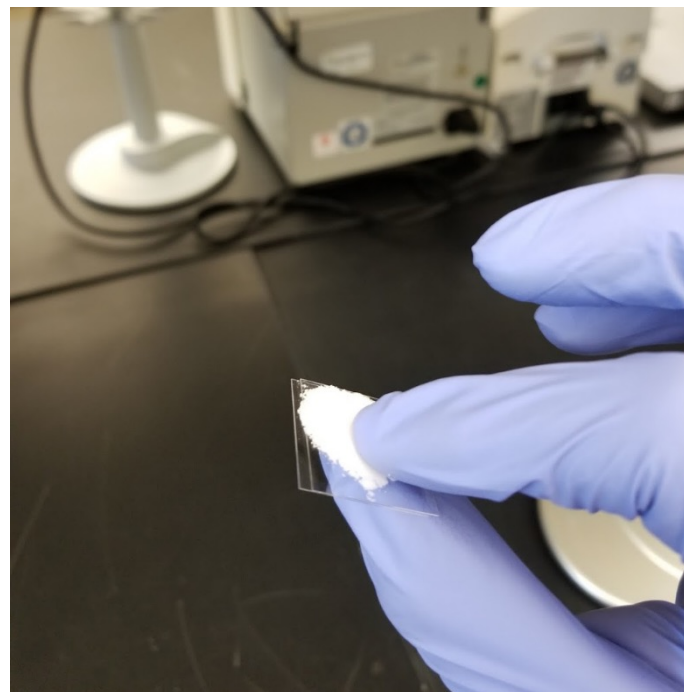
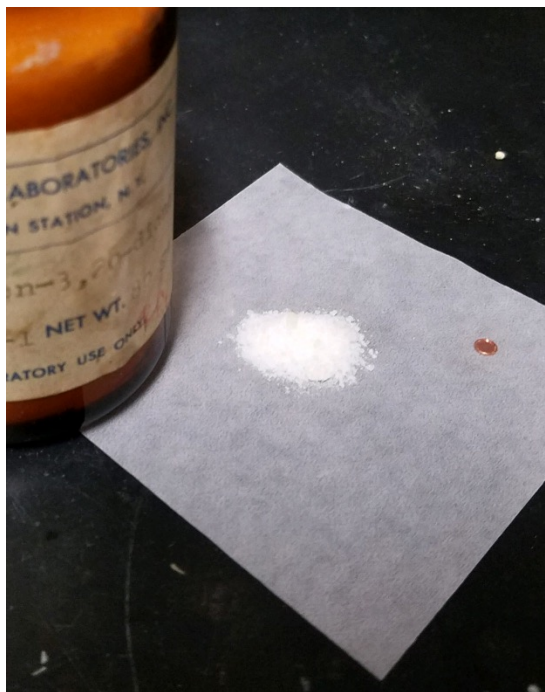
What thickness?

● 120 kV ● 200 kV ● 300 kV



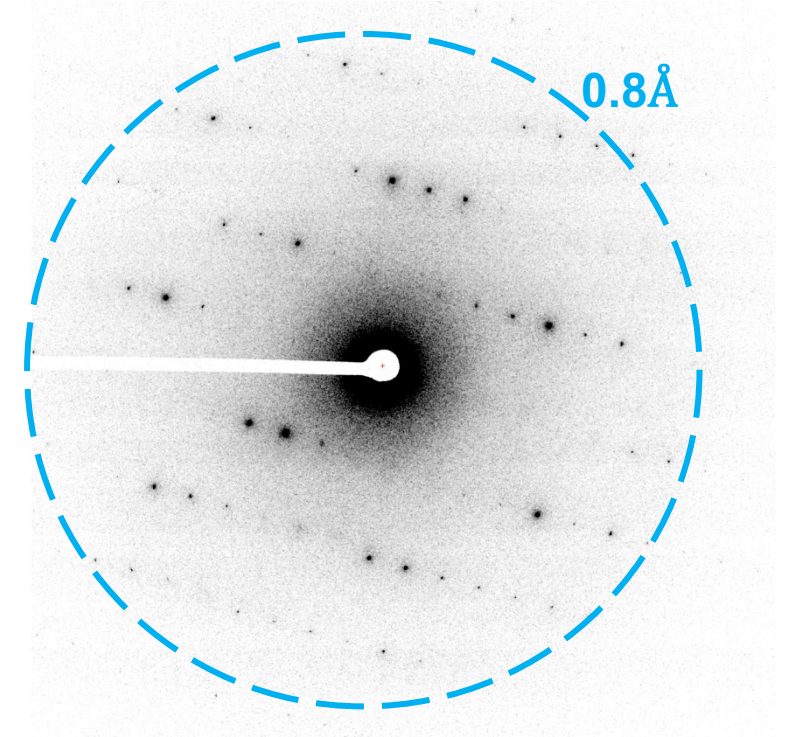
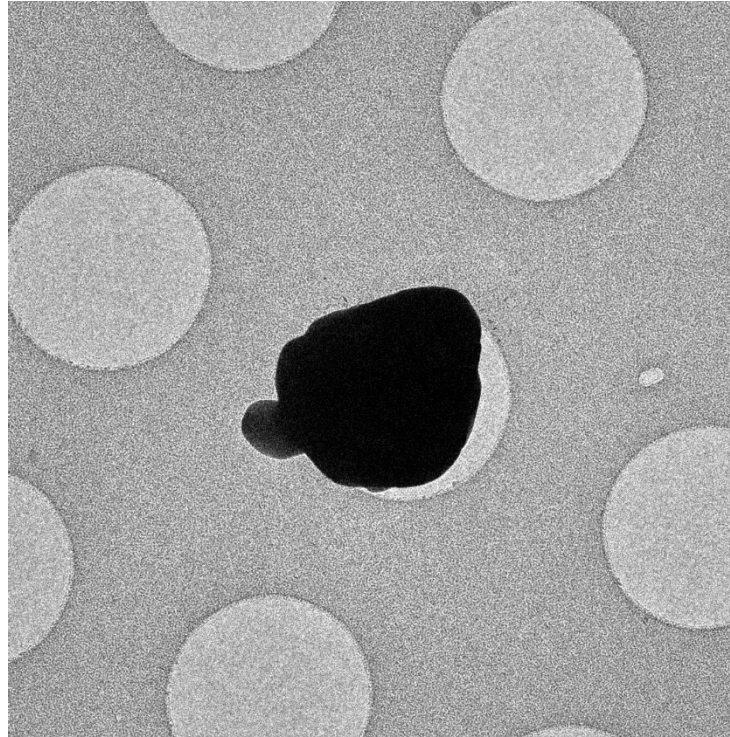
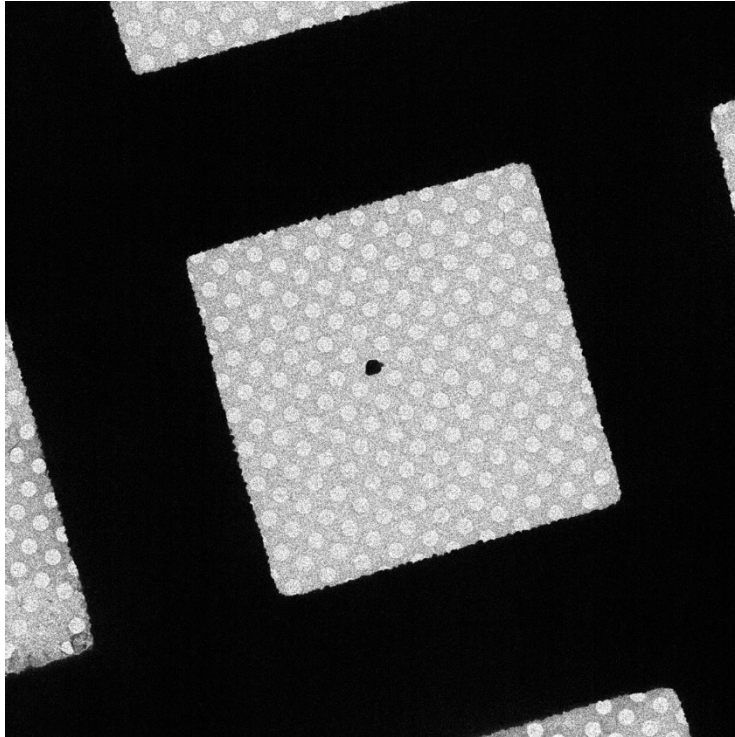
Martynowycz et al. PNAS 2021a

Small molecules – straight forward



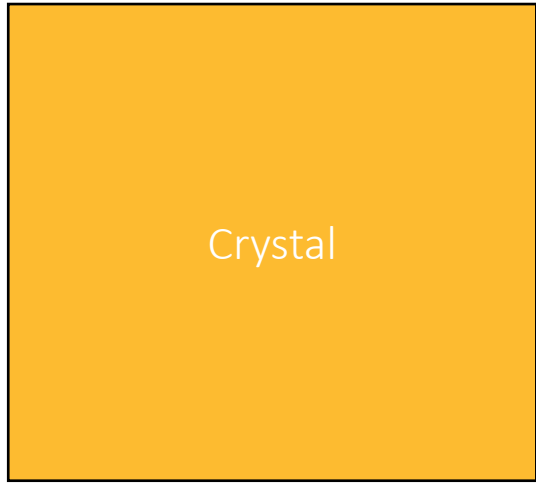
Jones et al. ACS Central Science 2018
Gruene et al. Angew. Chem. Int. 2018

Small molecules – straight forward



Martynowycz, et al. *JOVE* (2021)

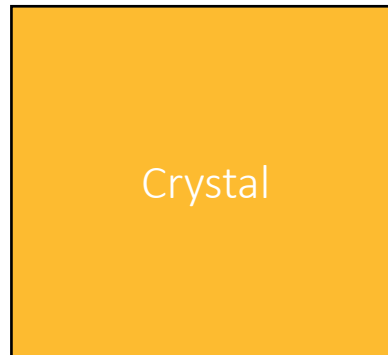
Protein crystals



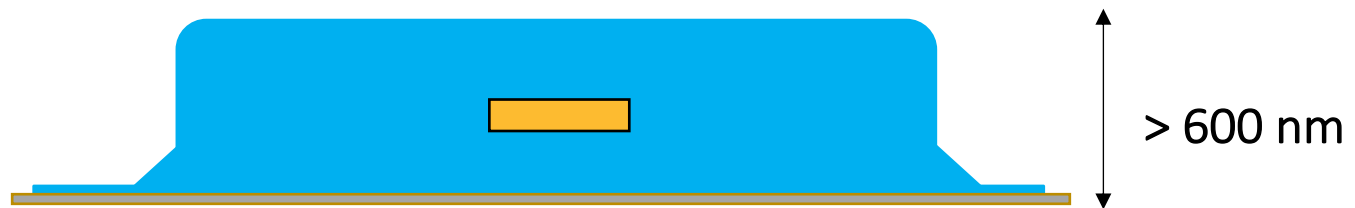
$> 10 \mu\text{m}$



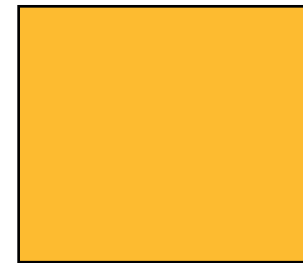
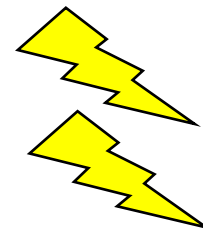
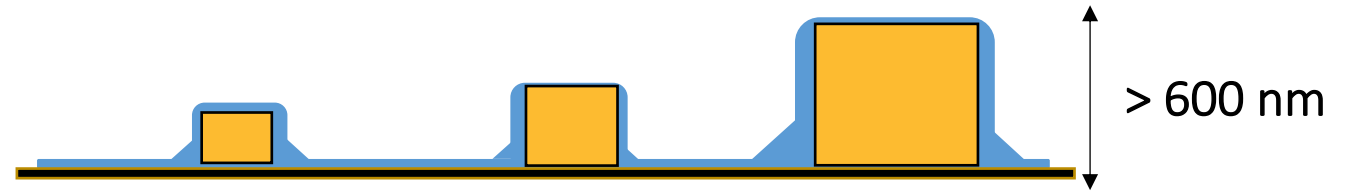
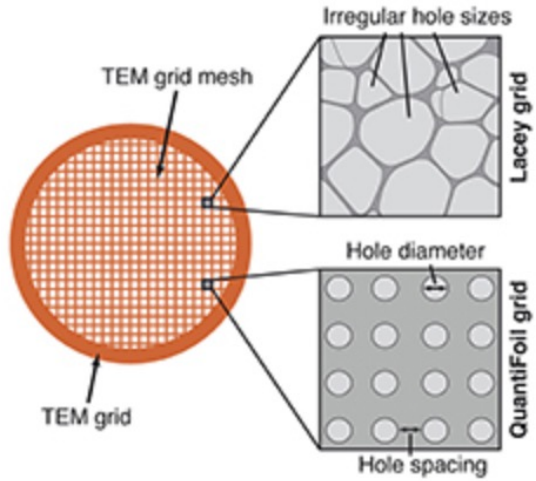
$600 \text{ nm} < \text{????????} < 10 \mu\text{m}$



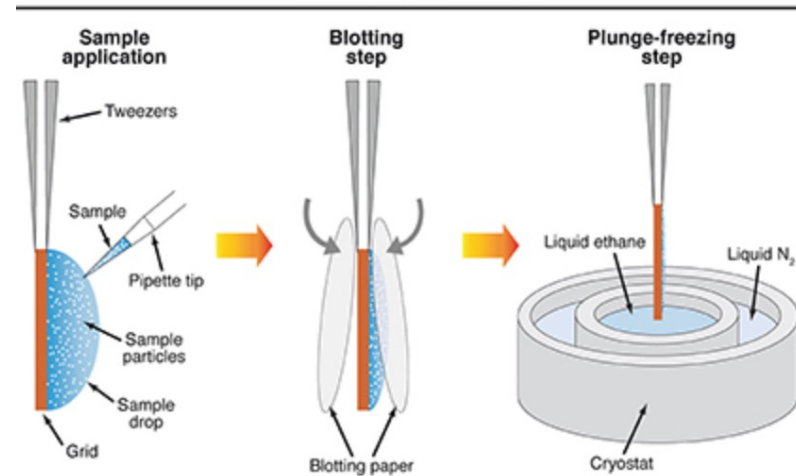
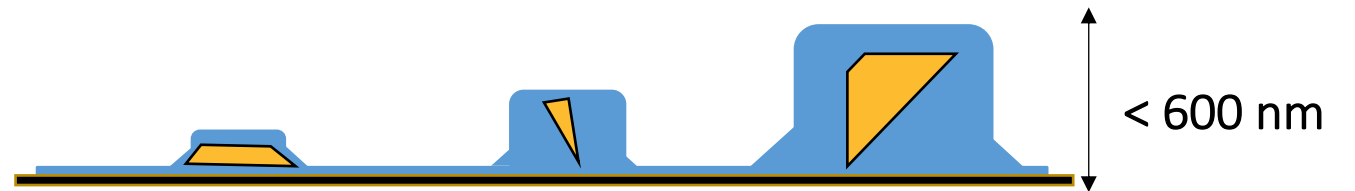
Protein crystals



Fragmentation

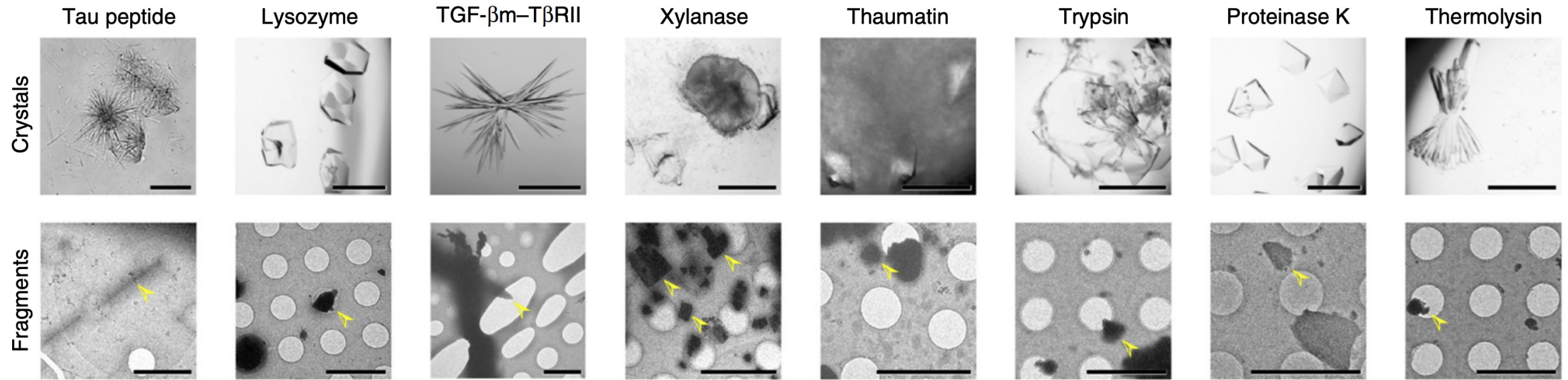


Fragmentation



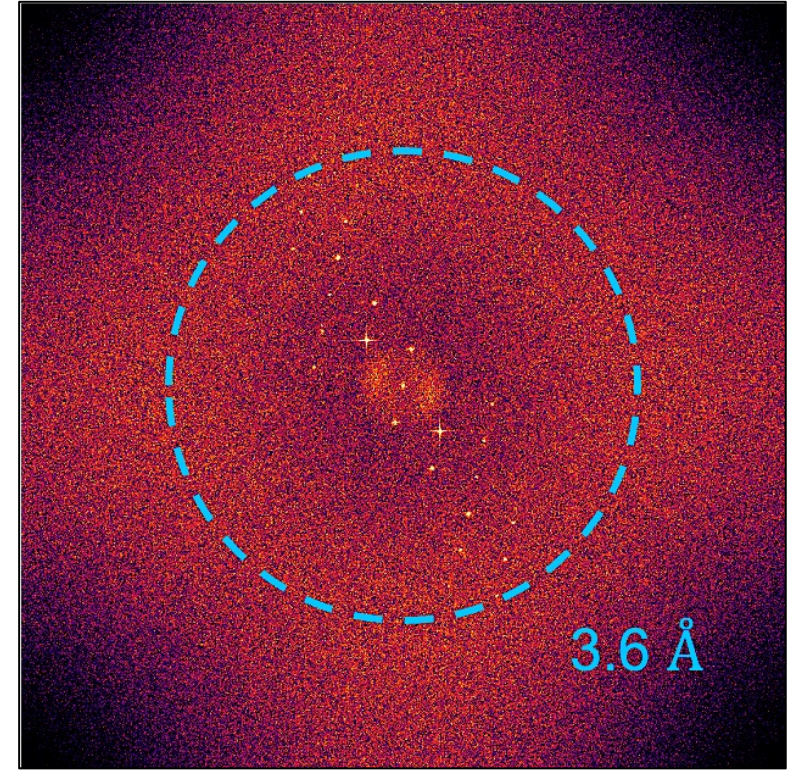
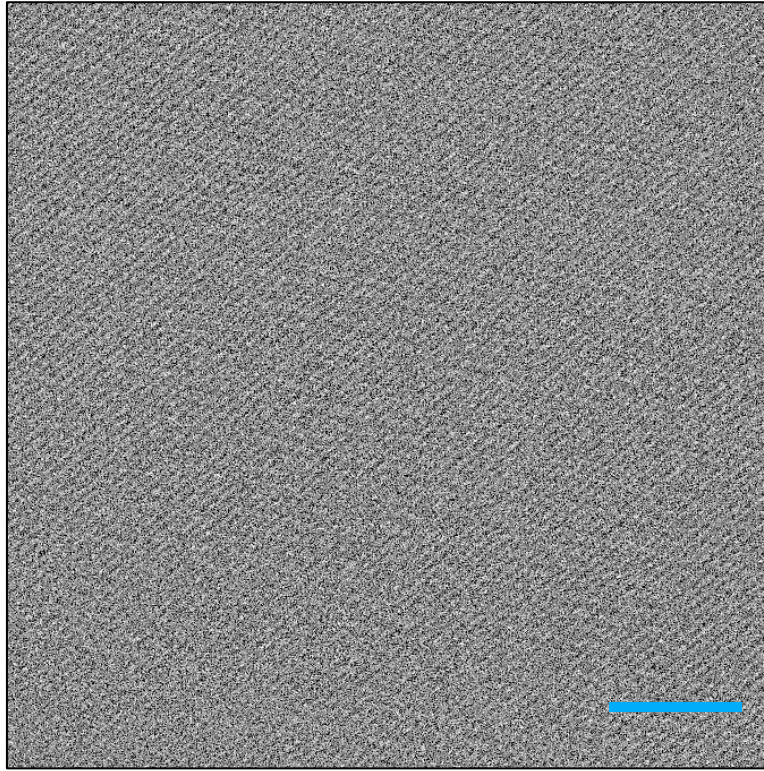
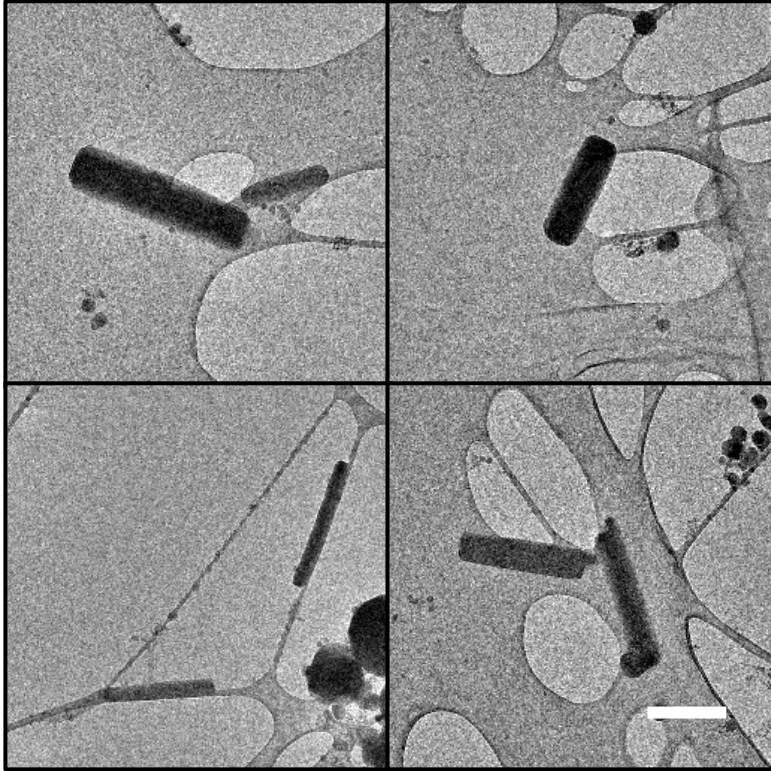
de la Cruz et al. nature methods. 2017
Sgro GG & Costa TR Frontiers in molecular biosciences. 2018

Fragmentation

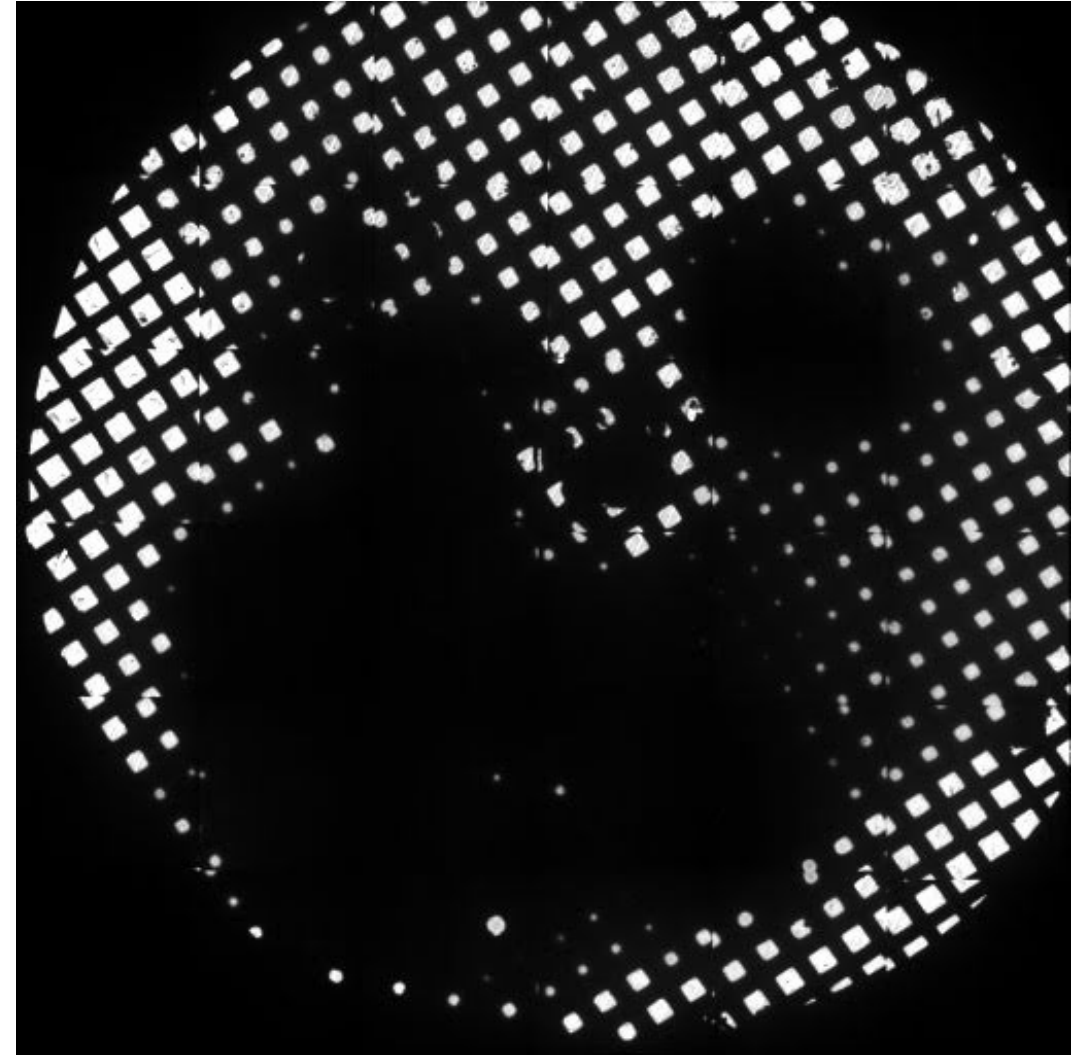
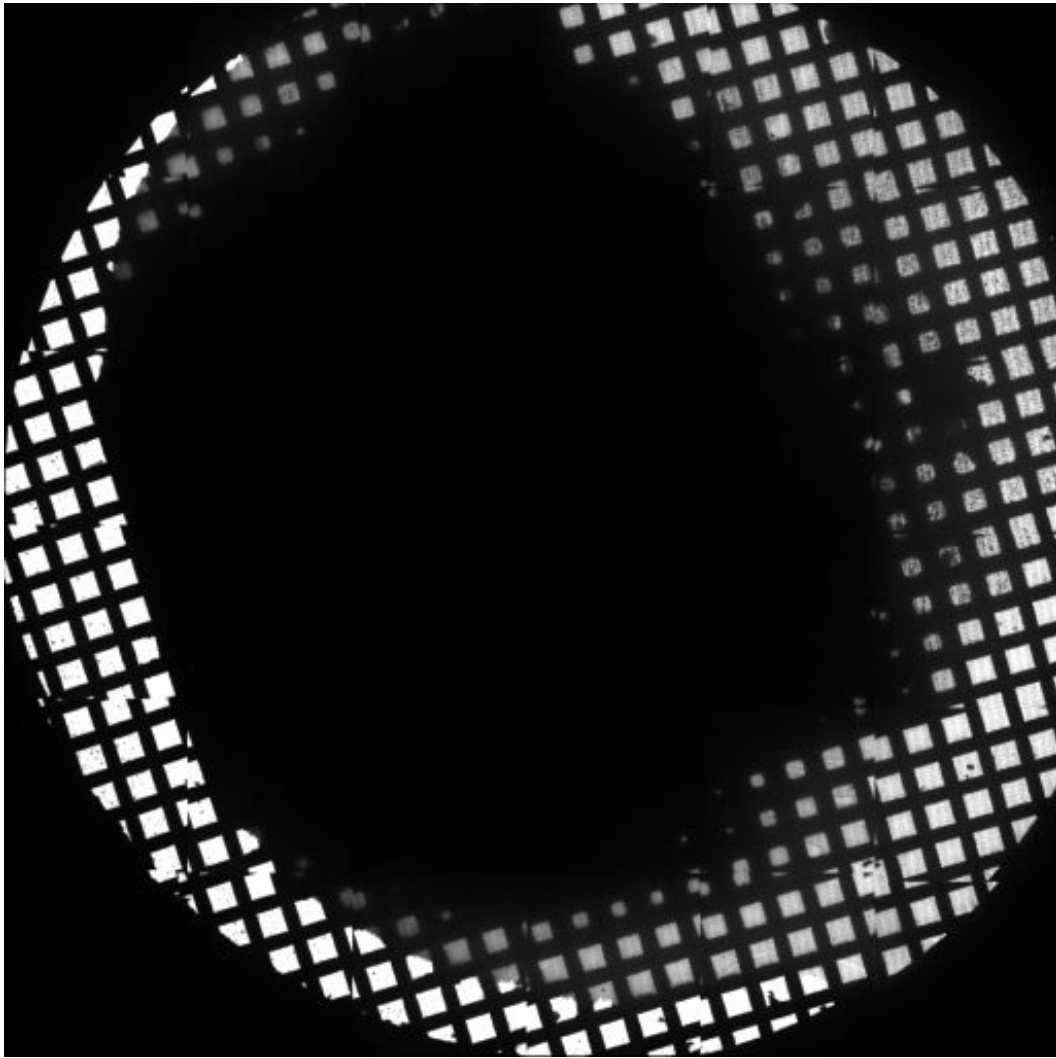


de la Cruz et al. Nature Methods 2017

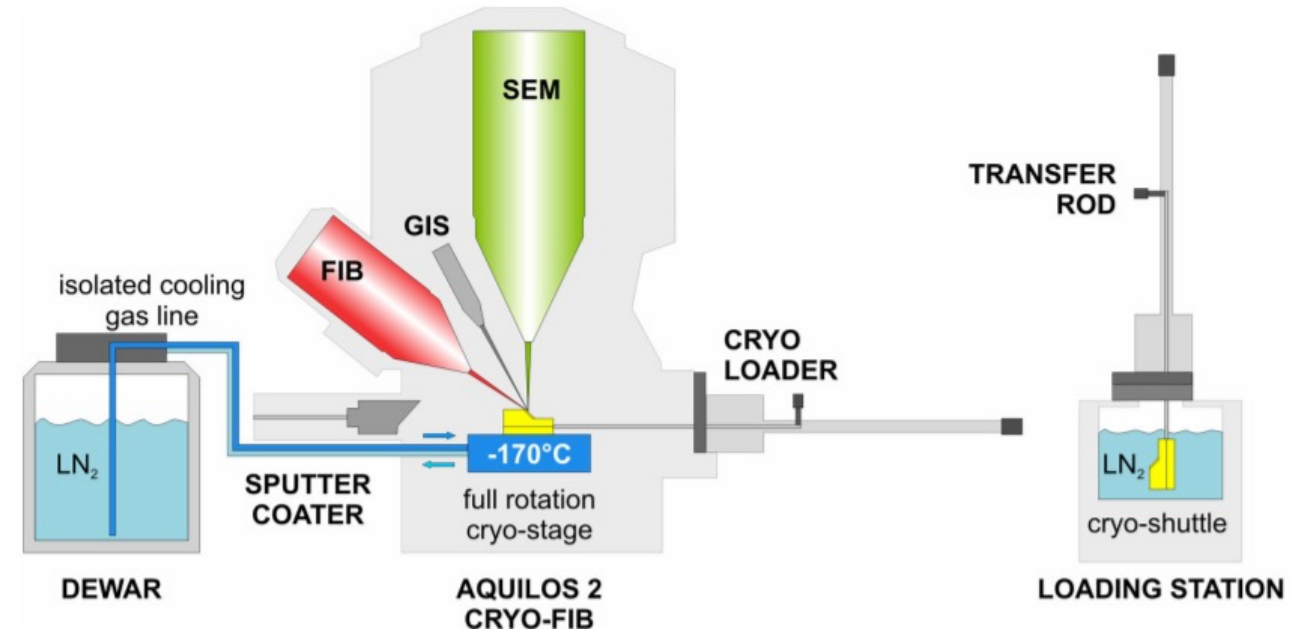
Blotting – sometimes, it works great!



Blotting – other times, the stuff of nightmares

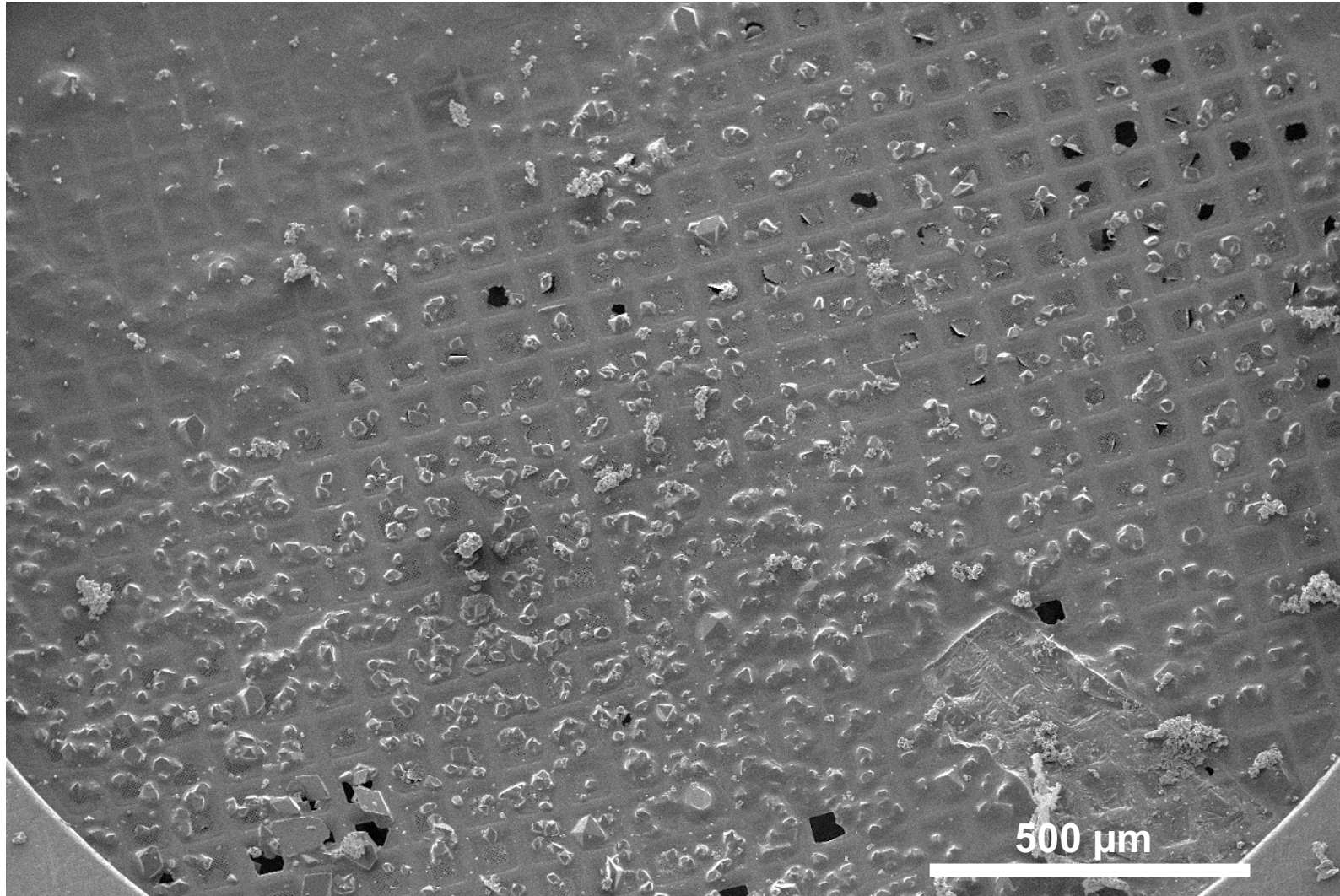


A different view – dual beam microscopes

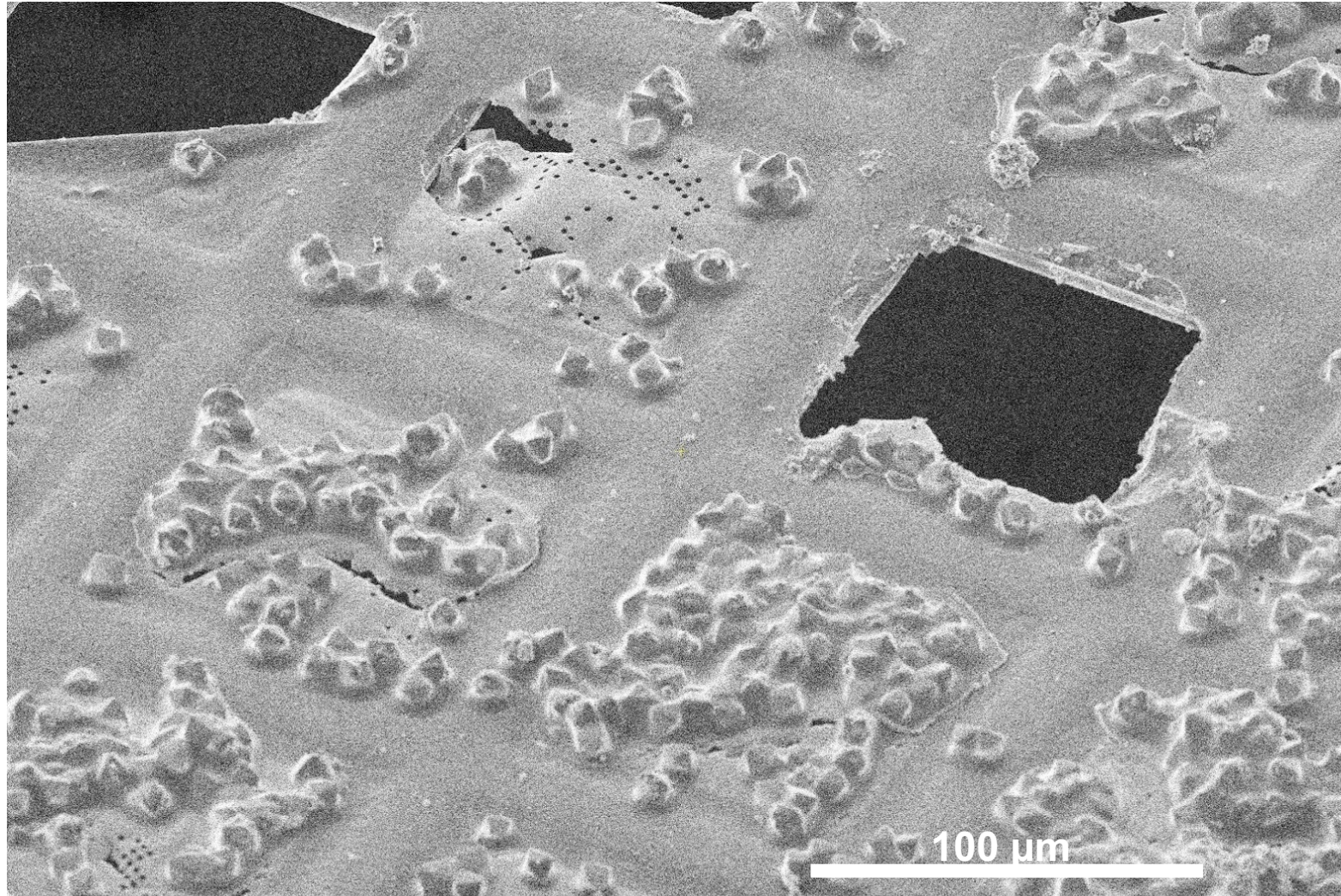


Kuba et al. Journal of Microscopy 2020

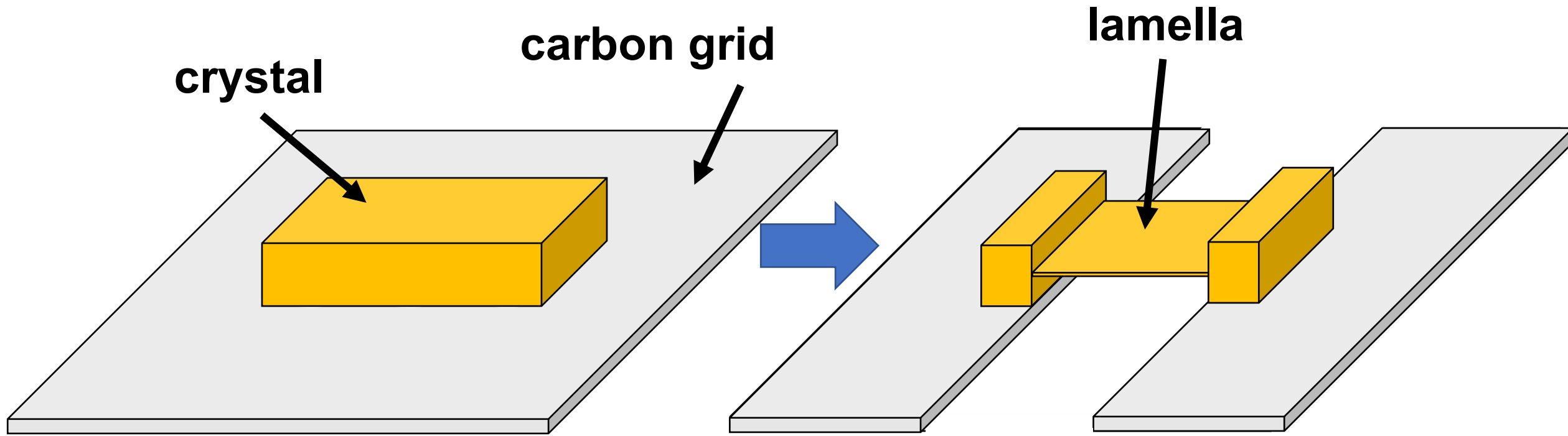
Where are the crystals?



Where are the crystals?

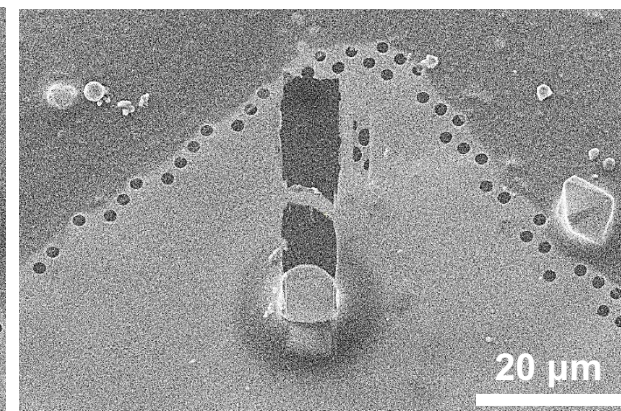
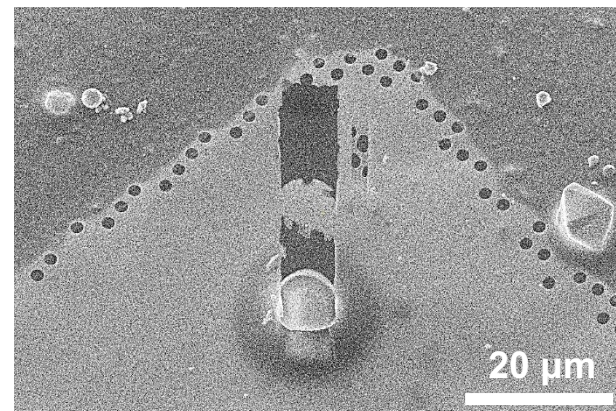
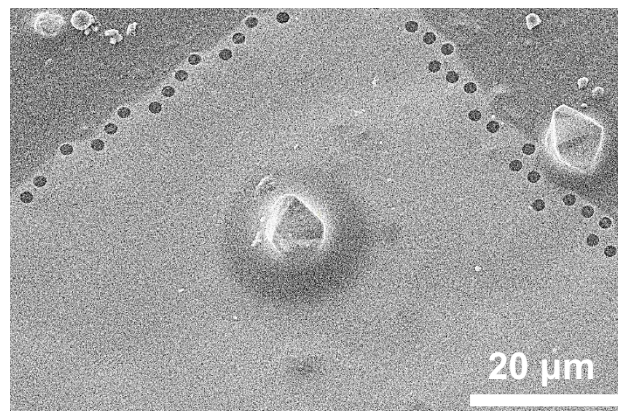
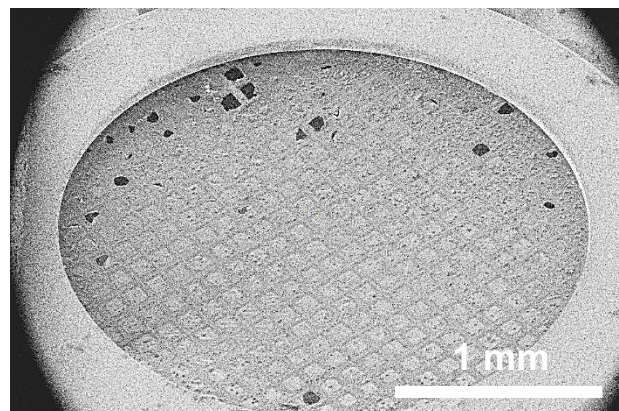


Milling – when life gives you crystal lemons

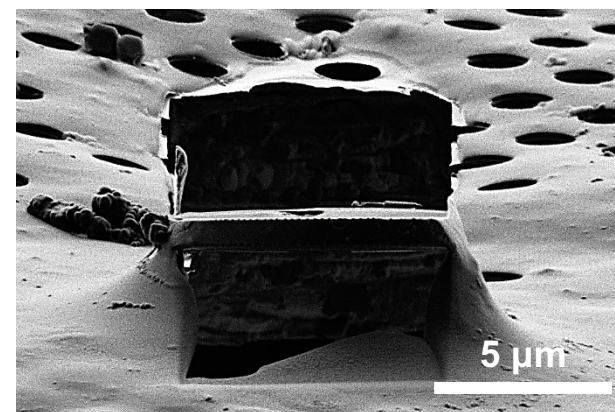
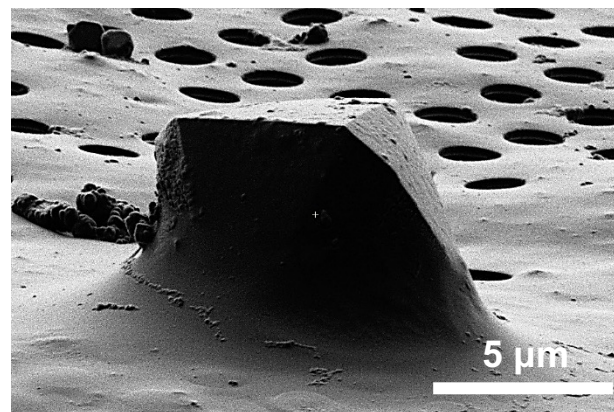
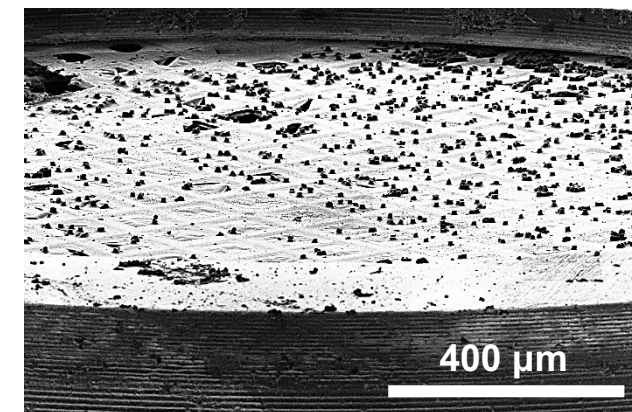


Milling protein crystals

SEM

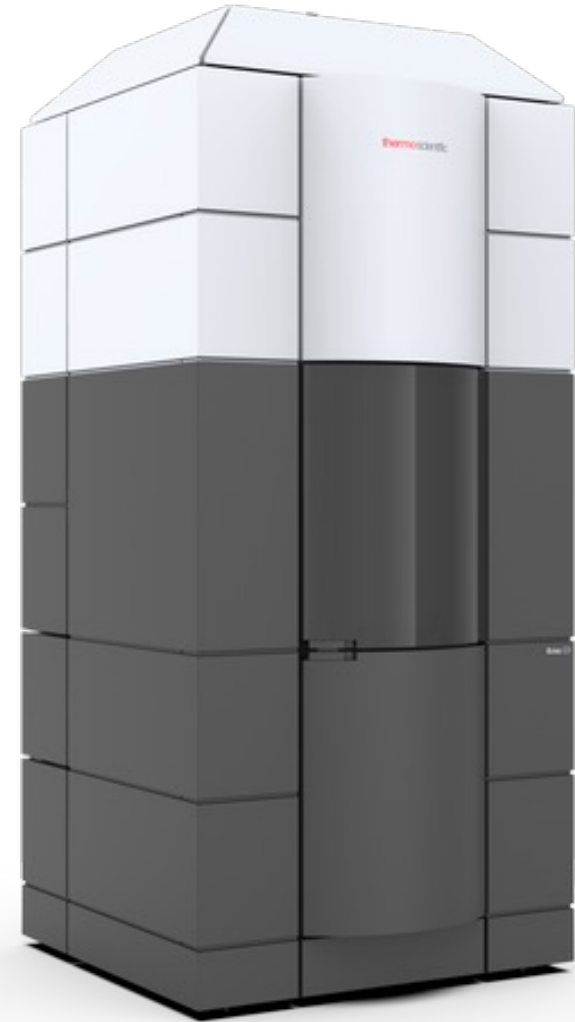
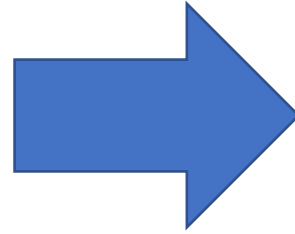


FIB

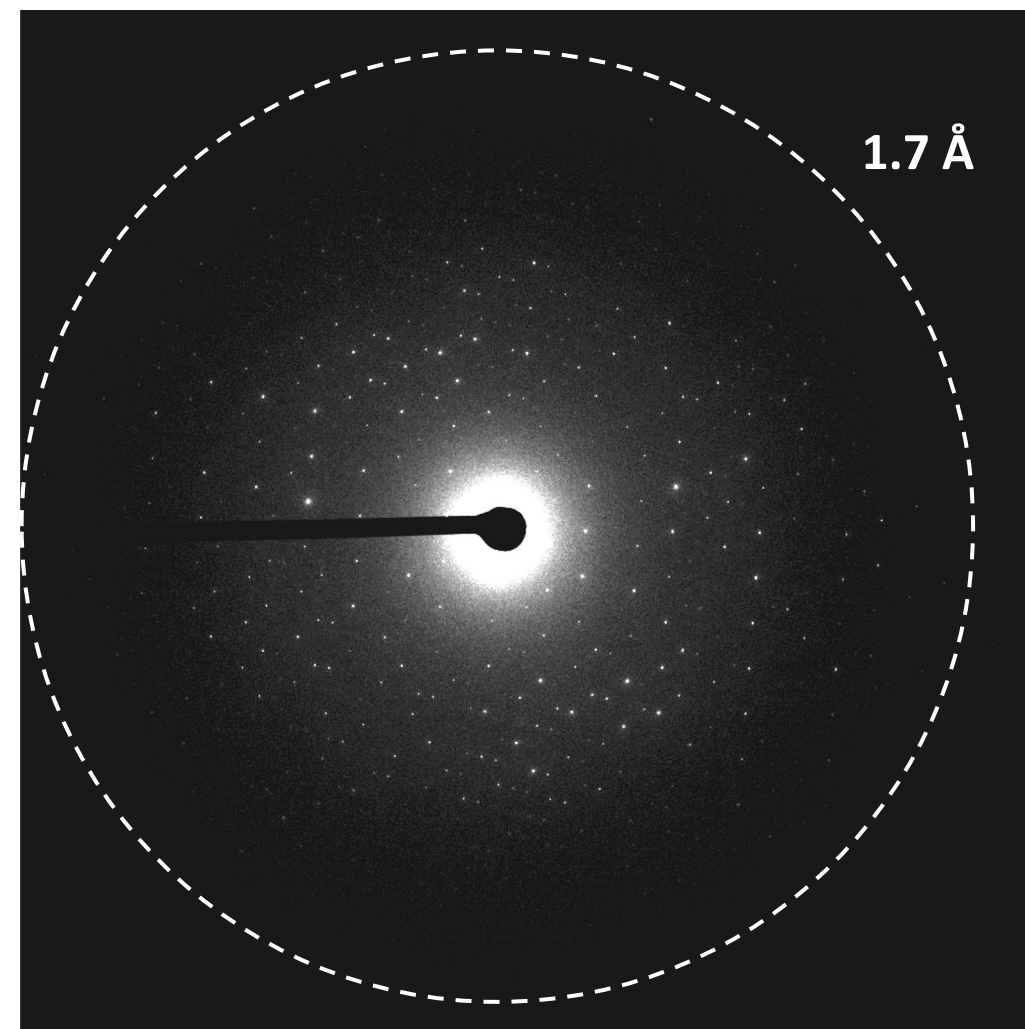
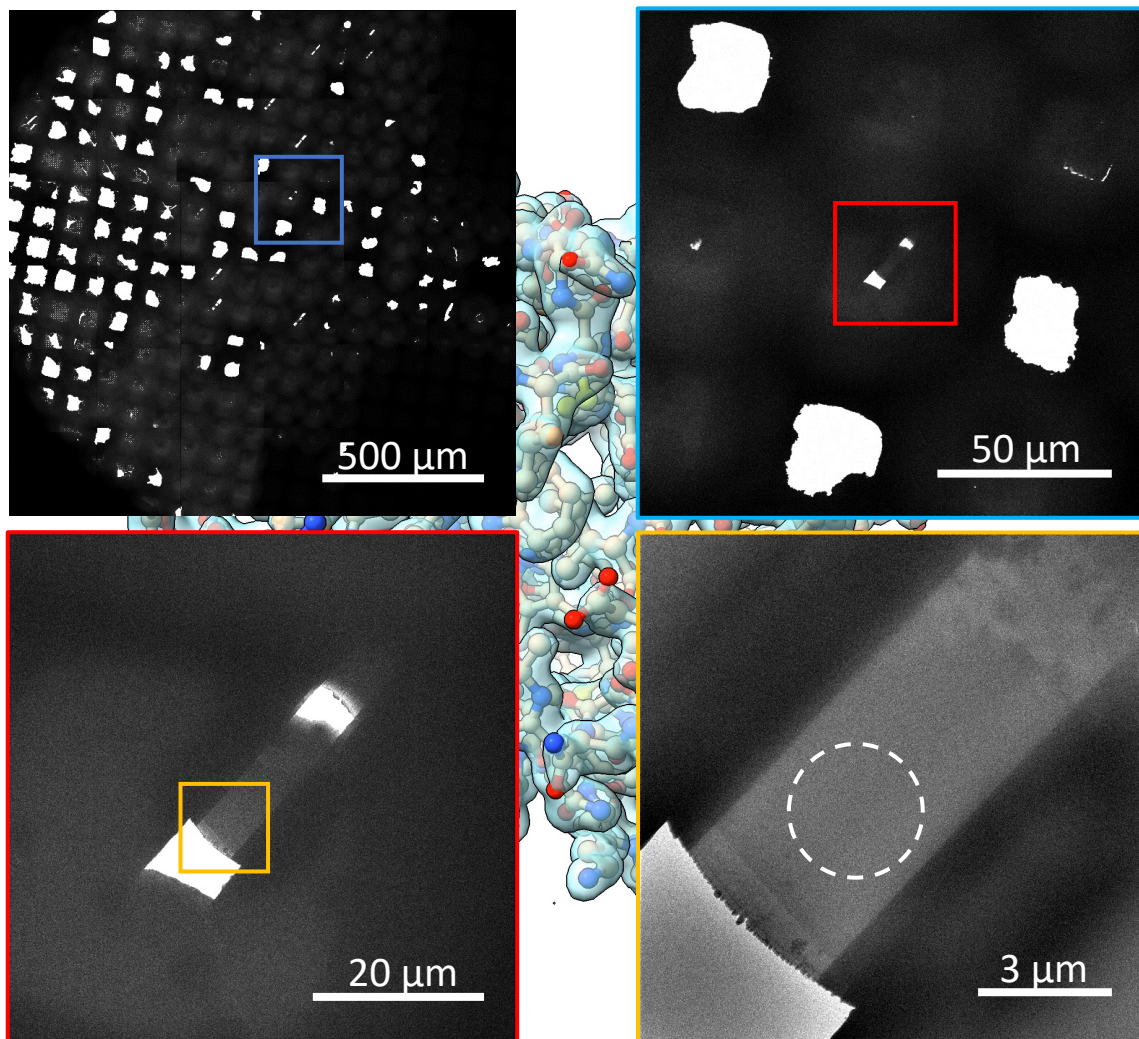


Martynowycz *et al.* Meth Mol Biol. (2020)

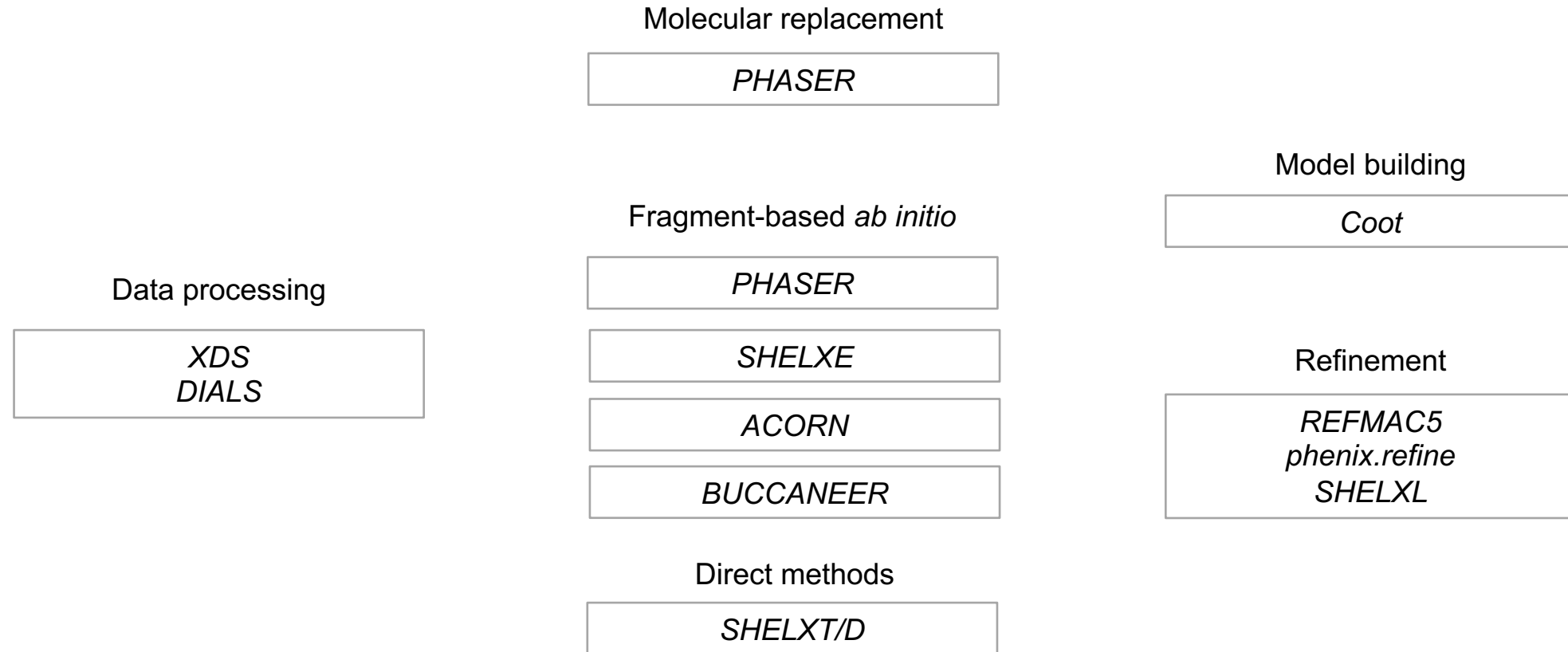
Milling protein crystals



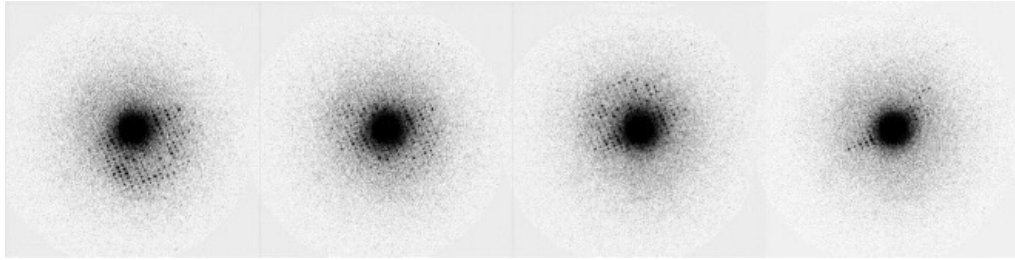
Milling – when life gives you crystal lemons



Data processing, structure solution, model building and refinement

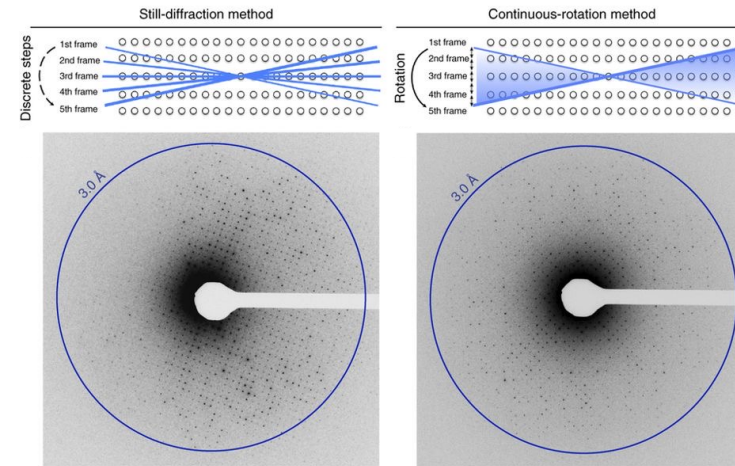


The continuous rotation method is analogous to X-ray crystallography



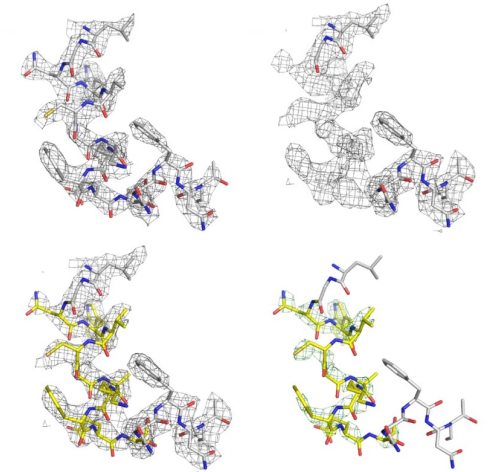
Continuous rotation electron diffraction data collection of protein nanocrystals using a hybrid pixel detector

Nederlof *et al.*, *Acta Cryst.* D69, 1223-1230 (2013)

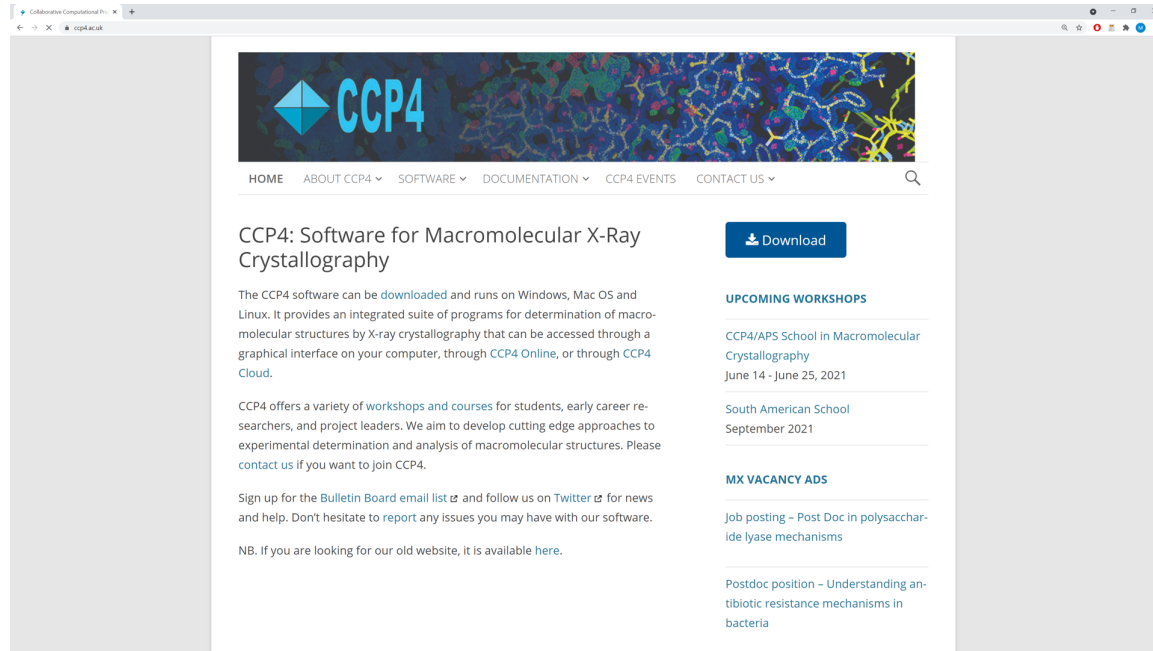


Continuous rotation MicroED data collection and structure determination from protein microcrystals

Nannenga *et al.*, *Nat. Methods* 11, 927-930 (2014)

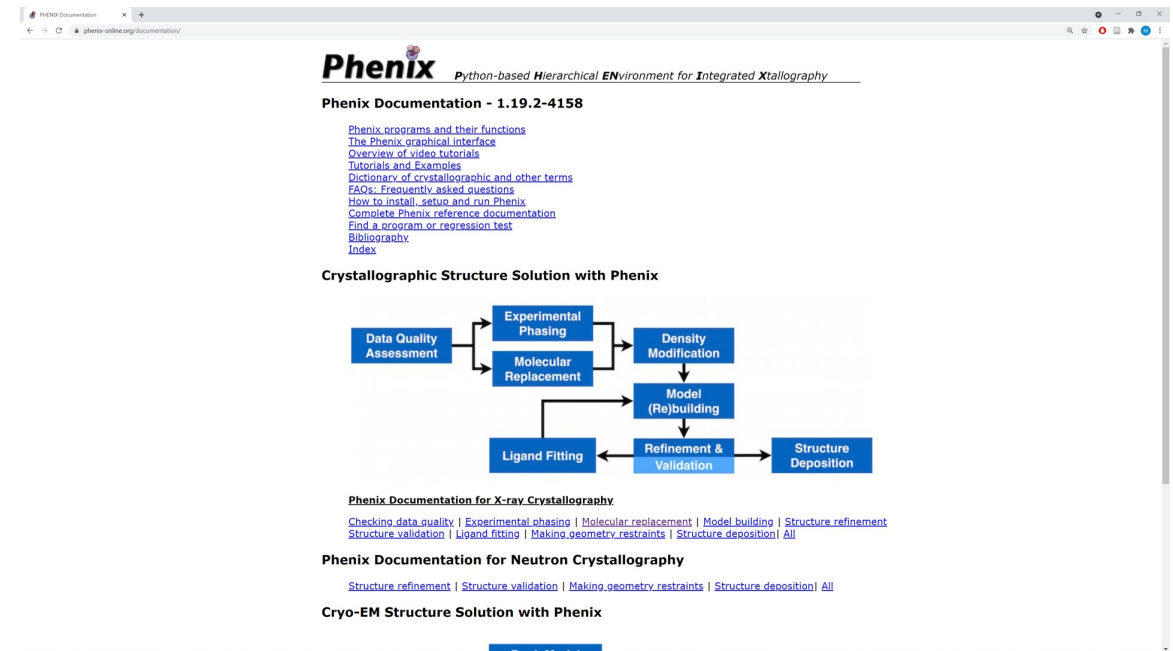


Crystallographic software are compatible with MicroED data



The screenshot shows the CCP4 website. At the top is a navigation bar with links: HOME, ABOUT CCP4, SOFTWARE, DOCUMENTATION, CCP4 EVENTS, and CONTACT US. Below the navigation bar is a large banner image with the CCP4 logo and a molecular structure. The main content area is titled "CCP4: Software for Macromolecular X-Ray Crystallography". It includes a "Download" button and a section for "UPCOMING WORKSHOPS" listing the CCP4/APS School in Macromolecular Crystallography (June 14 - June 25, 2021) and the South American School (September 2021). There is also a section for "MX VACANCY ADS" with links to job postings and a postdoc position.

CCP4 software package <https://www.ccp4.ac.uk/>
Winn *et al.*, *Acta Cryst.* D67, 235-242 (2011)



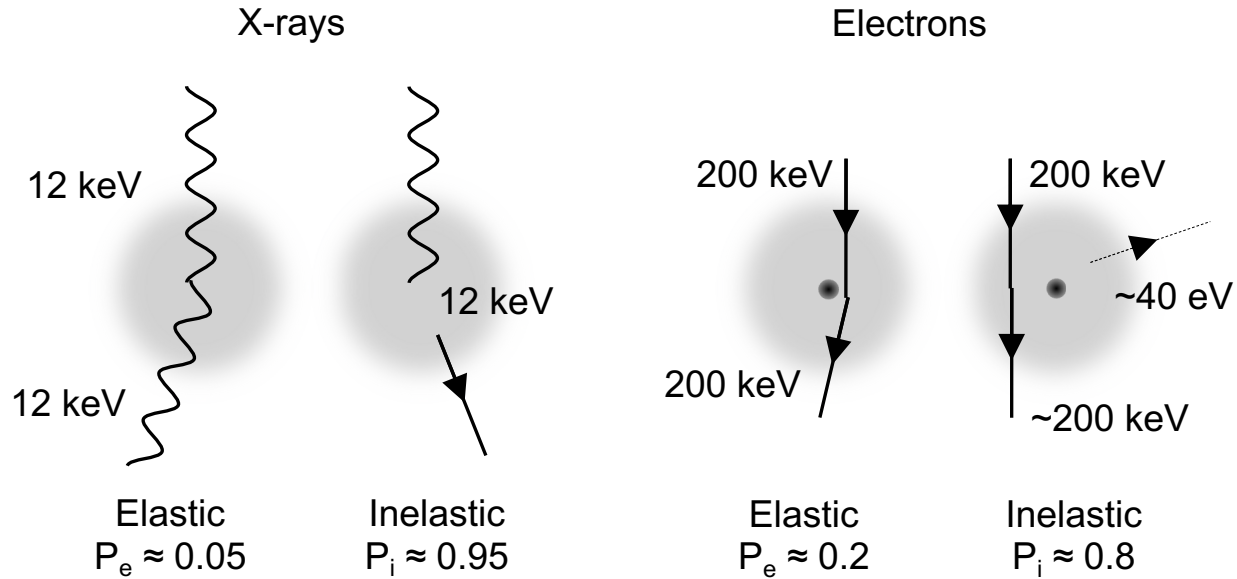
The screenshot shows the Phenix website. At the top is a navigation bar with links: HOME, ABOUT PHENIX, SOFTWARE, DOCUMENTATION, PHENIX EVENTS, and CONTACT US. Below the navigation bar is a large banner image with the Phenix logo and a molecular structure. The main content area is titled "Phenix Documentation - 1.19.2-4158". It includes a list of links for documentation, a section for "Crystallographic Structure Solution with Phenix" with a flowchart, and sections for "Phenix Documentation for X-ray Crystallography" and "Phenix Documentation for Neutron Crystallography".

```
graph LR
    A[Data Quality Assessment] --> B[Experimental Phasing]
    A --> C[Molecular Replacement]
    B --> D[Density Modification]
    C --> D
    D --> E[Model (Re)building]
    E --> F[Refinement & Validation]
    F --> G[Structure Deposition]
    F --> H[Ligand Fitting]
```

Phenix software suite <https://phenix-online.org/>
Adams *et al.*, *Acta Cryst.* D58, 1948-1954 (2002)

Electron and X-ray diffraction

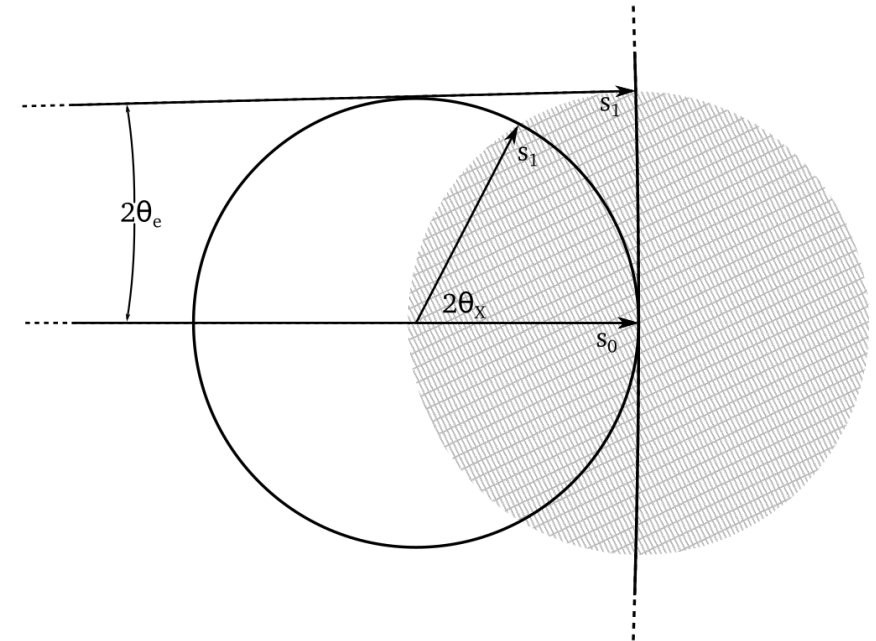
Electrons interact with the electrostatic potential



± 200 keV collateral damage
per useful diffracted quantum

± 200 eV collateral damage
per useful diffracted quantum

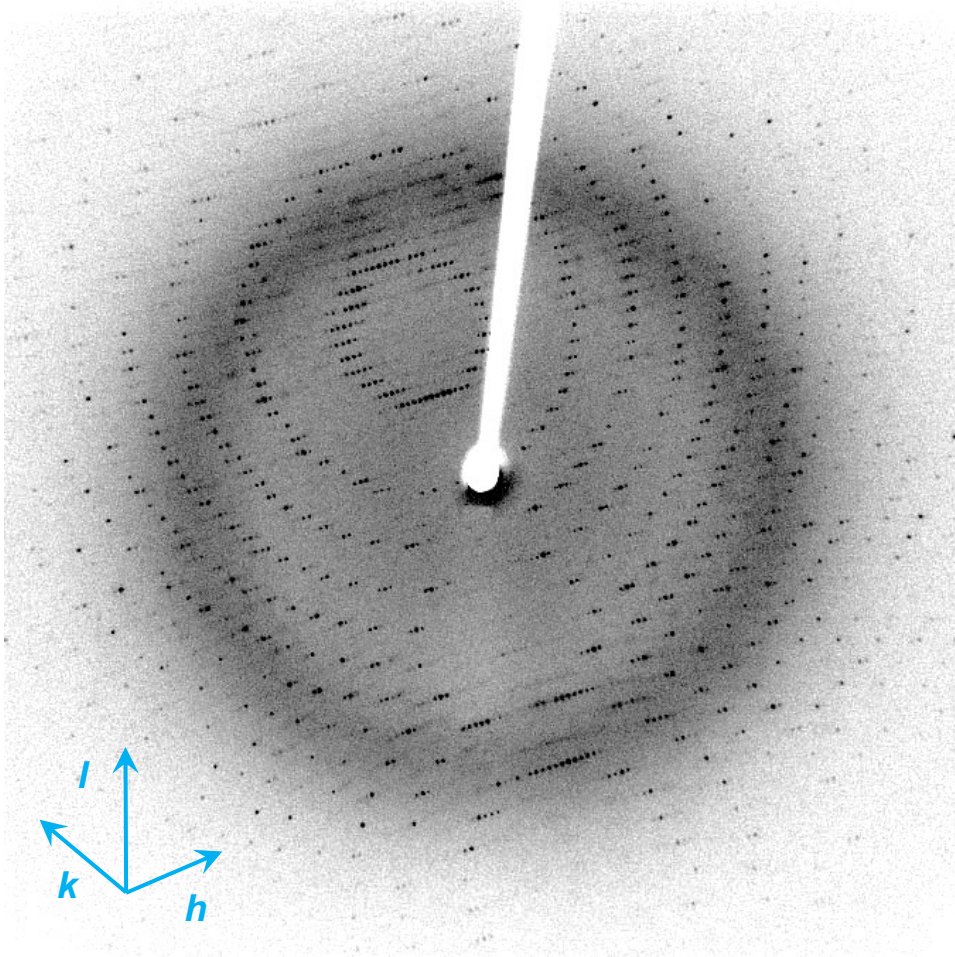
Diffraction geometry and Ewald construction



The Ewald construction is virtually flat for electron diffraction

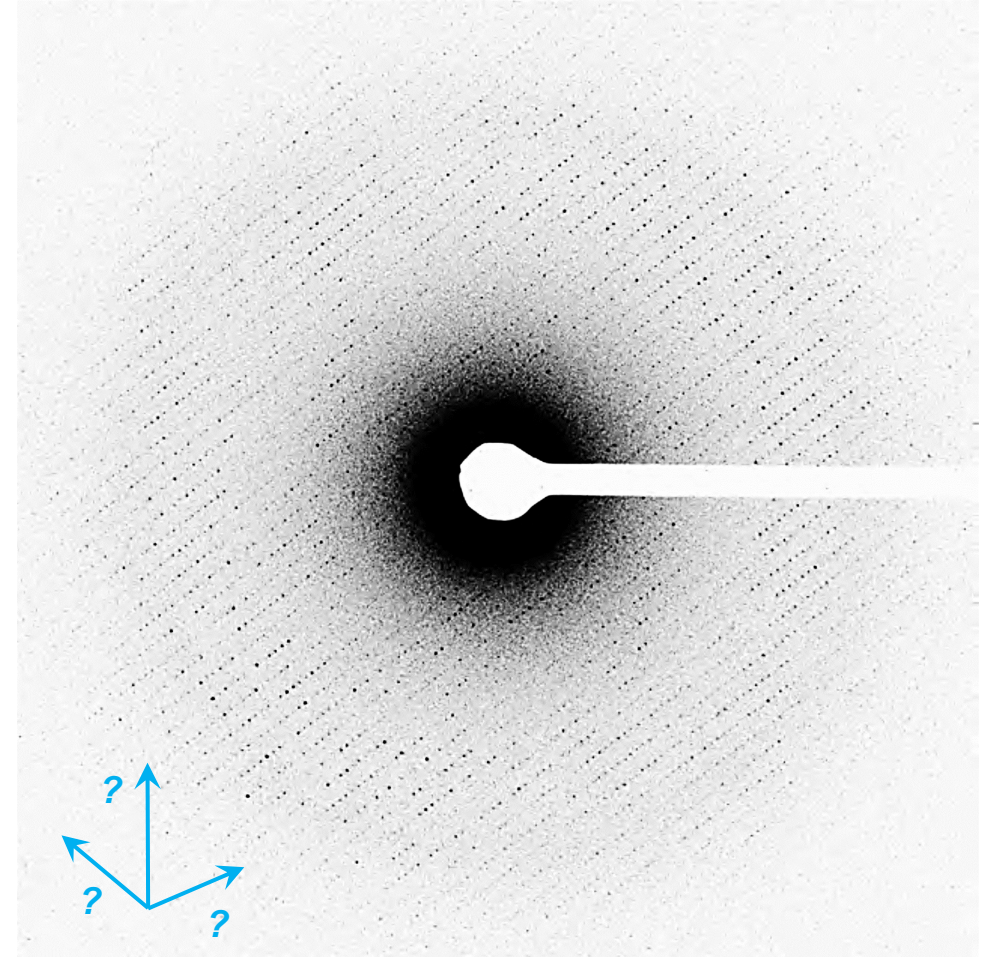
Indexing X-ray and electron diffraction patterns

X-ray diffraction



12 keV X-ray photons, $\lambda = 1.0332 \text{ \AA}$
Scattering angle $2\theta_x \approx 62.2^\circ$ at $1 \text{ \AA}$

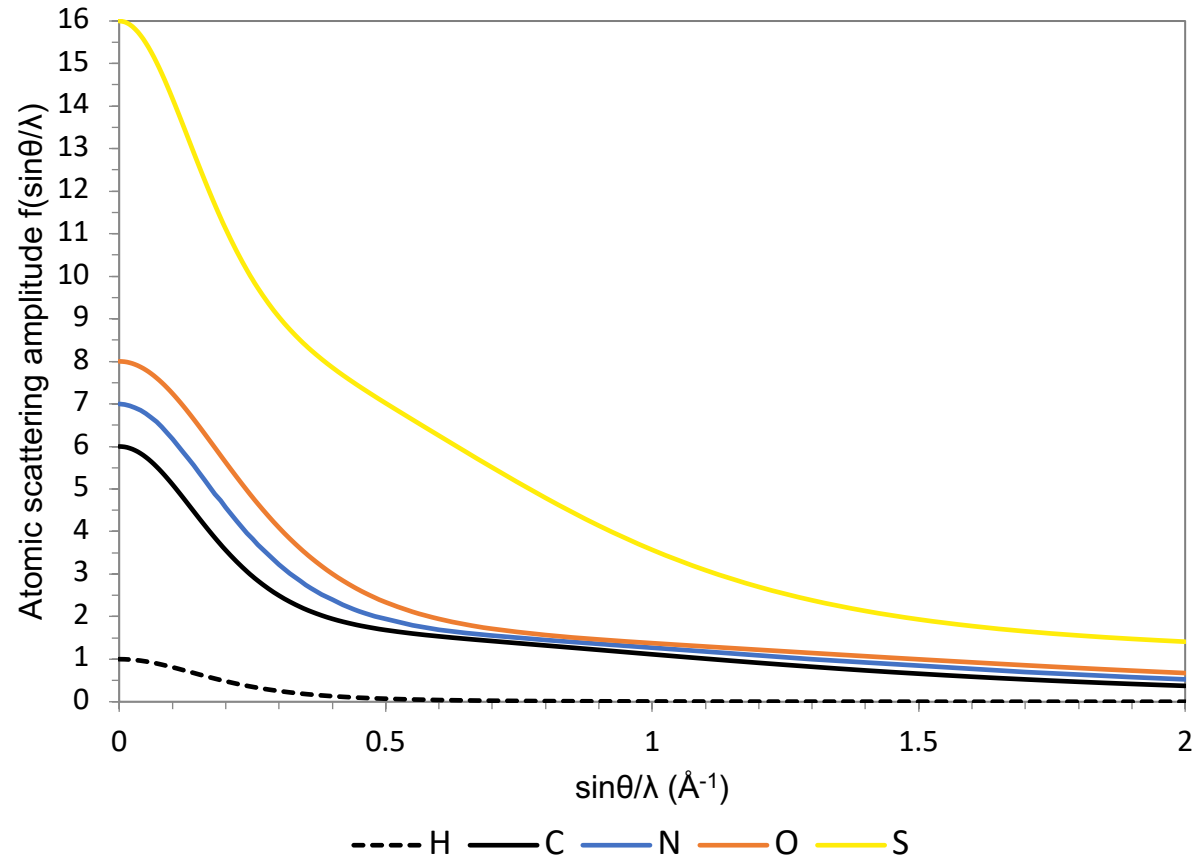
Electron diffraction



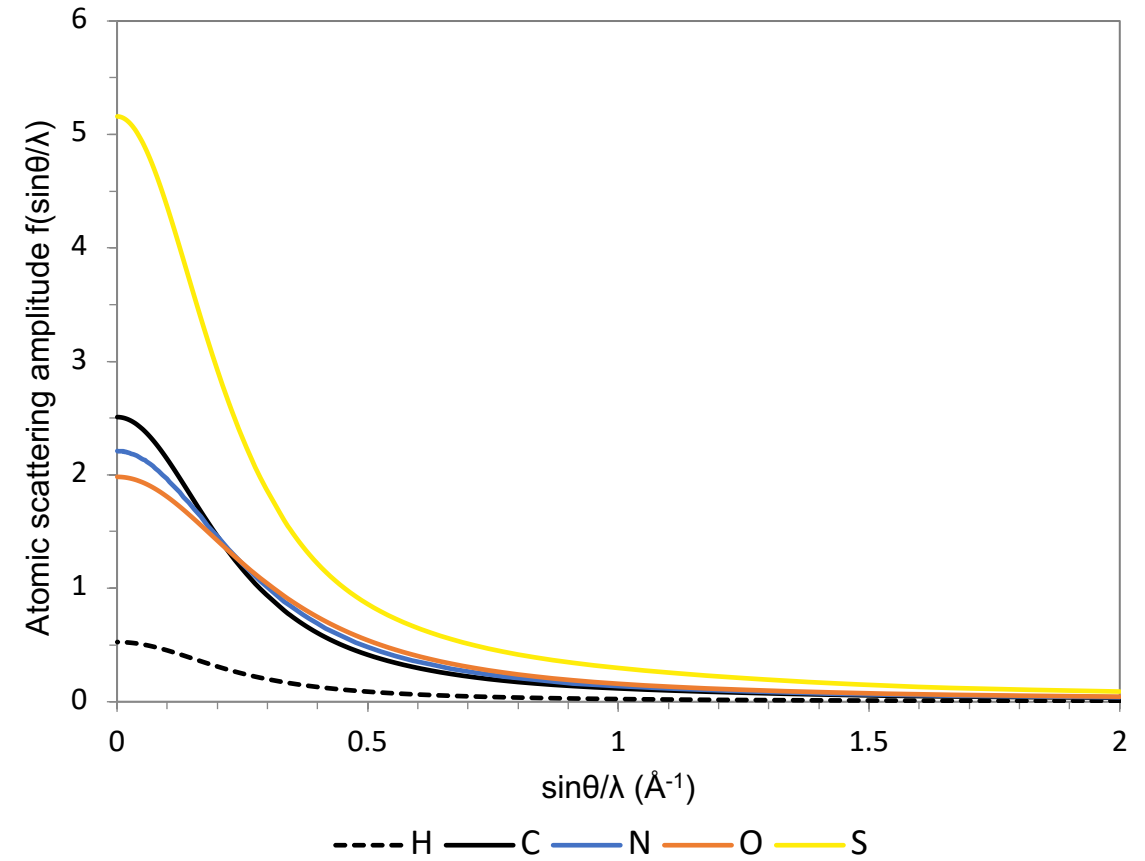
200 kV electrons, $\lambda = 0.0251 \text{ \AA}$
Scattering angle $2\theta_e \approx 1.4^\circ$ at $1 \text{ \AA}$

Atomic scattering factors

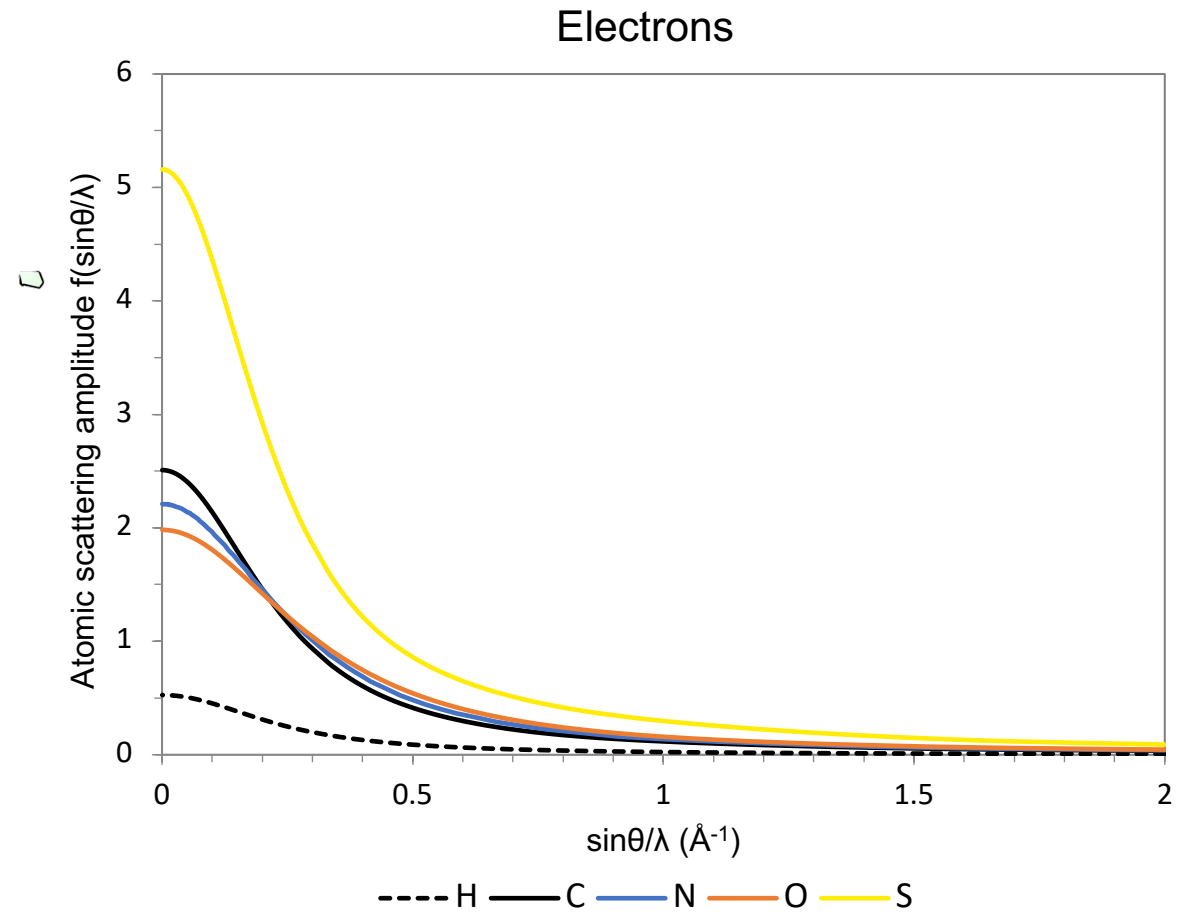
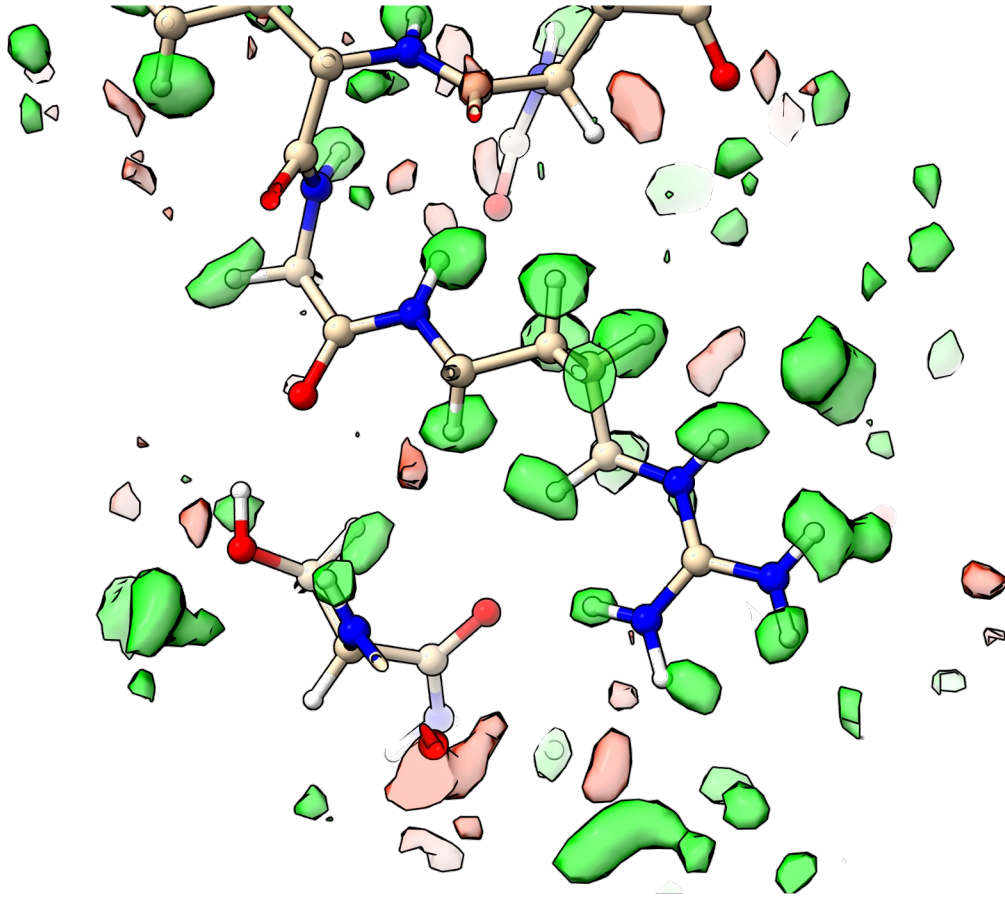
X-rays



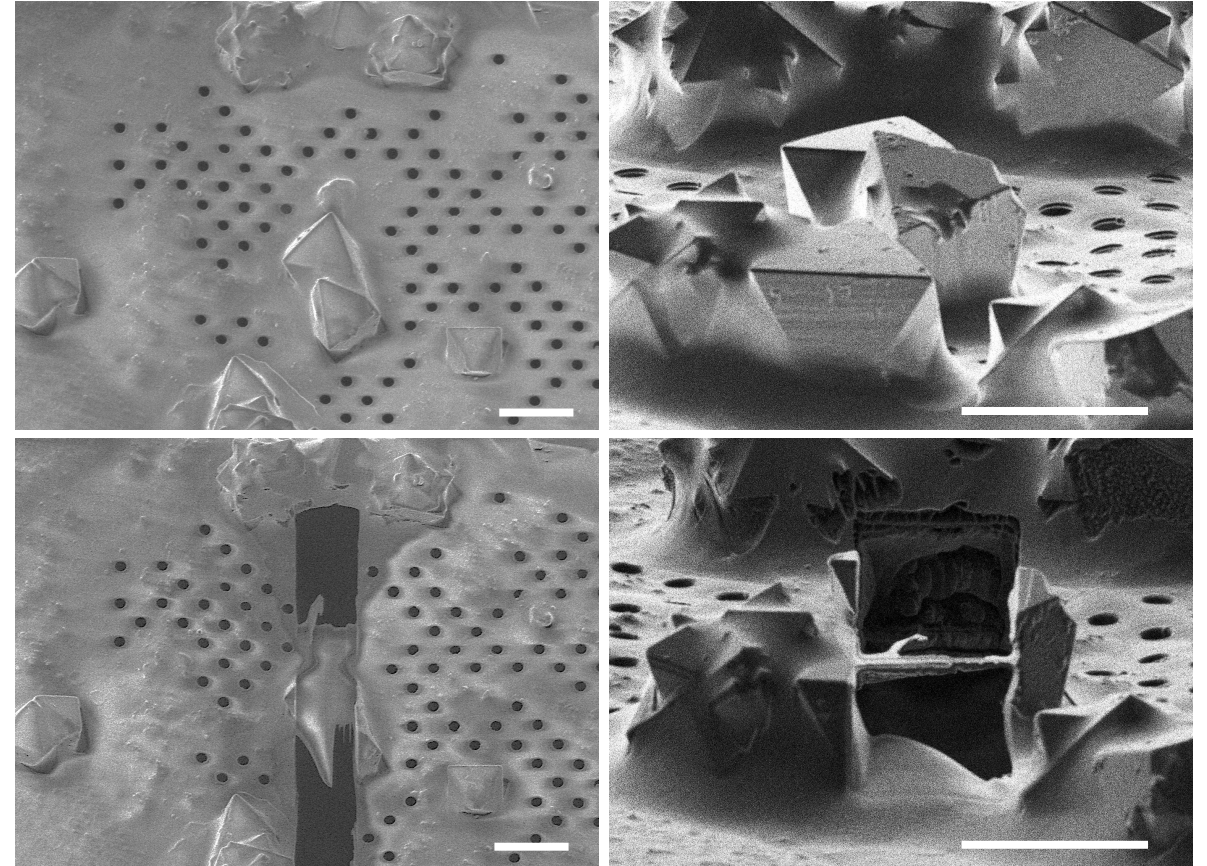
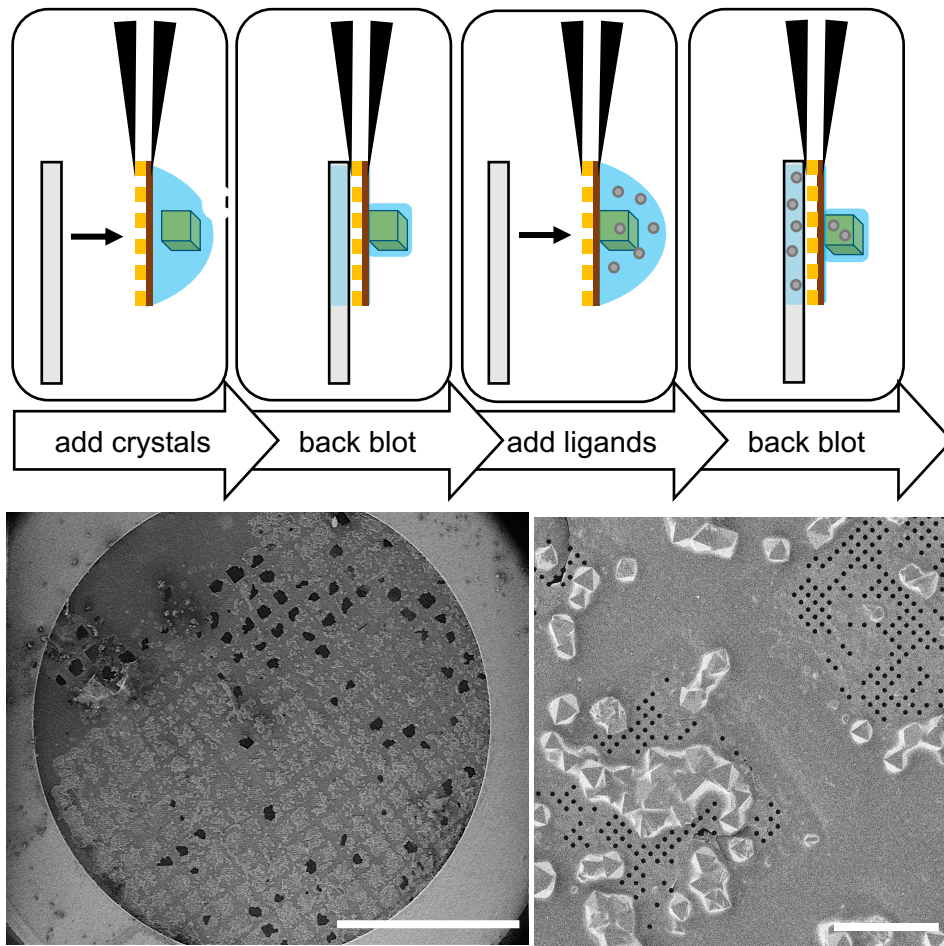
Electrons

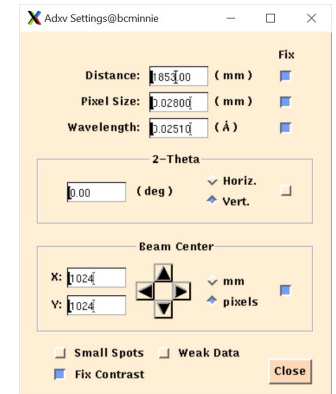
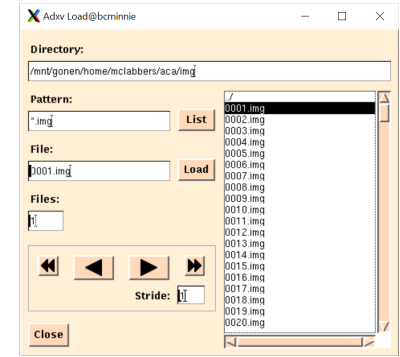
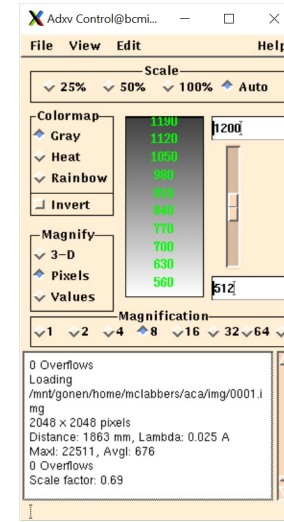
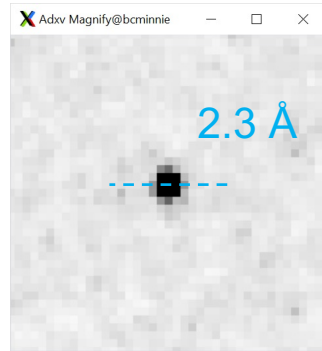


Atomic scattering factors



Example 1: Efficient ligand incorporation into microcrystals of proteinase K by on-grid soaking





<https://www.scripps.edu/tainer/arvai/adxv.html>

Data processing software

XDS

DIALS

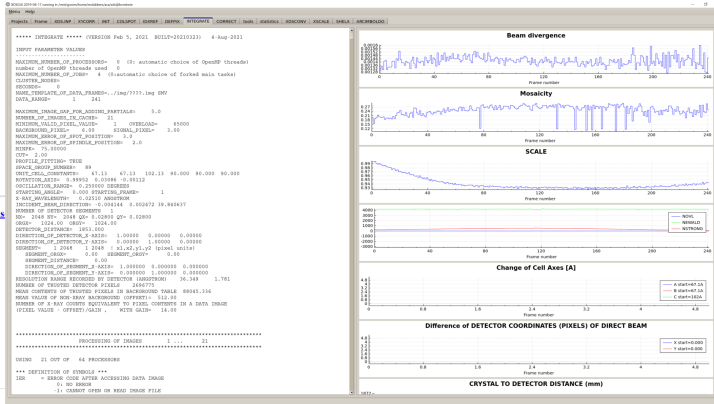
XDS Program Package

for processing single-crystal monochromatic diffraction data recorded by the rotation method.

Version: Feb 5, 2021 [Release Notes](#)
Author: Wolfgang Kabsch

- [Downloading and Installation](#)
- [Getting started](#)
- [Ads](#)
 - [Input parameters](#)
 - [Files](#)
- [Scale](#)
 - [Input parameters](#)
 - [Files](#)
- [xdsconv](#)
 - [Input Parameters](#)
 - [Output formats](#)
- [2cbf](#)
- [microCbf](#)
- [CellScan](#)
- [PixLab](#)
- [References](#)

© 2005-2021, MPI for Medical Research, Heidelberg [Imaging Data](#)
<http://www.kabsch@mpi-czr.uni-heidelberg.mpg.de>
page last updated: Feb 5, 2021



DIALS
Diffraction Integration for Advanced Light Sources

Home
About
Installation
Documentation
Tutorials
How to
Knowledge Base
Workshops
Publications
License

DIALS: Diffraction Integration for Advanced Light Sources

X-ray crystallography for structural biology has benefited greatly from a number of advances in recent years including high performance pixel array detectors, new beamlines capable of delivering micro- and sub-micron focus and new light sources such as XFELs. The DIALS project is a collaborative endeavour to develop new diffraction integration software to meet the data analysis requirements presented by these recent advances. There are three end goals: to develop an extensible framework for the development of algorithms to analyse X-ray diffraction data; the implementation of algorithms within this framework and finally a set of user facing tools using these algorithms to allow integration of data from diffraction experiments on synchrotron and free electron sources.

Contact

For feature requests, bug reports or any other information, please contact the DIALS developers at dials-report@lists.sourceforge.net, or post an issue.

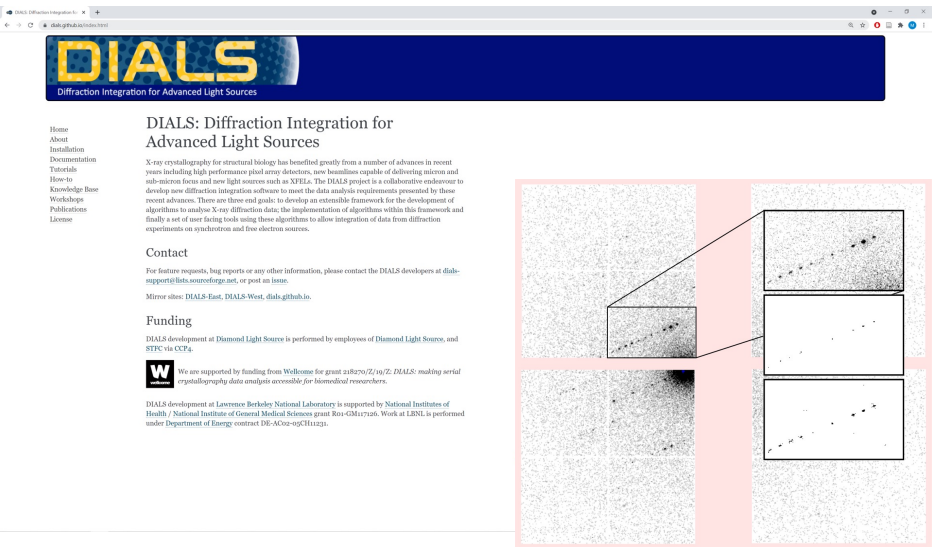
Mirror sites: DIALS-East, DIALS-West, dials.github.io.

Funding

DIALS development at Diamond Light Source is performed by employees of Diamond Light Source, and ESRF, via CITA.

W We are supported by funding from Wellcome for grant 218790/Z/19/Z. DIALS: making serial crystallography data analysis accessible for biomedical researchers.

DIALS development at Lawrence Berkeley National Laboratory is supported by National Institutes of Health / National Institute of General Medical Sciences grant R01GM127366. Work at LBNL is performed under Department of Energy contract DE-AC02-05CH11432.



Kabsch, *Acta Cryst.* D66, 125-132 (2010)

<https://xds.mpimf-heidelberg.mpg.de>

<https://strucbio.biologie.uni-konstanz.de/xdswiki>

<https://homepage.univie.ac.at/tim.gruene/research/seminars/index>

Winter *et al.*, *Acta Cryst.* D74, 85-97 (2018)
Clabbers *et al.*, *Acta Cryst.* D74, 506-518 (2018)

<https://dials.github.io/index.html>

https://dials.github.io/documentation/tutorials/dials_for_ed.html

Setting up data processing in *XDS*

```
JOB=XYCORR INIT COLSPOT IDXREF DEFPIX INTEGRATE CORRECT
```

Default in *XDS* is to run all steps including spot finding (COLSPOT), indexing (IDXREF), data integration (INTEGRATE), refinement and scaling (CORRECT)

```
DETECTOR=ADSC OFFSET=512 GAIN=14  
DIRECTION_OF_DETECTOR_X-AXIS=1.0 0.0 0.0  
DIRECTION_OF_DETECTOR_Y-AXIS=0.0 1.0 0.0  
NX=2048 NY=2048 QX=0.028 QY=0.028  
TRUSTED_REGION=0.0 1.41
```

Parameters describing the detector setup, offset for negative pixel values, and detector gain

```
ORGX=1024 ORGY=1024  
DETECTOR_DISTANCE=1853
```

```
ROTATION_AXIS=1.0 0.0 0.0  
OSCILLATION_RANGE=0.25  
X-RAY_WAVELENGTH=0.0251
```

Rotation axis and angular increment in deg/frame during MicroED data collection

```
REFINE(IDXREF)=BEAM AXIS ORIENTATION CELL !POSITION  
REFINE(INTEGRATE)=!POSITION BEAM ORIENTATION !CELL AXIS  
REFINE(CORRECT)=BEAM AXIS ORIENTATION CELL !POSITION
```

Parameter refinement during indexing (IDXREF), integration (INTEGRATE) and refinement (CORRECT), detector distance (POSITION) should preferably be decoupled from lattice refinement (CELL)

```
MINIMUM_I/SIGMA=1.0
```

```
STRONG_PIXEL=6.0  
MINIMUM_NUMBER_OF_PIXELS_IN_A_SPOT=6
```

Spot finding parameters (COLSPOT) for pixels to be included that exceed the background by at least 6 standard deviations, and have at least 6 pixels to be considered as a spot

Refinement of the experimental and lattice parameters

```

mv GXPARM.XDS XPARAM.XDS

JOB=DEFPIX INTEGRATE CORRECT

DETECTOR=ADSC OFFSET=512 GAIN=14
DIRECTION_OF_DETECTOR_X-AXIS=1.0 0.0 0.0
DIRECTION_OF_DETECTOR_Y-AXIS=0.0 1.0 0.0
NX=2048 NY=2048 QX=0.028 QY=0.028
TRUSTED_REGION=0.0 1.41

ORGX=1024 ORGY=1024
DETECTOR_DISTANCE=1853

ROTATION_AXIS=1.0 0.0 0.0
OSCILLATION_RANGE=0.25
X-RAY_WAVELENGTH=0.0251
    
```

```

REFINE(IDXREF)=BEAM AXIS ORIENTATION CELL !POSITION
REFINE(INTEGRATE)=POSITION BEAM ORIENTATION !CELL AXIS
REFINE(CORRECT)=BEAM AXIS ORIENTATION CELL !POSITION
    
```

```
MINIMUM_I/SIGMA=1.0
```

```
STRONG_PIXEL=6.0
```

```
MINIMUM_NUMBER_OF_PIXELS_IN_A_SPOT=6
```

```

BEAM_DIVERGENCE=    0.01486  BEAM_DIVERGENCE_E.S.D.=    0.00149
REFLECTING_RANGE=   1.29714  REFLECTING_RANGE_E.S.D.=    0.18531
    
```

```

XPARAM.XDS  VERSION Feb 5, 2021  BUILT=20210323
  1          0.0000    0.2500  0.999529  0.030622 -0.001849
    0.025100    -0.004082    0.002496    39.840637
89    67.1354    67.1354   102.0969  90.000   90.000   90.000
 -54.352264    33.359516   -20.979265
-25.849874    -3.203901    61.876282
  45.235233    88.466454    23.478504
    1          2048    2048    0.028000    0.028000
1024.000000   1024.000000   1853.000000
  1.000000    0.000000    0.000000
  0.000000    1.000000    0.000000
  0.000000    0.000000    1.000000
    1          1          2048    1          2048
    
```

A key feature of data processing with XDS is iterative recycling of the refined experimental and lattice parameters (XPARAM.XDS) suggested during integration and refinement

Estimated reflecting range/rocking curve for a reflection in degrees

Estimated mosaicity of the crystal in degrees



Refined unit cell parameters and intensity statistics

REFINED VALUES OF DIFFRACTION PARAMETERS DERIVED FROM 30765 INDEXED SPOTS

REFINED PARAMETERS: AXIS BEAM ORIENTATION CELL

STANDARD DEVIATION OF SPOT POSITION (PIXELS) 1.84

STANDARD DEVIATION OF SPINDLE POSITION (DEGREES) 0.11

Data are integrated in Laue group 4/mmm

SPACE GROUP NUMBER 89

UNIT CELL PARAMETERS 67.133 67.133 102.087 90.000 90.000 90.000

E.S.D. OF CELL PARAMETERS 1.2E-01 1.2E-01 6.5E-02 0.0E+00 0.0E+00 0.0E+00

REC. CELL PARAMETERS 0.014896 0.014896 0.009796 90.000 90.000 90.000

COORDINATES OF UNIT CELL A-AXIS 25.847 3.201 -61.875

COORDINATES OF UNIT CELL B-AXIS 54.348 -33.364 20.977

COORDINATES OF UNIT CELL C-AXIS -45.241 -88.453 -23.474

CRYSTAL MOSAICITY (DEGREES) 0.180

LAB COORDINATES OF ROTATION AXIS 0.999534 0.030477 -0.001864

DIRECT BEAM COORDINATES (REC. ANGSTROM) -0.004092 0.002480 39.840637

DETECTOR COORDINATES (PIXELS) OF DIRECT BEAM 1017.20 1028.12

DETECTOR ORIGIN (PIXELS) AT 1024.00 1024.00

CRYSTAL TO DETECTOR DISTANCE (mm) 1853.00

LAB COORDINATES OF DETECTOR X-AXIS 1.000000 0.000000 0.000000

LAB COORDINATES OF DETECTOR Y-AXIS 0.000000 1.000000 0.000000

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE ≥ -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR COMPARED expected	I/SIGMA	R-meas	CC (1/2)
5.26	3955	913	1038	88.0%	17.4%	13.5%	3948	8.39	96.5*
3.75	7308	1504	1651	91.1%	15.3%	13.2%	7301	9.51	97.1*
3.07	9728	1948	2064	94.4%	17.6%	15.7%	9720	7.91	97.2*
2.66	11663	2279	2415	94.4%	26.2%	24.9%	11654	5.46	93.9*
2.38	13259	2568	2702	95.0%	36.7%	40.2%	13243	3.78	89.8*
2.18	14681	2817	2966	95.0%	47.9%	56.2%	14663	2.94	81.2*
2.02	15908	3047	3204	95.1%	63.5%	79.5%	15884	2.17	72.1*
1.89	17179	3266	3435	95.1%	95.6%	132.0%	17156	1.33	51.5*
1.78	17426	3435	3638	94.4%	157.2%	245.8%	17351	0.69	25.5*
total	111107	21777	23113	94.2%	30.8%	34.4%	110920	3.72	96.8*



Setting up data processing in *DIALS*

```
dials.import template=./img/####.img\  
    distance=1853 detector.panel.pedestal=512 detector.panel.gain=14  
  
dials.find_spots imported.expt\  
    sigma_strong=6 nproc=32  
  
dials.search_beam_position imported.expt strong.refl\  
    n_macro_cycles=3  
  
dials.index optimised.expt strong.refl\  
    beam.fix=all detector.fix=distance  
  
dials.refine indexed.{expt,refl}\  
    beam.fix=all detector.fix=distance  
  
dials.integrate refined.{expt,refl}\  
    prediction.d_min=1.7 nproc=32  
  
dials.scale integrated.{expt,refl}\  
    d_min=1.78 min_Ih=1.0
```

Importing the data and defining additional detector parameters that are not listed in the metadata

Spot finding parameters for pixels to be included that exceed the background by at least 6 standard deviations

Data indexing and refinement of the lattice parameters

Data integration and scaling

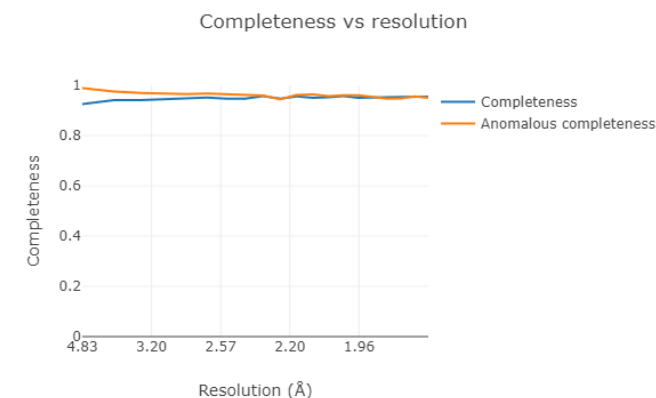
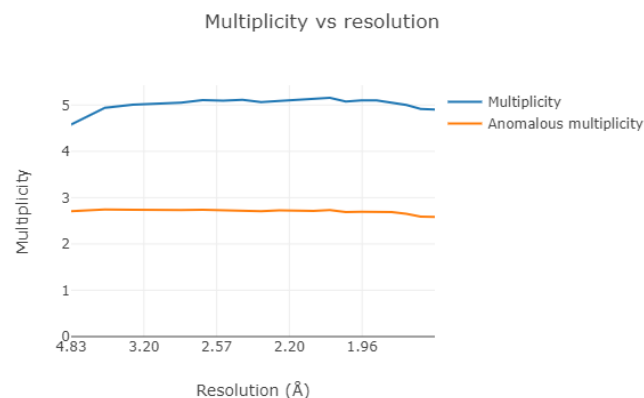
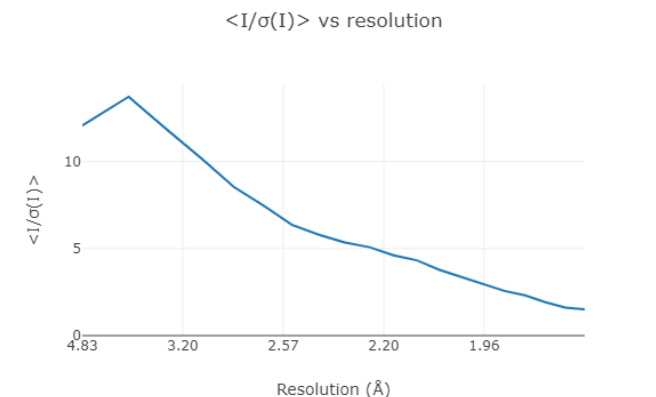
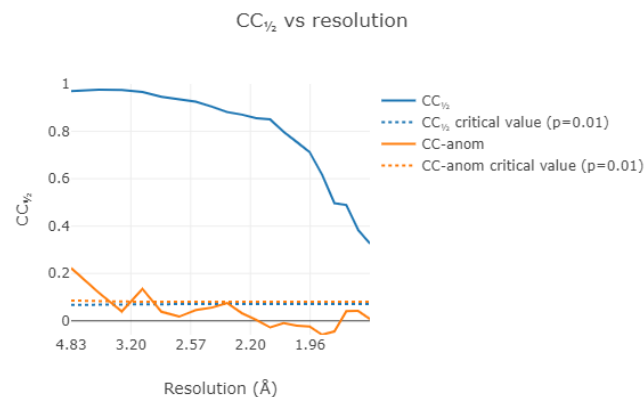
Refined unit cell parameters and intensity statistics

Unit cell: (67.1851(16), 67.1851(16), 102.221(4), 90.0, 90.0, 90.0)

Space group: P 4 2 2

Summary of merging statistics

	Overall	Low resolution	High resolution
Resolution (Å)	43.08 - 1.78	43.09 - 4.83	1.81 - 1.78
Observations	110649	5566	5358
Unique reflections	22007	1216	1093
Multiplicity	5.0	4.6	4.9
Completeness	94.87%	92.54%	95.46%
Mean $I/\sigma(I)$	5.8	12.1	1.5
R_{merge}	0.256	0.175	0.958
R_{meas}	0.286	0.198	1.077
R_{pim}	0.125	0.089	0.481
CC_{1/2}	0.972	0.970	0.327



Data merging statistics

	<i>XDS</i>			<i>DIALS</i>		
	Overall	InnerShell	OuterShell	Overall	InnerShell	OuterShell
Low resolution limit	28.05	28.05	1.82	56.16	56.16	1.81
High resolution limit	1.78	9.08	1.78	1.78	9.07	1.78
Rmerge (within I+/I-)	0.285	0.161	1.795	0.241	0.155	1.025
Rmerge (all I+ and I-)	0.320	0.176	1.944	0.274	0.171	1.152
Rmeas (within I+/I-)	0.353	0.200	2.237	0.299	0.193	1.280
Rmeas (all I+ & I-)	0.359	0.203	2.186	0.307	0.197	1.292
Rpim (within I+/I-)	0.205	0.118	1.315	0.174	0.113	0.754
Rpim (all I+ & I-)	0.157	0.099	0.971	0.135	0.095	0.571
Rmerge in top intensity bin	0.134	-	-	0.129	-	-
Total number of observations	111107	685	5720	108664	789	6058
Total number unique	21777	188	1206	21984	209	1198
Mean(I)/sd(I)	3.7	8.0	0.5	5.5	11.1	1.3
Mn(I) half-set correlation CC(1/2)	0.968	0.963	0.162	0.971	0.959	0.325
Completeness	95.6	88.5	93.9	95.8	95.0	91.9
Multiplicity	5.1	3.6	4.7	4.9	3.8	5.1

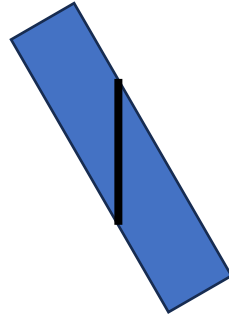
STARANISO Server

<https://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>



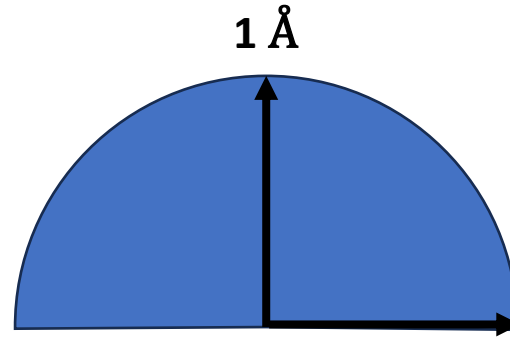
Tilt= 0°

Thickness = t
resolution = 1 Å

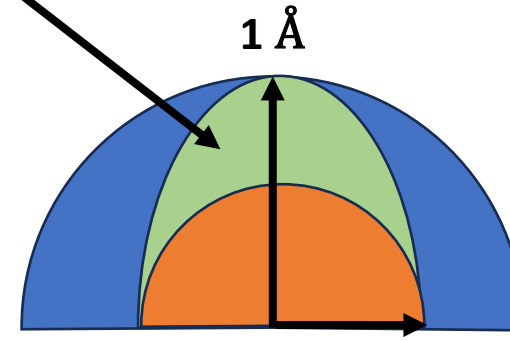


Tilt= 60°

Thickness = $2t$
resolution = 1.6 Å



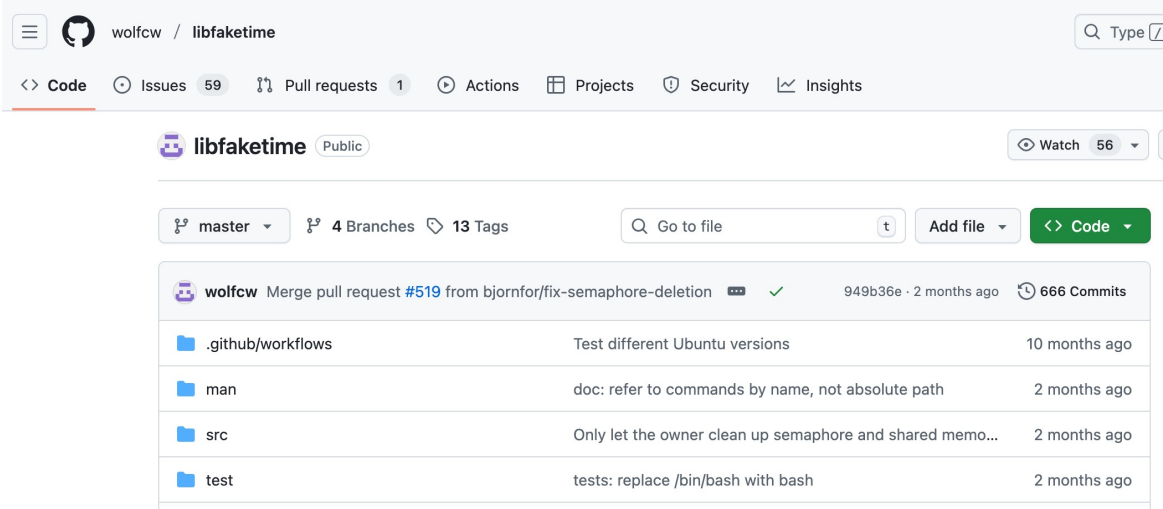
1 Å



1 Å

1.6 Å

A word on XDS versions



<https://wiki.uni-konstanz.de/pub/xds/>
<https://pypi.org/project/faketime/>
<https://github.com/wolfcw/libfaketime>

	2020-apr17/	2021-02-19 21:53	-
	2020-jan31/	2020-06-28 21:36	-
	2021-Feb05/	2021-03-23 10:13	-
	2021-Mar23/	2022-01-24 11:12	-
	2022-Aug20/	2023-09-01 17:00	-
	2022-Feb20/	2022-12-05 09:54	-
	2022-Jan20/	2022-06-02 08:34	-
	2023-Jun30/	2024-07-29 15:42	-
	2024-Jun23/	2024-10-02 14:37	-
	2024-Oct02/	2025-02-14 15:52	-
	2025-Apr09/	2025-05-03 13:31	-
	2025-Apr30/	2025-07-20 13:35	-
	2025-Feb24/	2025-03-24 17:26	-
	2025-Jan19/	2025-03-24 15:39	-
	2025-Jul14/	2025-10-22 15:16	-
	2025-Mar20/	2025-03-28 15:32	-
	2025-Mar27/	2025-04-10 18:10	-
	VIEW.linux	2007-12-11 15:24	72K
	XDS-Apple_M1.tar.gz	2025-07-14 14:32	5.3M
	XDS-OSX_64.tar.gz	2025-07-14 14:32	9.9M
	XDS-gfortran_Linux_x86_64.tar.gz	2025-07-14 14:32	9.7M
	XDS_html_doc.tar.gz	2025-07-14 14:32	130K
	XDS_html_doc/	2025-03-26 22:27	-
	wget.rc	2025-10-22 15:18	339



Phaser

Phaser (Phase) - Home
phaser.phar.com | phaser.phar.com
log in

[page](#)
[discussion](#)
[view source](#)
[history](#)

Molecular Replacement

Quicklink to example scripts → MR using keyword

Quicklink to phaser.famoc find_all_etc_gym_mats documentation → Famoc

Phaser should be able to solve most structures with the Automated Molecular Replacement mode, and this is the first mode that you should try. Give Phaser your data ([How to Define Data](#)) and your models ([How to Define Models](#)), tell Phaser what to search for, and a set of possible spacings (in the same grid group).

If this doesn't work (see [Phaser Solved?](#)!), you can try selecting peaks of lower significance in the rotation function case as the real orientation was not within the selection criteria. By default peaks above 75% of the top peak are selected (see [How to Select Peaks](#)). See [What to do in Difficult Cases](#) for more hints and tips. If the automated molecular replacement mode doesn't work even with non-default input you need to run the models by **Phaser separately**. The possibilities are endless – you can even try exhaustive searches (iterations of all orientations) if you want, but experience has shown that most structures that can be solved by Phaser can be solved by relatively simple strategies.

Automated Molecular Replacement

Automated Molecular Replacement combines the anisotropy correction, likelihood enhanced fast rotation function, likelihood enhanced fast translation function, packing and refinement modes for multiple search methods and a set of possible spacings to automatically solve a structure by molecular replacement. Top solutions are output to the files FILEROT.out, FILEROT.mtz and FILEROT.pdb where "FILEROT" refers to the sorted solution number, 1 being the best, and any 1's output by default). Many structures can be solved by running an automated molecular replacement search with defaults, giving the ensembles that you expect to be easiest to find first.

At the completion of Molecular Replacement you may wish to place your solutions on a common origin with a previous solution, for which [Famoc](#) can be used.

Sites

- Phaser/PH home
- News
- Downloads
- Manuals
- Tutorials
- FACG
- The Top Ten
- PUBLICATIONS
- External Links

Tools

- MRF Phasing
- SAD Phasing

Webpages

- Phaser interface
- Contact Developers
- Developer Pages
- Licenses
- Source Code
- Feedback
- GitHub

Search

Search PhaserWiki

Links

- What links here
- Related changes
- Special pages
- Printable version
- Permanent link
- Page information

```

graph TD
    A[All data corrected] --> B[Anisotropy Correction]
    A --> C[TRMS Correction]
    B --> D[Expected LLOs]
    C --> D
    D --> E[Fast Rotation Function]
    E --> F[eLLG (ortho) ensemble + resolution limited data]
    E --> G[All peak selection criteria]
    F --> H[Loop over ensembles]
    G --> H
    H --> I[Fast Translation Function]
    I --> J[Packing]
    J --> K[Refinement and Phasing]
    K --> L[Best N solutions for 1st ensemble, Best ensembles]
    L --> M[Fast Translation Function]
    M --> I
    N[All data corrected] --> O[New LLG (ortho) ensemble + resolution limited data]
    O --> P[Fast Translation Function]
    P --> Q[Packing]
    Q --> R[Refinement and Phasing]
    R --> S[Best solutions for complete structure]
    S --> L
    
```

Should Phaser Solve It?

The difficulty of a molecular replacement problem depends primarily on two major factors: how well the model will be able to explain the diffraction data (which depends both on the accuracy of the model and on its completeness), and how many reflections can be explained, at least in part. One reflection provides a piece of information that helps to identify correct MR solutions.

It is possible to make a reasonable prediction of whether or not a solution will be found. If the quality of the model (its accuracy and completeness) can be estimated, then the expected contribution of each reflection to the total LLO can also be estimated. From a large battery of tests, we know that an LLO of 40 or greater usually indicates a correct solution (at least in the absence of complicating factors such as translational non-crystallographic symmetry, INCS). Building on this understanding, if it is estimated that the LLO will be 60 or less, then Phaser will assume that the problem is a difficult one, and will implement search procedures for difficult problems.

What Resolution of Data Should be Used?

Contents [hide]

- 1 Automated Molecular Replacement
- 2 Should Phaser Solve It?
- 3 What Resolution of Data Should be Used?
- 4 Has Phaser Solved It?
- 5 Anisotropy
- 6 History
- 7 What to do in Difficult Cases
- 8 How to Define Models
- 9 How to Define Spacings
 - 9.1 Working an Ensemble from Electron Density
 - 9.1.1 Coordinate Editing
 - 9.1.1.1 METAREFINLANDS
 - 9.1.1.2 WATER
- 10 How to Define Composition
 - 10.1 Default Composition
 - 10.2 Composition by Sequence
 - 10.3 Composition by Atom
 - 10.4 Composition by Solvent Content
 - 10.5 Composition by Number of Residues in ASU
 - 10.6 Composition by Molecular Weight
 - 10.7 Composition by Percentage Scattering
- 11 How to Define Solutions
 - 12.1 "No" Files
 - 12.2 "Yes" Files
 - 12.3 Fixed partial structure
- 12 How to Select Peaks
- 13 How to Control Output
- 14 Translational Non-crystallographic Symmetry
 - 15.1 Default INCS detection and correction
 - 15.1.1 No INCS
 - 15.1.2 Pairs of molecules
 - 15.2 Non-default INCS correlation
 - 15.2.1 INCS involving multimers
 - 15.2.2 Higher order INCS
 - 15.2.3 Complex INCS

McCoy *et al.*, *J. Appl. Cryst.* 40, 658-674 (2007)

<https://www.phaser.cimr.cam.ac.uk/index.php/Phaser> Crystallographic Software

Finding a suitable search model

Setting up molecular replacement in *Phaser*

Electron scattering factors can be selected from the drop-down menu

The screenshot shows the 'Configure' window of the Phaser-MR software. The 'Input and general options' tab is selected. The 'Title' field contains 'mr_protk_6cl7_truncated_dials'. The 'Phaser mode' is set to 'Automated molecular replacement'. The 'Input files' section shows the 'Data file' as '/mnt/gonen/home/mclabbers/aca/dials/a89.mtz' and the 'Unit cell' as '67.144 67.144 102.478 90 90 90'. The 'Space group' is set to 'P 4 2 2'. The 'Data labels' are set to 'IMEAN,SIGIMEAN'. The 'High resolution' field is empty. The 'Use partial solution from previous job' is set to '---'. The 'Number of parallel threads' is set to '63'. The 'Output' section shows the 'Output directory' as '/mnt/gonen/home/mclabbers/aca/phenix'. The status bar at the bottom indicates 'Idle' and 'Project: aca'.

As this stage we have only narrowed down the space group to *P 4 2 2*

The screenshot shows the 'Configure' window of the Phaser-MR software, with the 'Search procedure' tab selected. The 'Protocol' section shows the 'Strategy' with a list of searches. The first search is 'Component 1' with a value of '1'. The 'Search options' section shows 'Also try alternative space group(s)' set to 'All possible in same pointgroup' and 'Search method' set to 'Fast'. The status bar at the bottom indicates 'Idle' and 'Project: aca'.

Phaser can search for the most likely solution in all other possible space groups within the same point group

Finding a well-contrasting solution

Phaser-MR (full-featured) (Project: aca)@bcminnie

File Actions Settings Utilities Help

Preferences Help Run Abort Save Xtriage Ensembler Sculptor Simple GUI

Configure **phaser_27**

Run status Graphs **Results**

Summary

Phaser has found 1 MR solution(s).

Top LLG: 7068.359

Top TFZ: 70.6

Spacegroup: P 43 21 2

Run phenix.refine

Run AutoBuild

Run MR-SAD

Output files

Title: mr_protk_6cl7_truncated_dials

Directory: /mnt/gonen/home/mclabbers/aca/phenix/phaser_27

File name	Contents
phaser.1.mtz	Phases and maps
phaser.1.pdb	Output model

Open in Coot Open in PyMOL

Idle Project: aca

TF Z-score	Have I solved it?
less than 5	no
5 - 6	unlikely
6 - 7	possibly
7 - 8	probably
more than 8*	definitely
*6 for 1st model in monoclinic space groups	

https://www.phaser.cimr.cam.ac.uk/index.php/Molecular_Replacement

Structure refinement

REFMAC5

Murshudov *et al.*, *Acta Cryst.* D67, 355-367 (2011)

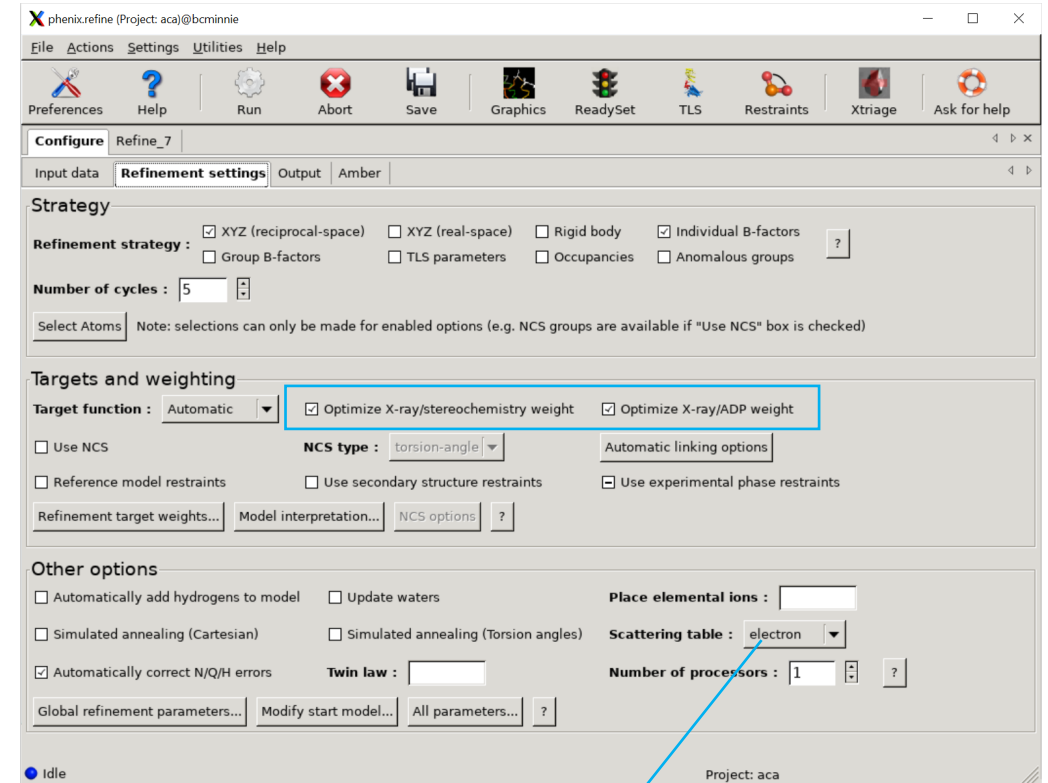
```
refmac5 \
  xyzin phaser.1.pdb \
  hklin a.mtz \
  xyzout refmac.pdb \
  hklout refmac.mtz \
  << eof | tee refmac.log
make check NONE
make -
  peptide NO -
  hydr all -
  cispeptide YES -
  ssbridge YES -
  symmetry YES -
  sugar YES -
  connectivity NO -
  link NO -
  source electron mb
  ridg dist sign 0.02
refi -
  type REST -
  resi MLKF -
  meth CGMAT -
  bref ISOT
ncyc 100
scal -
  type SIMP -
  LSSC -
  ANISO -
  EXPE
solvent YES
weight MATRIX auto
monitor MEDIUM -
  torsion 10.0 -
  distance 10.0 -
  angle 10.0 -
  plane 10.0 -
  chiral 10.0 -
  bfactor 10.0 -
  bsphere 10.0 -
  rbond 10.0 -
  ncsr 10.0
labin FP=FREE SIGFP=SIGF FREE=FreeR_flag
labout FC=FC FWT=FWT PHIC=PHIC PHWT=PHWT DELFWT=DELFWT PHDELWT=PHDELWT FOM=FOM
PNAME f4pk
DNAME 1
RSIZE 80
END
eof
```

Electron scattering factors

Electron scattering factors can be set using the keyword SOURCE ELECTRON MB, the keyword MAPC FREE EXCLUDE ensures that no missing reflections will be filled in with weighted estimates of Fc for map calculations

phenix.refine

Afonine *et al.*, *Acta Cryst.* D68, 352-367 (2012)



Electron scattering factors

Electron scattering factors can be selected from the drop-down menu, *phenix.refine* writes out two 2mFo-DFc maps, with and without restoring missing reflections with weighted estimates of Fc for map calculations

Rigid body refinement of the molecular replacement solution

phenix.refine (Project: aca)@bcminnie

File Actions Settings Utilities Help

Preferences Help Run Abort Save Graphics ReadySet TLS Restraints Xtriage Ask for help

Configure Refine_7

Results MolProbity Real-space correlation Atomic properties

Output files

Title: refine_6cl7_tpvh_dials

Directory: /mnt/gonen/home/mclabbers/aca/phenix/Refine_7

File name	Contents
refine_7.log	phenix.refine log file
refine_7.geo	Geometry restraints before refine...
refine_7.eff	Effective parameters for this run
refine_7.pdb	Refined model
refine_7.cif	mmCIF model for PDB deposition

Convert map coefficients MTZ to CCP4 maps

Refinement statistics

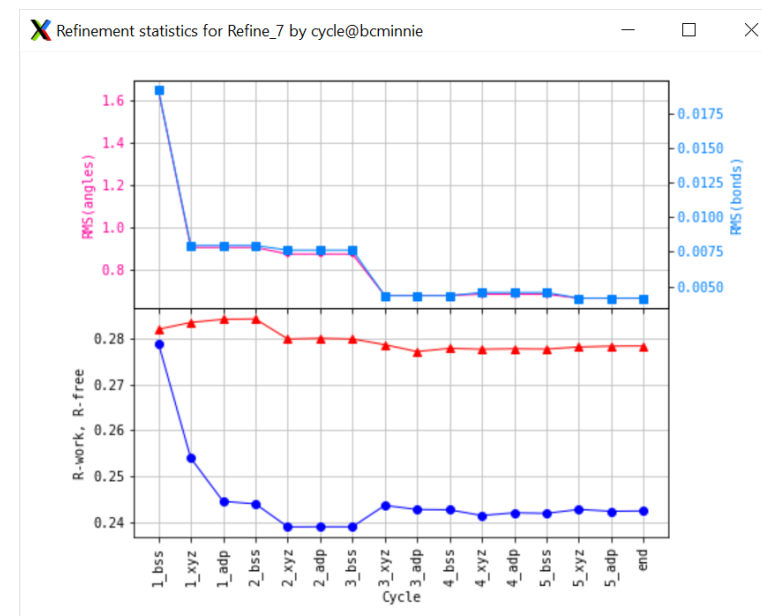
Compare statistics Plot statistics by cycle Plot statistics by resolution

Before and after refinement:

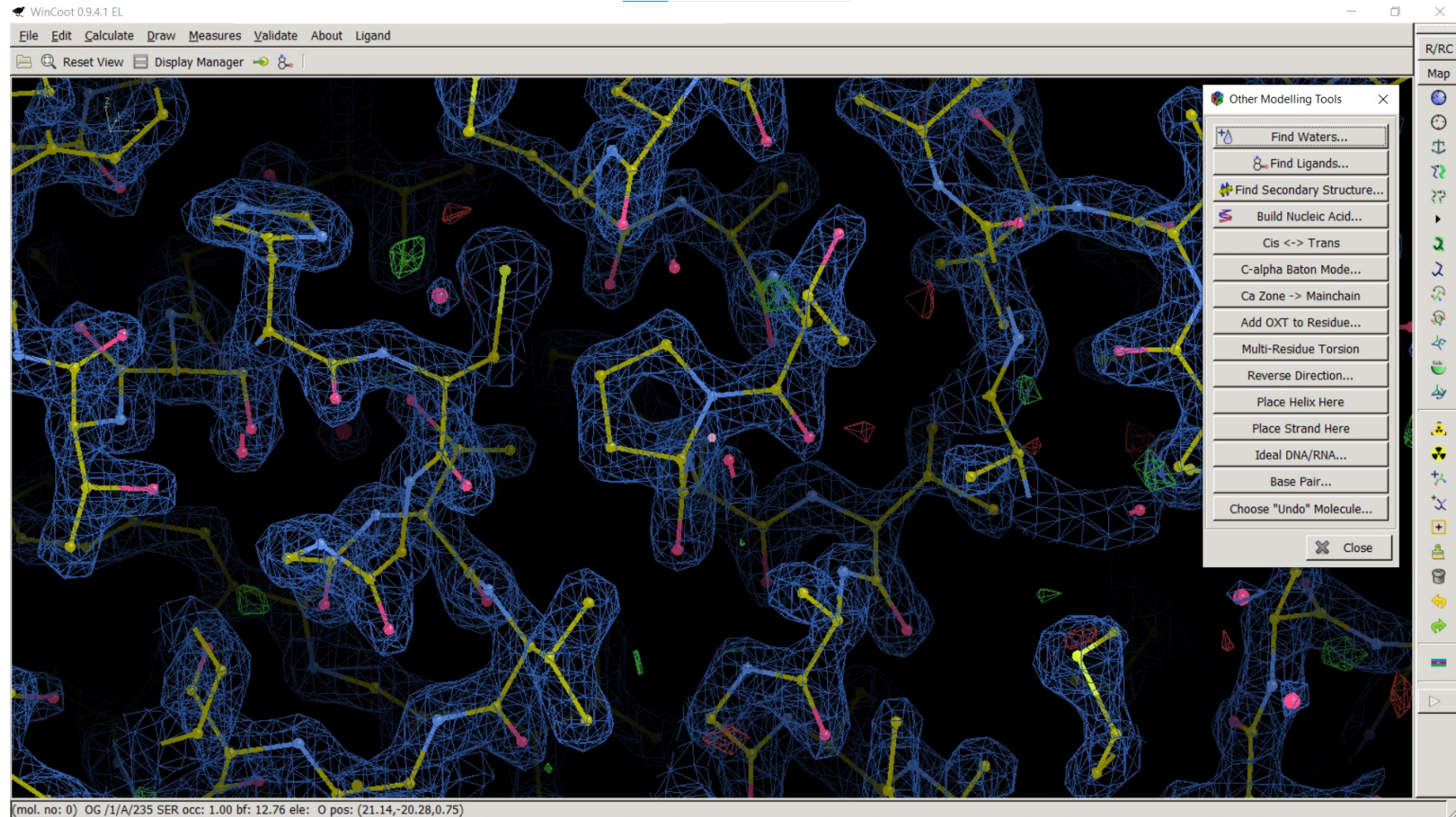
	Starting	Final
R-work	0.2789	0.2424
R-free	0.2820	0.2783
Bonds	0.019	0.004
Angles	1.650	0.666

X-ray statistics by resolution bin:

Idle Project: aca

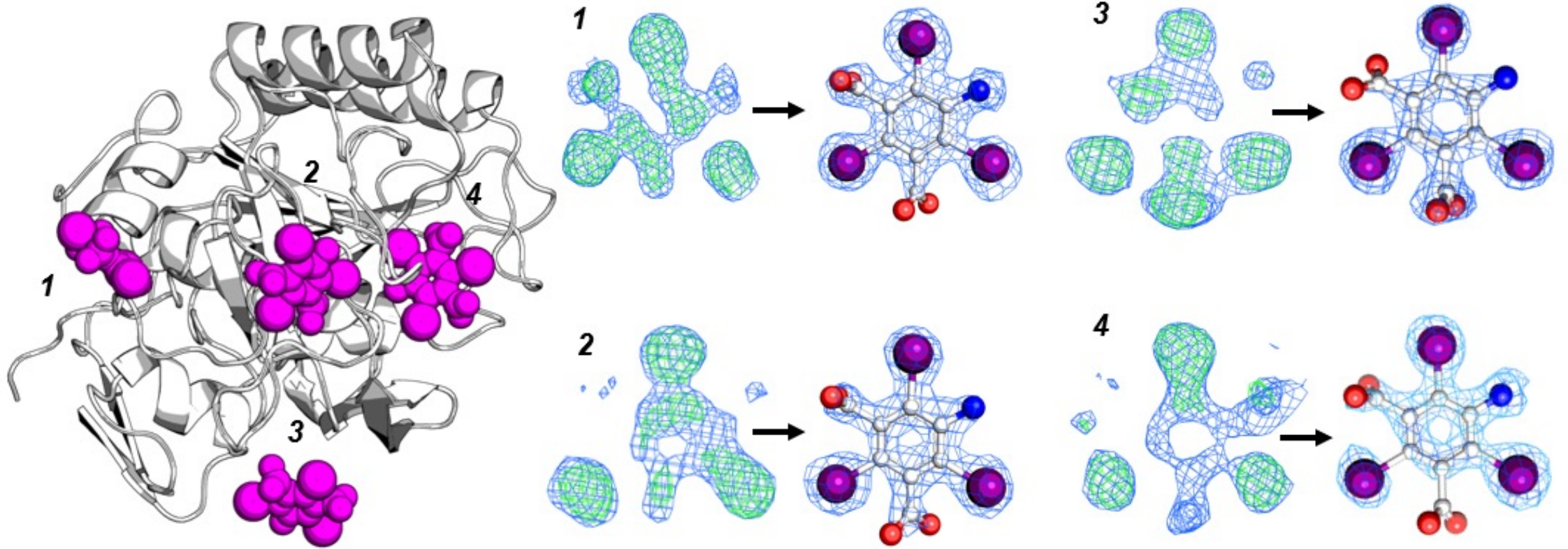


Model building using Coot

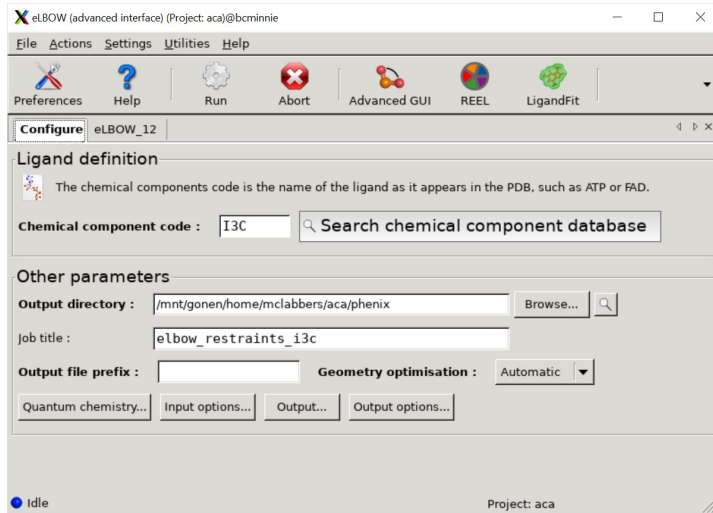


Emsley *et al.*, *Acta Cryst. D*66, 486-501 (2010)

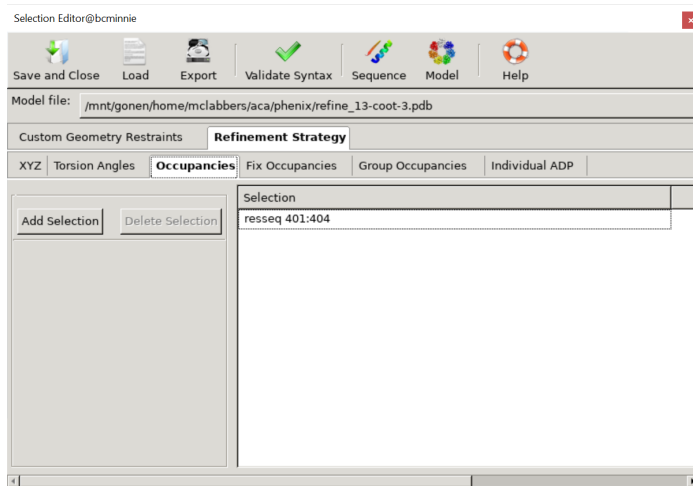
Placing ligands based on the difference map



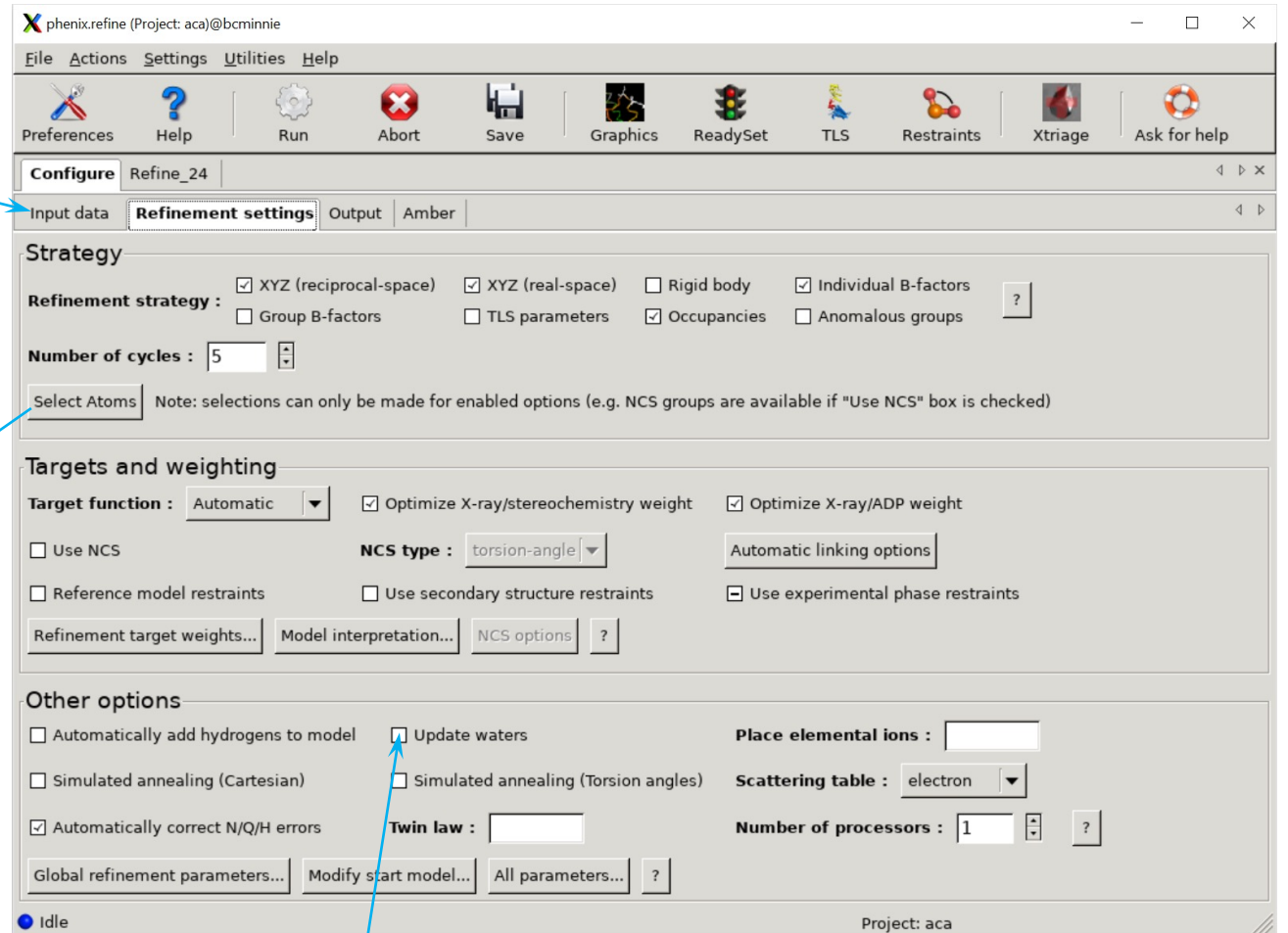
Structure refinement



Generating ligand restraints for refinement

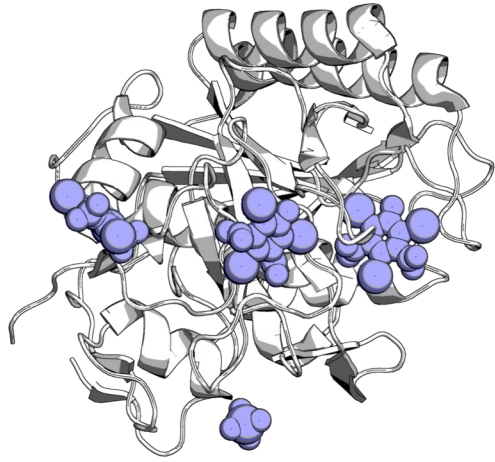


Selecting atoms/residues for occupancy refinement



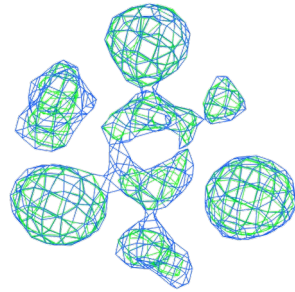
Update waters automatically within phenix.refine

High occupancy ligands



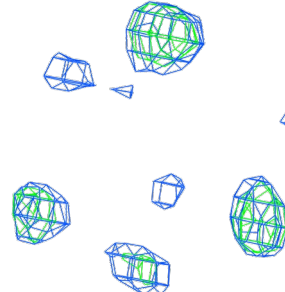
X-ray

1



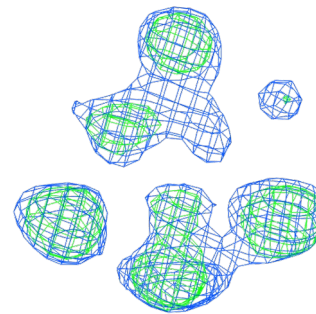
75% occupancy

2



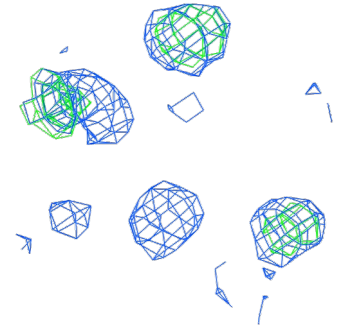
21% occupancy

3



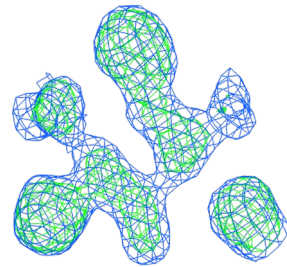
0% occupancy

4

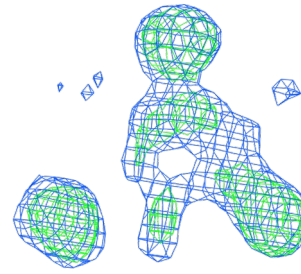


12% occupancy

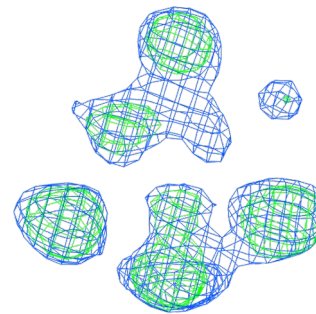
MicroED



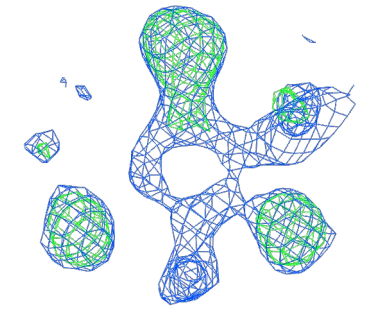
75% occupancy



58% occupancy



61% occupancy



59% occupancy

Finalizing the refinement

phenix.refine (Project: aca)@bcminnie

File Actions Settings Utilities Help

Preferences Help Run Abort Save Graphics ReadySet TLS Restraints Xtriage Ask for help

Configure **Refine_25**

Results MolProbity Real-space correlation Atomic properties Hydrogens

Output files

Title: refine_6cl7_tpvh_dials_rebuild_elbow_occupancy
Directory: /mnt/gonen/home/mclabbers/aca/phenix/Refine_25

File name	Contents
refine_25.log	phenix.refine log file
refine_25.geo	Geometry restraints before refine...
refine_25.eff	Effective parameters for this run
refine_25.pdb	Refined model
refine_25.cif	mmCIF model for PDB deposition

Convert map coefficients MTZ to CCP4 maps

Open in Coot
Open in PyMOL
Sequence viewer

Refinement statistics

Compare statistics **Plot statistics by cycle** **Plot statistics by resolution**

Before and after refinement:

	Starting	Final
R-work	0.1906	0.1750
R-free	0.2252	0.2224
Bonds	0.003	0.010
Angles	0.626	0.870

X-ray statistics by resolution bin:

	R-work	R-free	%complete	FOM	Phase error	Scale factor	#work	#test
56.16 - 3.56	0.1534	0.1718	93.7%	0.91	12.35	1.00	2755	156
3.56 - 2.82	0.1566	0.2117	94.7%	0.90	15.37	1.01	2654	131
2.82 - 2.47	0.1730	0.2308	94.8%	0.87	19.71	1.01	2615	143
2.47 - 2.24	0.1723	0.2299	95.1%	0.88	20.33	0.99	2585	138
2.24 - 2.08	0.1761	0.2494	95.2%	0.88	21.84	1.02	2573	152
2.08 - 1.96	0.2005	0.2598	95.4%	0.86	22.57	1.06	2585	132
1.96 - 1.86	0.2141	0.2892	95.5%	0.85	26.25	0.98	2584	118
1.86 - 1.78	0.2626	0.3137	92.8%	0.84	25.35	0.90	2502	132

Idle Project: aca

phenix.refine (Project: aca)@bcminnie

File Actions Settings Utilities Help

Preferences Help Run Abort Save Graphics ReadySet TLS Restraints Xtriage Ask for help

Configure **Refine_25**

Results **MolProbity** Real-space correlation Atomic properties Hydrogens

Summary **Basic geometry** Protein Clashes

This page reports statistics for geometry restraints used in refinement. As a general rule, a fully refined structure should not have any outliers unless these are exceptionally clear in the electron density (usually at very high resolution). Be sure to also check the MolProbity validation results for this structure, as they are more sensitive to the geometric properties of proteins and nucleic acids (especially in the case of dihedral angles).

Bond length restraints

Number of restraints: 2134
RMS(deviation): 0.010
Max. deviation: 0.092
Number of outliers > 4sigma: 0

Bond angle restraints

Number of restraints: 2910
RMS(deviation): 0.872
Max. deviation: 9.723
Number of outliers > 4sigma: 0

Dihedral angle restraints

Number of restraints: 1222
RMS(deviation): 12.431
Max. deviation: 71.155
Number of outliers > 4sigma: 0

Chirality restraints

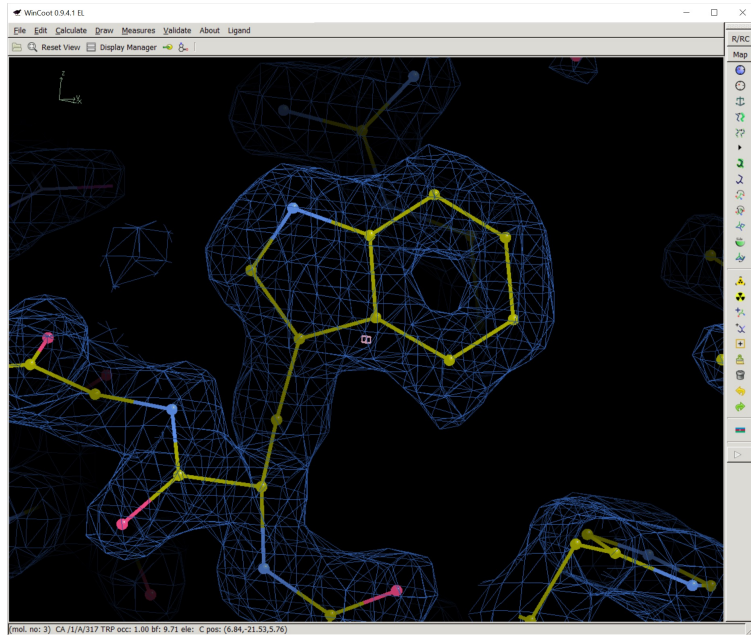
Number of restraints: 312
RMS(deviation): 0.057
Max. deviation: 0.251
Number of outliers > 4sigma: 0

Planarity restraints

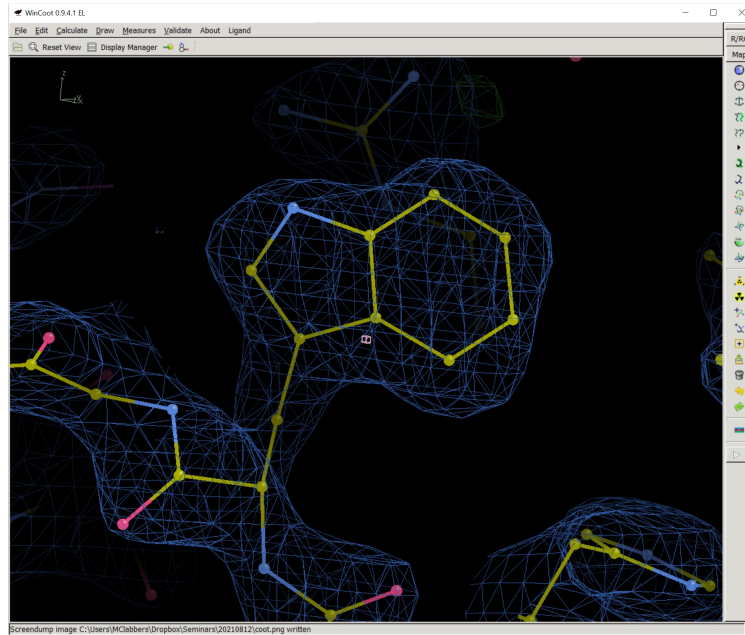
Number of restraints: 382
RMS(deviation): 0.007
Max. deviation: 0.050
Number of outliers > 4sigma: 0

Idle Project: aca

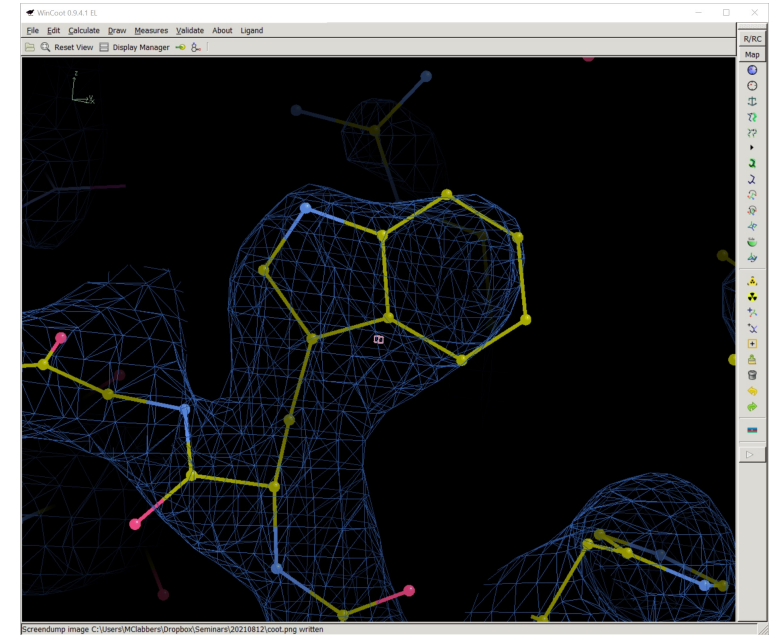
The best strategy for modelling and refinement depends on the data quality



1.78 Å



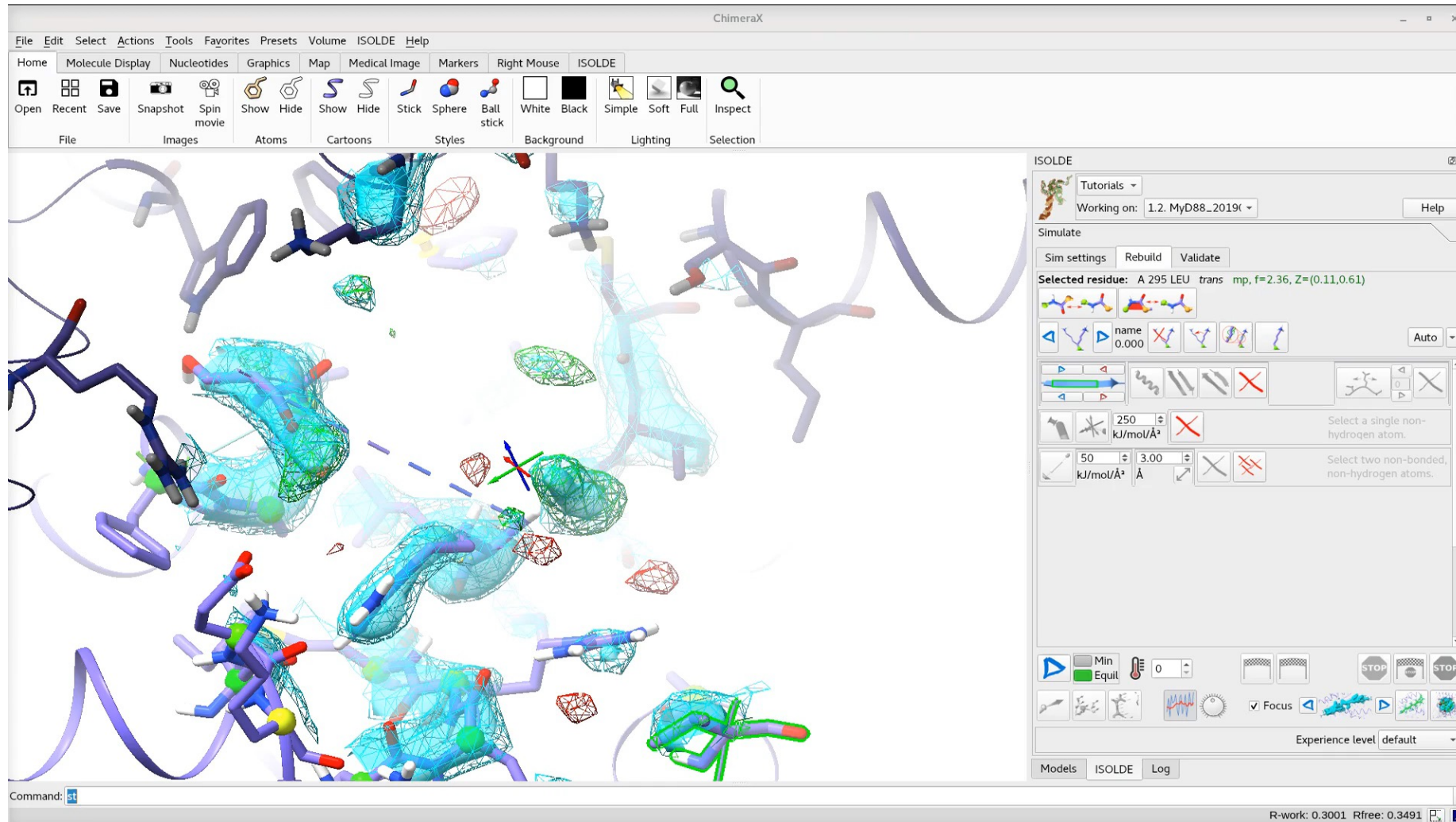
2.5 Å



3.5 Å

Electrostatic potential 2mFo-DFc maps are shown in blue, contoured at 1.6σ . Difference mFo-DFc maps are shown in green and red at a contour level of 2.8σ for positive and negative potential, respectively.

Interactive model building using medium to low resolution data with *ISOLDE*

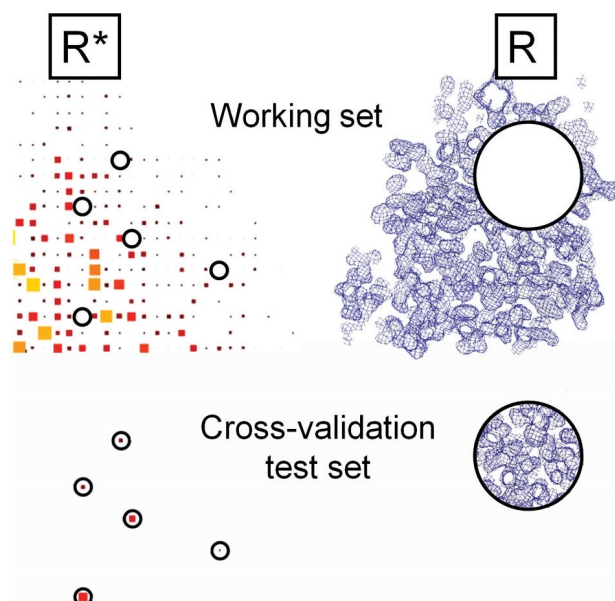


Screen capture courtesy of Tristan Croll. Data from Clabbers *et al.*, *Nat. Commun.* 12, 2578 (2021)

Croll, *Acta Cryst.* D74, 519-530 (2018); <https://isolde.cimr.cam.ac.uk/>

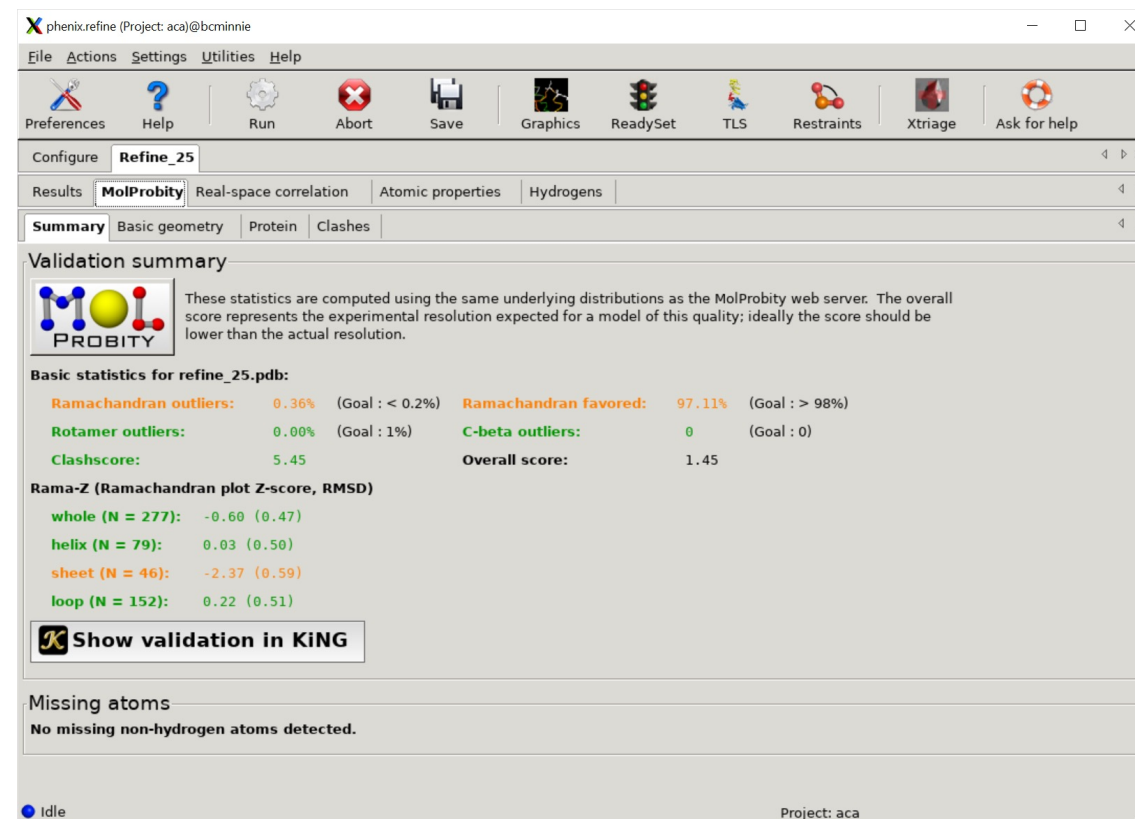
Validation of crystallographic data and model stereochemistry

Bernhard Rupp, Biomolecular Crystallography,
Garland Science, New York (2009)



Rwork/Rfree 0.175/0.222

SA composite omit map

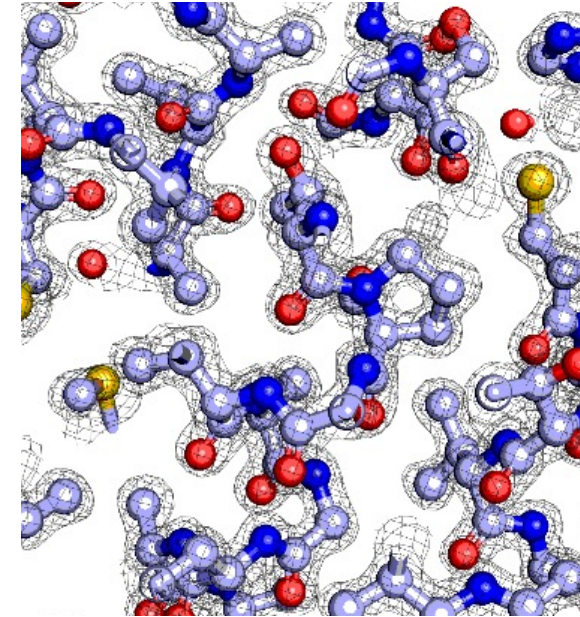
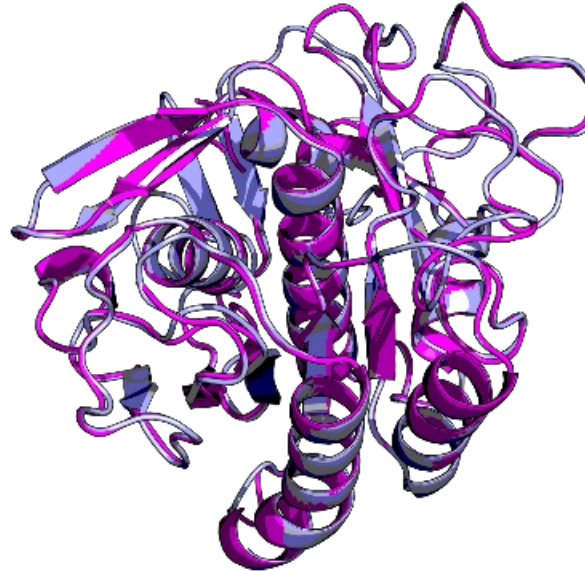
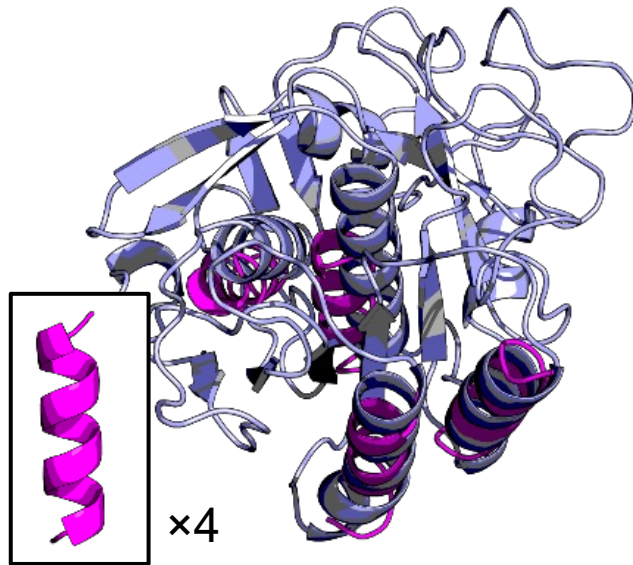


Example 2: Fragment based *ab initio* structure determination of proteinase K

Place ideal helical fragments

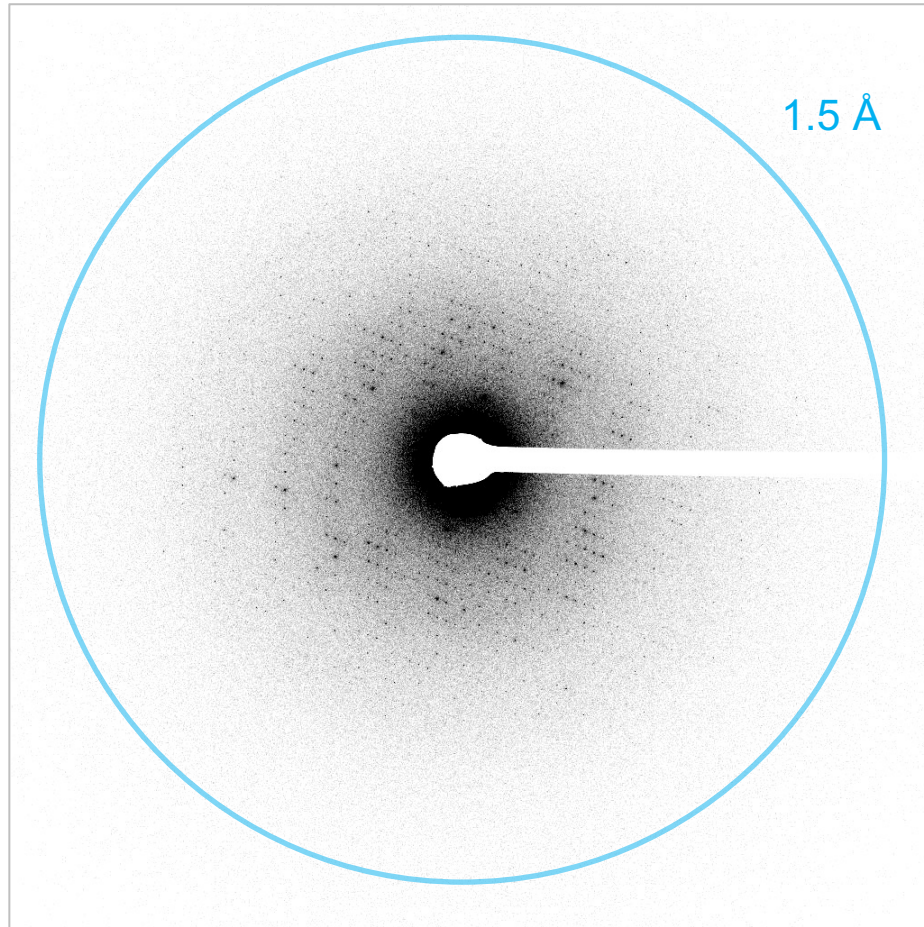
Density Modification
and Chain Tracing

Automatic Building

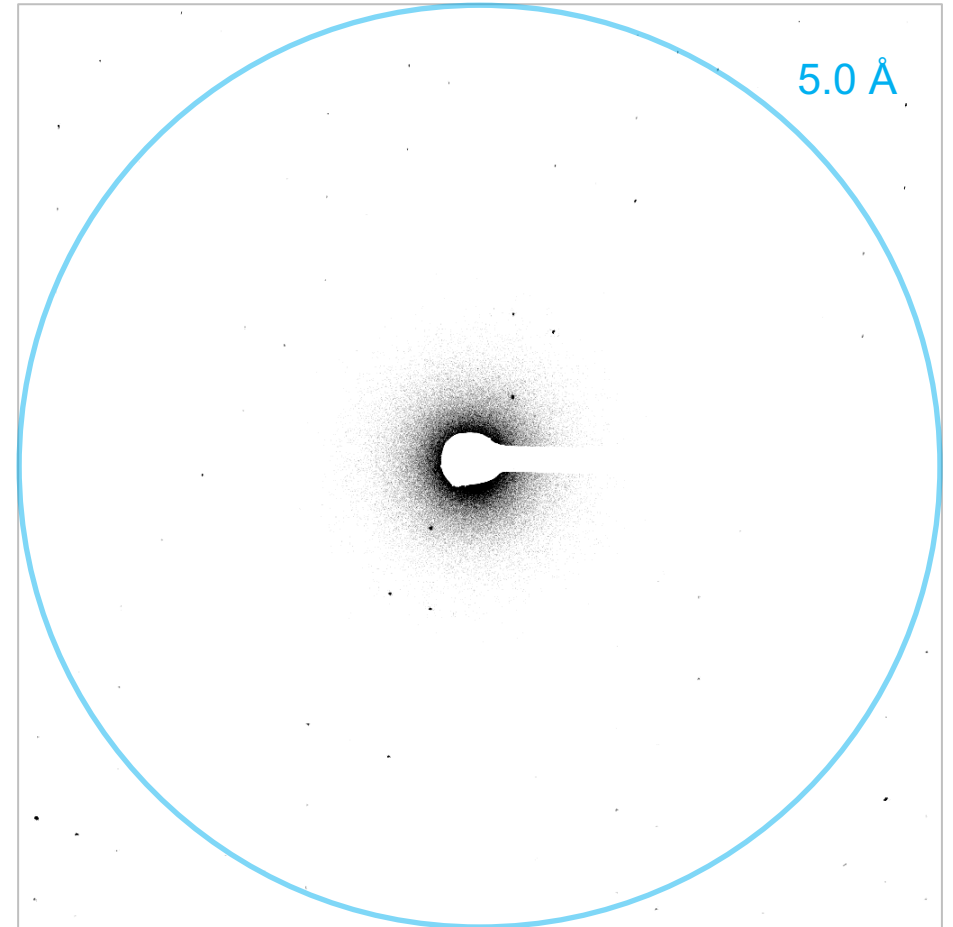


MicroED data collection on proteinase K crystal lamellae using electron counting

High resolution pass



Low resolution pass



Integrating the high resolution data using XDS

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE ≥ -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION	NUMBER OF REFLECTIONS			COMPLETENESS	R-FACTOR	R-FACTOR	COMPARED	I/SIGMA	R-meas	CC (1/2)
LIMIT	OBSERVED	UNIQUE	POSSIBLE	OF DATA	observed	expected				
4.40	7181	1529	1748	87.5%	11.4%	9.5%	7175	13.29	12.8%	98.3*
3.15	13637	2590	2807	92.3%	11.6%	10.3%	13628	13.08	12.8%	98.6*
2.58	17862	3275	3524	92.9%	18.0%	15.2%	17847	9.13	20.0%	97.0*
2.24	21341	3837	4120	93.1%	24.6%	22.7%	21321	6.57	27.1%	94.7*
2.01	24517	4353	4654	93.5%	30.2%	30.0%	24494	5.20	33.3%	92.3*
1.83	26894	4755	5084	93.5%	42.6%	49.5%	26866	3.50	47.0%	85.2*
1.70	29221	5137	5488	93.6%	65.6%	100.7%	29188	2.11	72.3%	68.3*
1.59	31369	5512	5877	93.8%	96.5%	210.4%	31333	1.33	106.3%	45.2*
1.50	31980	5801	6253	92.8%	134.7%	501.3%	31896	0.85	148.8%	25.4*
total	204002	36789	39555	93.0%	25.3%	36.2%	203748	4.67	28.0%	98.4*

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE ≥ -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION	NUMBER OF REFLECTIONS			COMPLETENESS	R-FACTOR	R-FACTOR	COMPARED	I/SIGMA	R-meas	CC (1/2)
LIMIT	OBSERVED	UNIQUE	POSSIBLE	OF DATA	observed	expected				
4.39	7146	1582	1755	90.1%	12.2%	10.1%	7138	11.73	13.7%	97.6*
3.15	13450	2665	2789	95.6%	12.4%	11.0%	13444	11.44	13.9%	98.0*
2.58	17781	3384	3528	95.9%	19.4%	16.4%	17772	8.07	21.6%	95.8*
2.24	21161	3985	4105	97.1%	26.9%	24.2%	21150	5.77	30.0%	92.5*
2.01	24081	4476	4601	97.3%	32.2%	33.1%	24065	4.62	35.8%	89.5*
1.83	26742	4929	5070	97.2%	44.2%	54.5%	26724	3.20	49.1%	80.6*
1.70	29089	5350	5475	97.7%	66.8%	96.8%	29067	1.97	74.2%	60.0*
1.59	31165	5719	5852	97.7%	95.4%	152.0%	31143	1.28	105.9%	43.9*
1.50	32179	6073	6231	97.5%	129.1%	360.8%	32115	0.84	143.6%	21.1*
total	202794	38163	39406	96.8%	26.8%	34.1%	202618	4.16	29.9%	97.6*



Integrating the low resolution data using *XDS*

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE ≥ -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION	NUMBER OF REFLECTIONS			COMPLETENESS	R-FACTOR	R-FACTOR	COMPARED	I/SIGMA	R-meas	CC (1/2)
LIMIT	OBSERVED	UNIQUE	POSSIBLE	OF DATA	observed	expected				
14.22	131	47	68	69.1%	22.5%	16.2%	119	5.60	26.5%	93.0*
10.36	286	81	94	86.2%	17.0%	16.8%	281	6.03	19.7%	95.7*
8.55	402	101	110	91.8%	17.1%	16.4%	400	6.75	20.2%	92.9*
7.44	470	111	123	90.2%	17.1%	16.8%	469	7.15	20.2%	91.0*
6.68	601	131	143	91.6%	17.0%	17.5%	601	7.22	19.0%	96.3*
6.11	683	142	153	92.8%	19.2%	18.5%	682	7.12	21.5%	93.5*
5.67	739	158	167	94.6%	17.8%	18.3%	738	6.90	19.8%	95.2*
5.31	759	165	171	96.5%	17.6%	21.2%	759	7.21	19.9%	96.3*
5.01	536	167	189	88.4%	15.0%	17.8%	519	5.84	17.8%	95.7*
total	4607	1103	1218	90.6%	17.6%	18.0%	4568	6.75	20.2%	94.3*

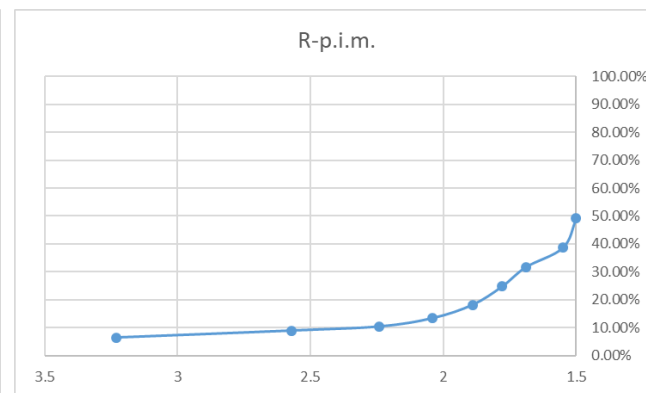
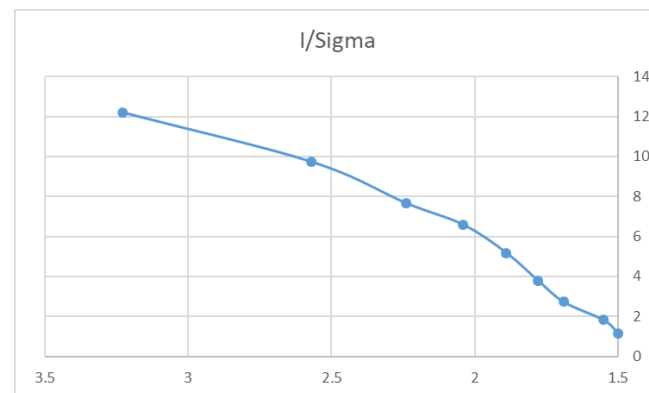
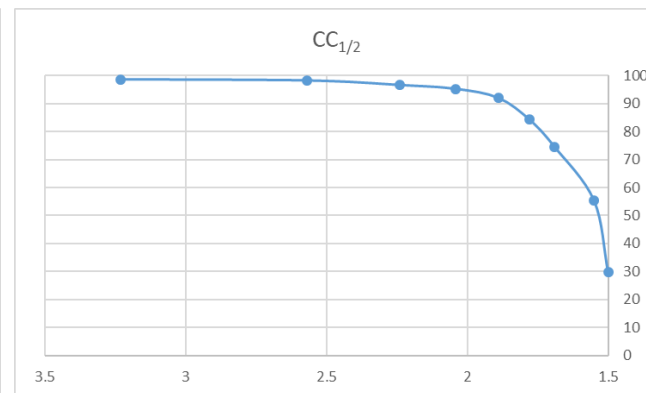
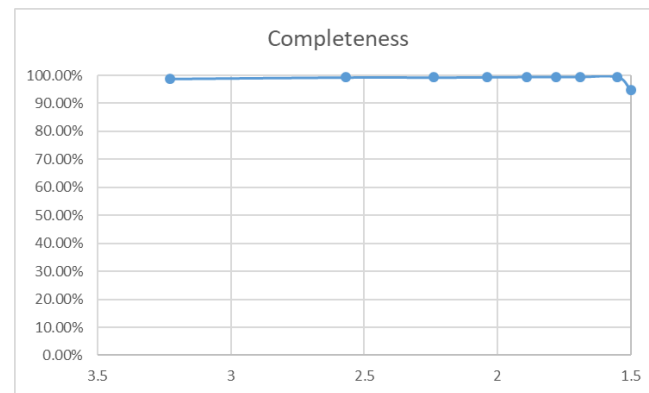
SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE ≥ -3.0 AS FUNCTION OF RESOLUTION

RESOLUTION	NUMBER OF REFLECTIONS			COMPLETENESS	R-FACTOR	R-FACTOR	COMPARED	I/SIGMA	R-meas	CC (1/2)
LIMIT	OBSERVED	UNIQUE	POSSIBLE	OF DATA	observed	expected				
14.27	137	51	67	76.1%	13.0%	10.5%	132	6.56	16.2%	98.6*
10.38	300	77	95	81.1%	16.6%	14.7%	297	6.04	19.1%	97.8*
8.55	432	99	111	89.2%	15.5%	13.3%	431	8.35	17.7%	98.0*
7.45	523	111	123	90.2%	16.2%	15.2%	523	8.17	18.2%	96.9*
6.68	641	133	144	92.4%	18.6%	17.9%	639	7.19	20.8%	95.0*
6.11	698	139	153	90.8%	19.1%	20.3%	698	6.60	21.2%	96.5*
5.66	803	153	167	91.6%	19.7%	22.2%	803	6.23	22.0%	95.2*
5.30	788	159	171	93.0%	19.7%	21.4%	787	6.52	22.2%	97.0*
5.00	550	159	190	83.7%	15.2%	20.2%	538	5.77	18.0%	95.6*
total	4872	1081	1221	88.5%	17.4%	17.9%	4848	6.77	19.7%	97.3*

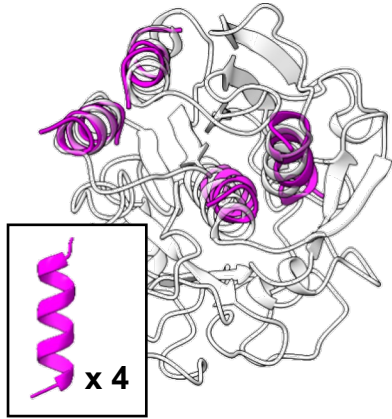


Merging statistics

No. of crystals	2
Space group	$P4_32_12$
Unit cell dimensions	
a, b, c (Å)	67.08, 67.08, 106.78
α, β, γ (°)	90, 90, 90
Resolution (Å)	43.35-1.50 (1.55-1.50)
Observed reflections	416133 (36794)
Unique reflections	39347 (3683)
Multiplicity	10.6 (9.9)
Completeness (%)	98.87 (94.41)
R_{merge}	0.277 (1.508)
R_{meas}	0.291 (1.590)
R_{pim}	0.087 (0.492)
$\langle I/\sigma I \rangle$	5.65 (1.12)
$CC_{1/2}$	0.989 (0.310)



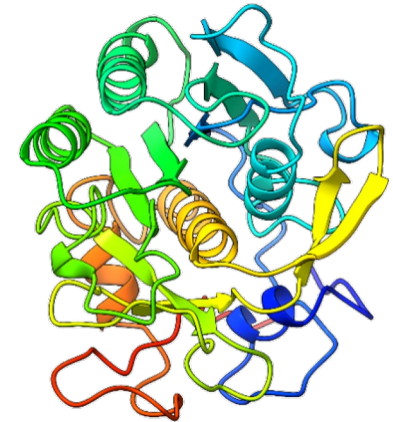
Ab initio structure determination of proteinase K at 1.5 Å resolution



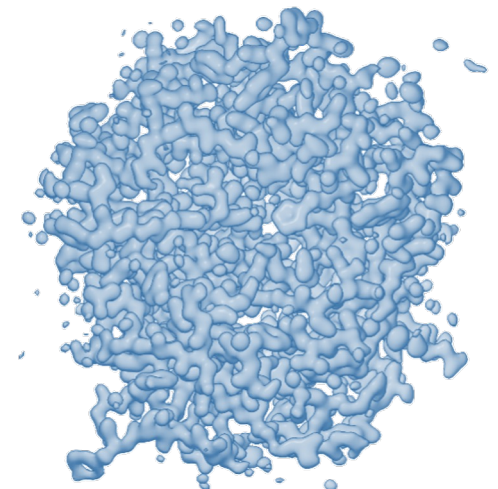
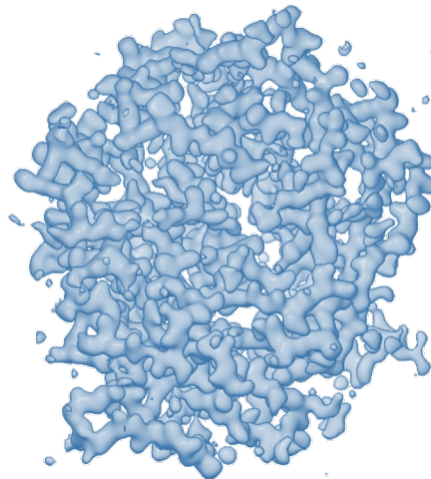
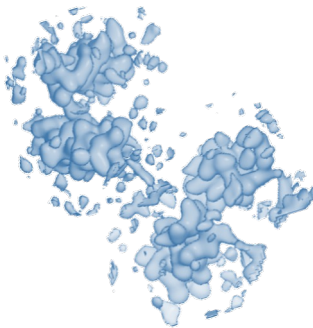
Placing fragments using *PHASER*



Chain tracing and density modification in *SHELXE*



Automated model building using *BUCCANEER/REFMAC5*



Acknowledgements

Our Lab

Purna Vasireddy
Tim Low-Beer
Sarah Kohne
Veronica Nickerson

Collaborators

Tamir Gonen, UCLA
Johan Hattne, UCLA
William Nicolas, Exelixis
Guanhong Bu, UCLA
Sara Weaver, Amgen
Hosea Nelson, CalTech
Brian Stoltz, CalTech

Jieye Lin, UCLA
Yasmeen Ruma, UCLA
Callie Saeher, UCLA
Anna Shiriaeva, UCLA
Johan Unge, Umea
Max Clabbers, Aarhus
Wei Zhao, CalTech
Michael Ostlie, NDSU

Brent Nannenga, ASU
Jeff Abramson, UCLA
Farha Khan, UCLA
Vadim Cherezov, USC
Jose Rodriguez, UCLA
Xuanrui Ge, USC
M. Jason de la Cruz, MSKCC
Grant Jensen, BYU
Hong Zhou, UCLA
Polly Roy, LSH
Xian Xia, UCLA
Po-Yu Sung, UCLA
Katherine Spoth, HWI
Devrim Acehan, HWI
Matthew Crawley, UB
Pierre Le Magueres, Rigaku
Joseph Ferrara, Rigaku
Jessica Burch, Rigaku

Michael W. Martynowycz

Assistant Professor
University at Buffalo
mmartyno@buffalo.edu
@Sciencemick
Martynowyczlab.com
630.415.8495



NIH 1P41GM136508



DTRA MCDC-2202-002



**Community Foundation
For Greater Buffalo**