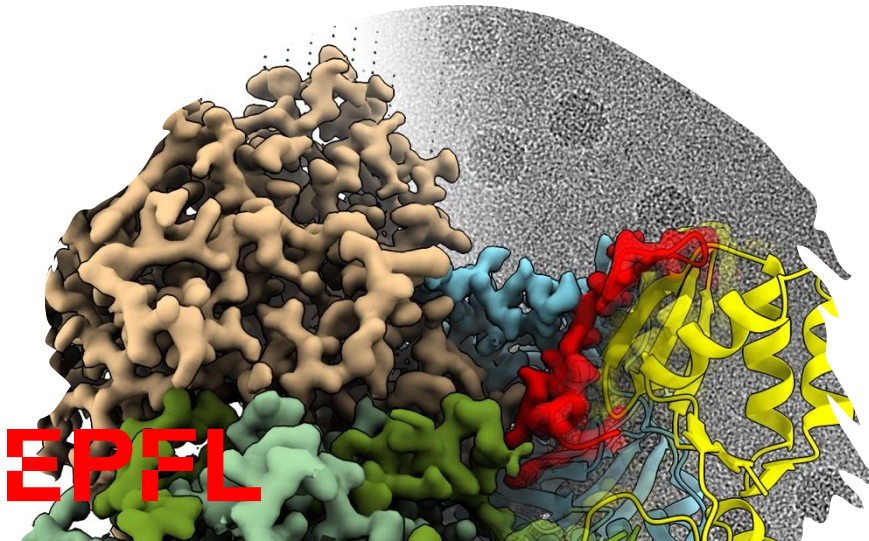


# EDMS BIO-643: Integrative Structural Biology for Life Sciences

- Luciano Abriata
- Kelvin Lau
- Yoan Duhoo
- Anna-Sophia Krebs
- Jonathan Schneider
- Florence Pojer
- Guest: Henning Stahlberg

**Protein Production and Structure Core Facility,  
SV-PTPSP**

**Fall Semester 2025**



## EDMS BIO-643

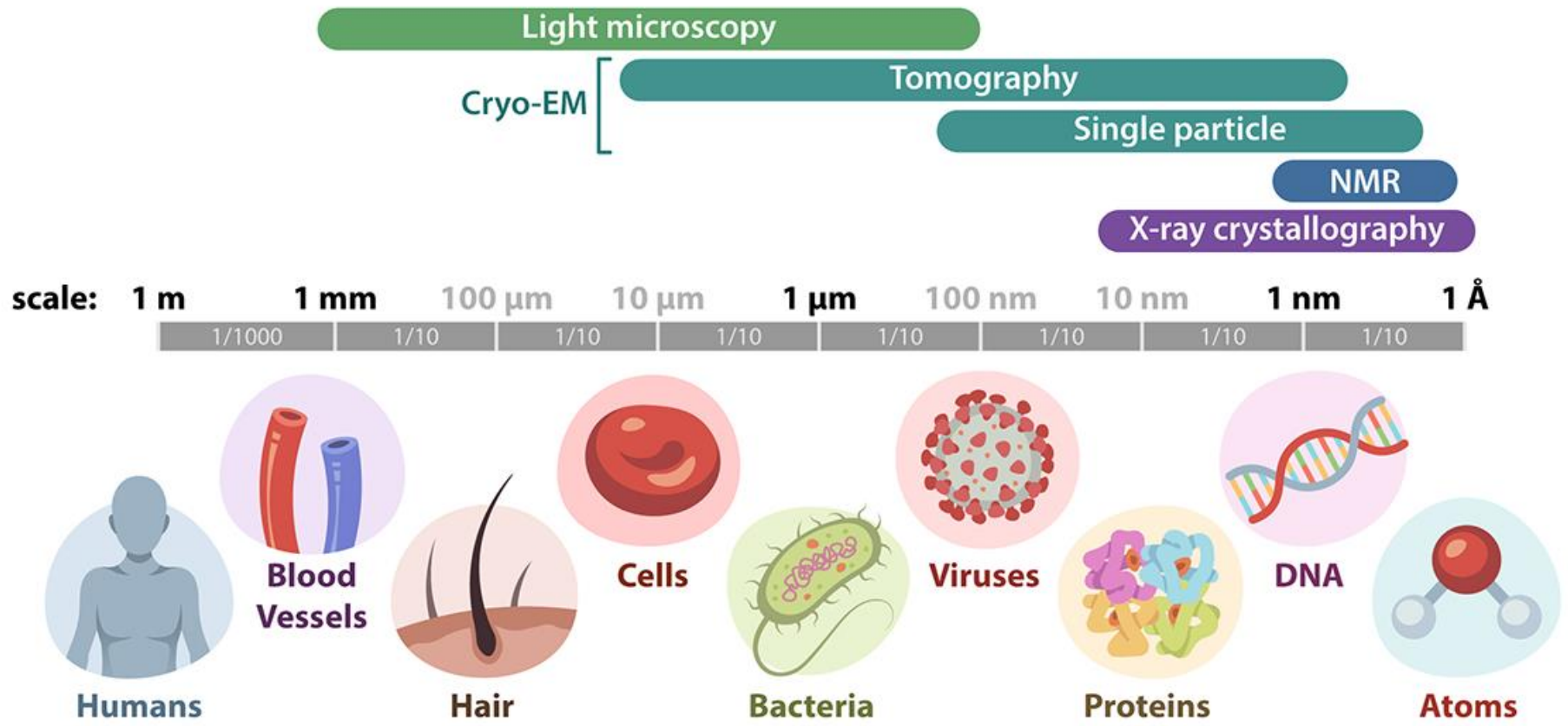
### Integrative structural biology for Life sciences

Fall semester 2025

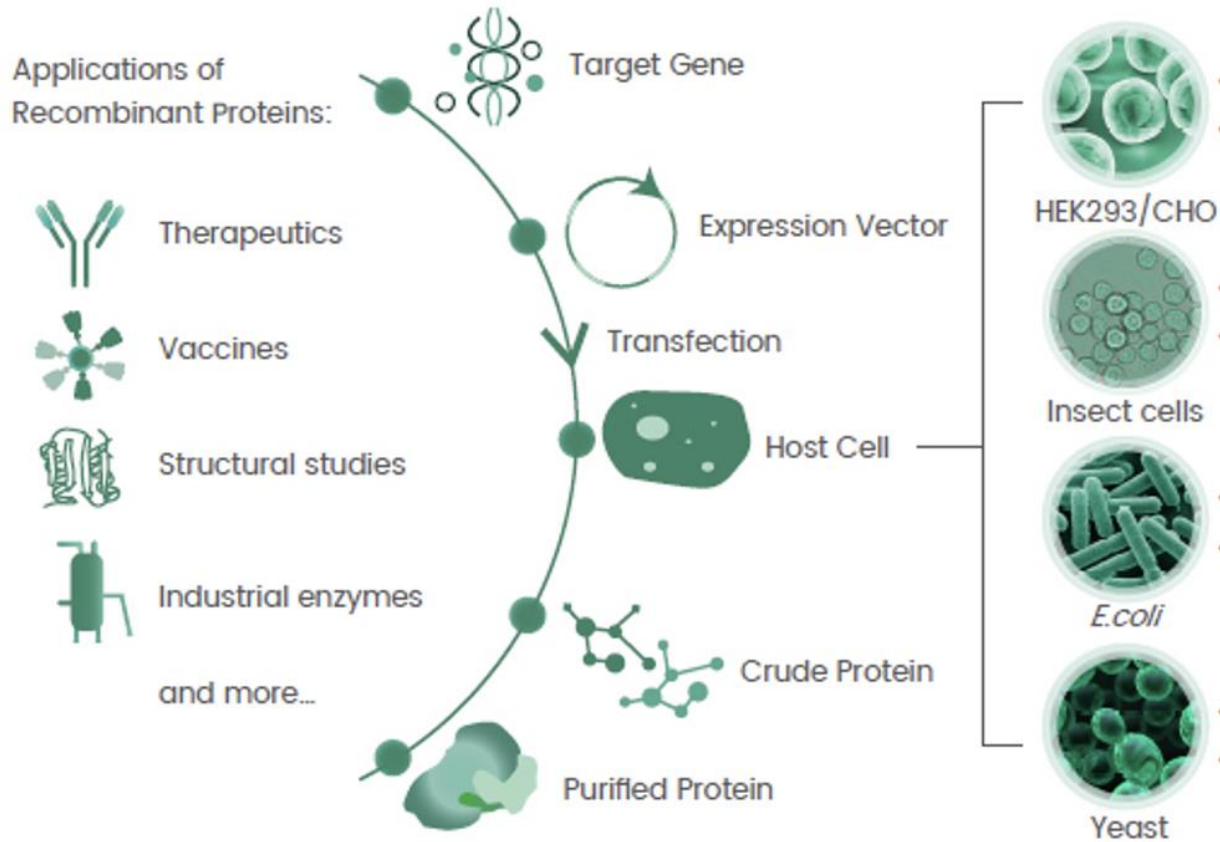
Date/Time: Thursdays 3:15 to 5pm

Location: AI 3142

Date	Topics	Speakers
Sept 11th	Course introduction	All teachers
Sept 18th	Modeling tools	Luciano
Sept 25th	Modeling tools -visualization in ChimeraX	Luciano, Yoan, Kelvin
Oct 2nd	X-ray crystallography theory + software	Florence/Kelvin
Oct 9th	no class	
Oct 16th	X-ray software	Kelvin/Yoan
Oct 23rd	no class (EPFL break)	
Oct 30th	X-ray software	Kelvin/Florence
Nov 6th	cryoEM theory	Henning
Nov 13th	cryoEM software	Yoan, Anna-Sophia, Jonathan
Nov 20th	cryoEM-ET software	Yoan, Anna-Sophia, Jonathan
Nov 27th	cryoEM-ET software + visit DCI	Yoan, Anna-Sophia, Jonathan
Dec 4th	Bio-NMR theory + software + visit	Luciano, Kelvin
Dec 11th	Students presentations	All teachers
dec 18th	Students presentations	All teachers



# Prerequisite for all techniques: a good sample



# Useful resources in protein sciences

## □ UniProt

Very informative and up to date: Function, Names & Taxonomy, Subcellular location, Pathology & Biotech, Interaction, Structure, Family & Domains, Sequence, ...  
AlphafoldDB included now in UniProt – models are to take with caution

## □ Protparam

Essential for biophysical parameters: Number of amino acids, Molecular weight, Theoretical PI, Amino acid composition, Atomic composition, Extinction coefficients.  
<https://web.expasy.org/protparam/>

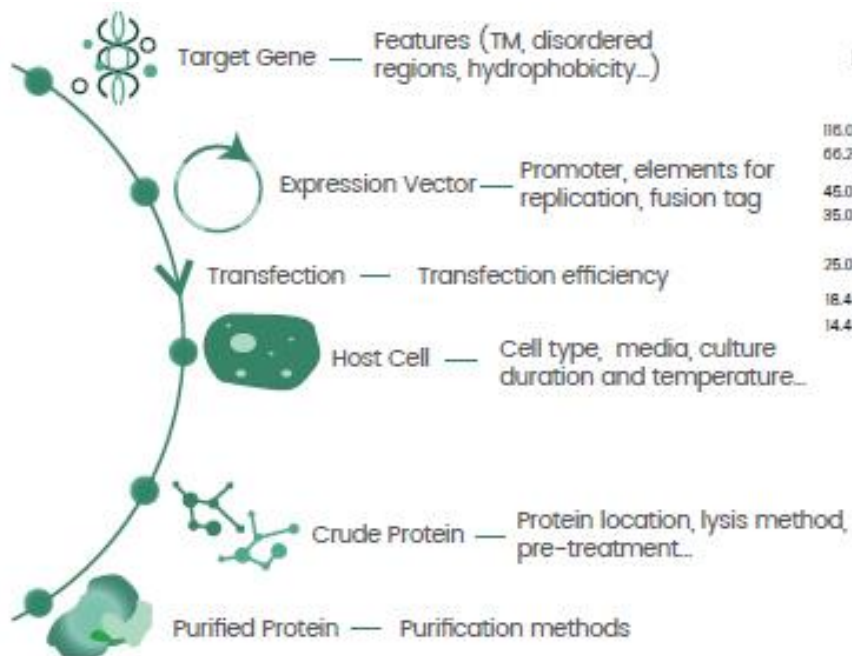
## □ Alphafold, SWISSMODEL and I-TASSER

Useful for quick 3D modelling of any proteins based on their AA sequence  
<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>  
<https://swissmodel.expasy.org>  
<https://zhanglab.ccmb.med.umich.edu/I-TASSER/>

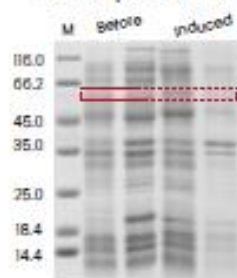
## □ HHPRED:

Useful for unknown function. It is a sequence database searching and structure prediction:  
<https://toolkit.tuebingen.mpg.de/tools/hhpred>

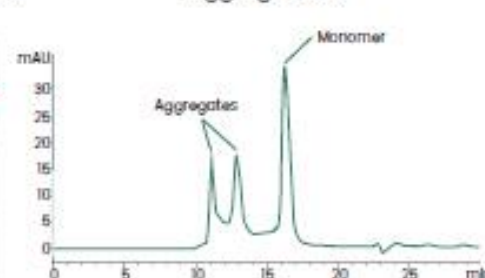
# Challenges of a good sample



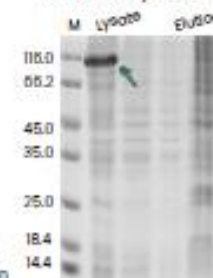
Low expression level



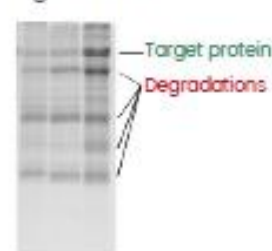
Aggregations



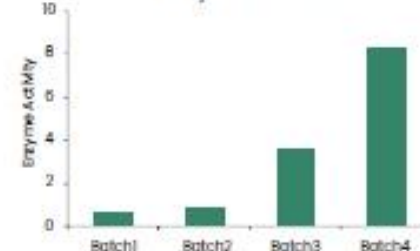
Insoluble protein



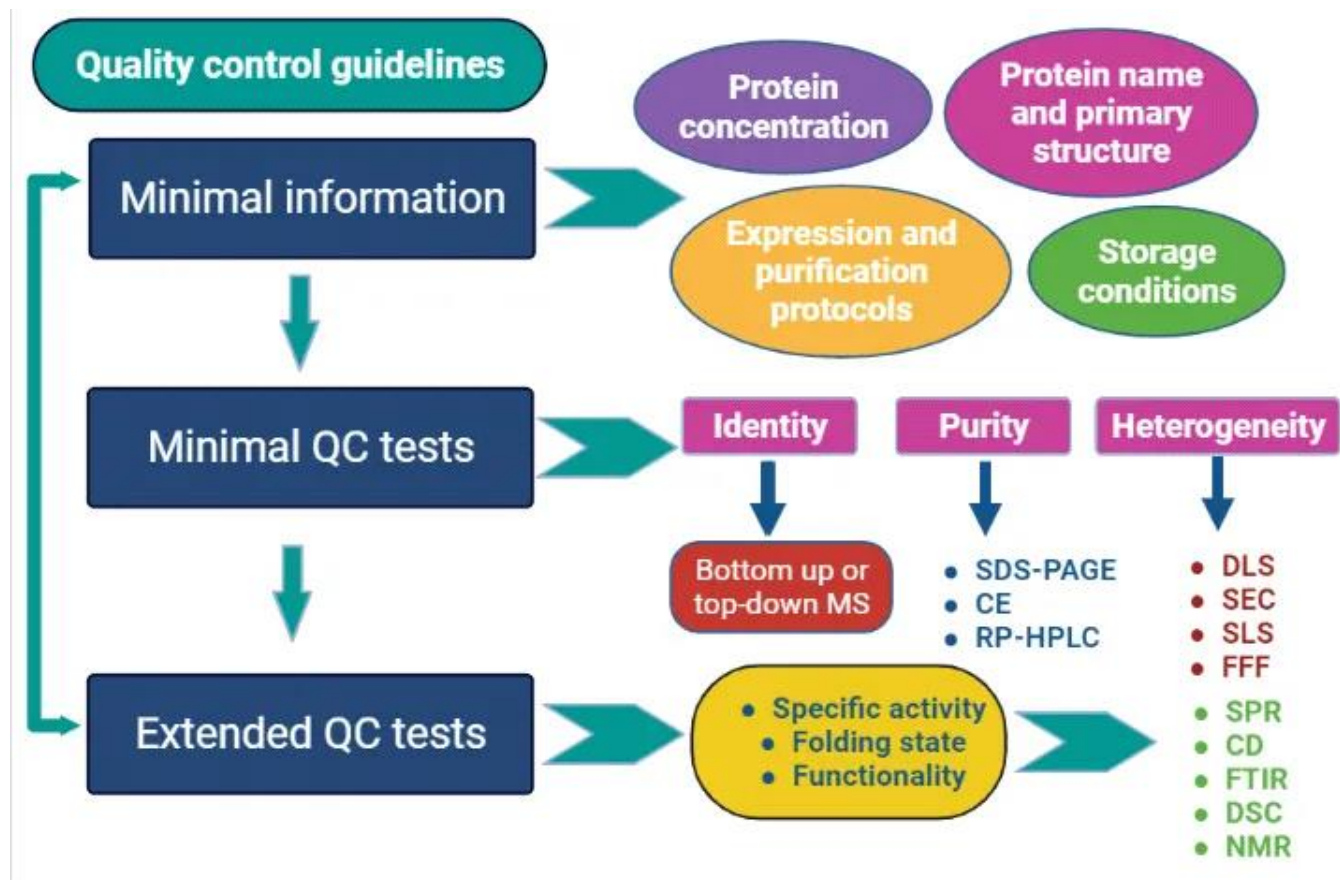
Degradation



Activity variations



# Essentiel to perform Quality Controls (QC)



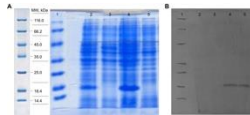
# Tools for Biophysical characterization at PTPSP

## Aktas (Cytiva)



Oligomeric state

## (SDS)-PAGE + Western blotting



Oligomeric state, purity, identity

## DLS Stunner (Unchained Bio)



Concentration, oligomer state,  
Purity Free

## SPR: Biacore 8K (Cytiva)



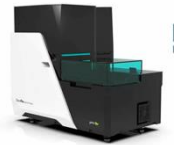
Binding assays and  
kinetics

## MicroCal PEAQ-ITC (Malvern)



Binding assays and  
thermodynamics

## BLI (GatorBio)



Binding assays and kinetics

## Mass photometer (Refeyn)



Purity, homogeneity,  
Oligomeric state

## WAVE system (Creoptix)



Grating-coupled  
interferometry (GCI)

## Uncle (Unchained Bio)



Stability, screening

## Chirscan V100 CD



Secondary structure,  
folding and stability

## Fida-1 (Fidabio)



Binding assays and kinetics

Quality control

Protein-protein/drug interactions for drug discoveries

EPFL

# Biophysical techniques

## Speed Dating

Meet 9 Advanced Instruments  
in a 1-Hour Private Tour at PTPSP

- Surface Plasmon Resonance (SPR)
- Isothermal Titration Calorimetry (ITC)
- BioLayer Interferometry (BLI)
- Grating-coupled Interferometry (GCI)
- Flow Induced Dispersion Analysis (FIDA)
- Mass Photometer (MP)
- Differential Scanning Fluorimetry (DSF)
- Circular Dichroism Spectrometer (CD)
- Dynamic Light Scattering and UV (DLS)



**Kelvin Lau**  
Scientist, Protein production  
and structure Core Facility



**Luciano Andres Abriata**  
Scientist, Protein production  
and structure Core Facility



**Yoan Duhoo**  
Scientist, Protein production  
and structure Core Facility

More infos & Sign-up

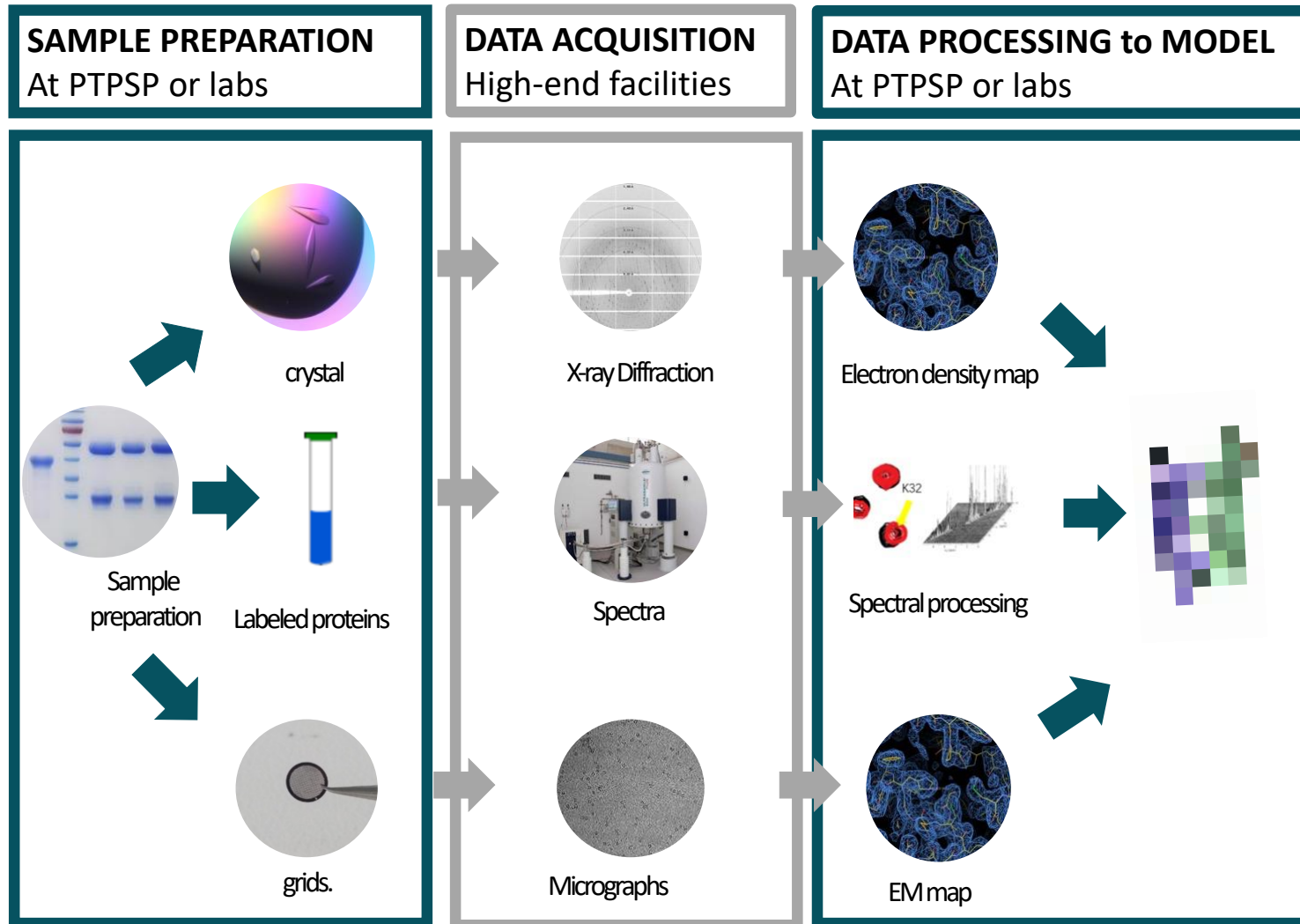
School  
of Life  
Sciences



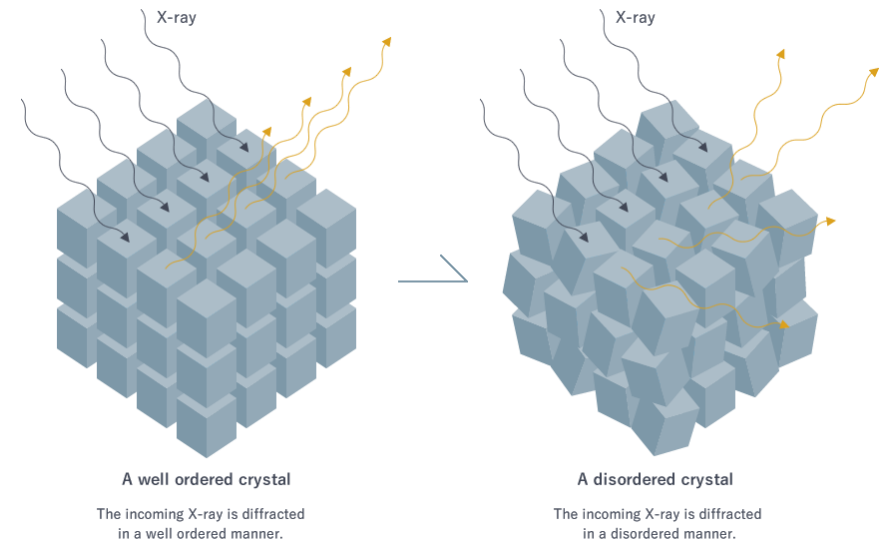
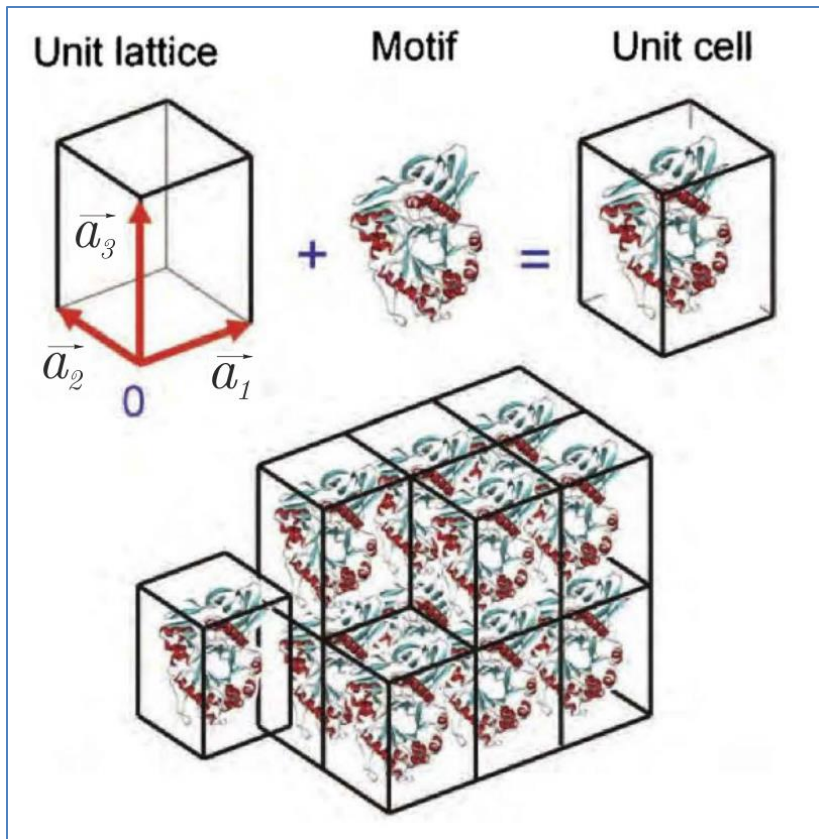
PTPSP

- Dates are available 4-6 times per year upon registration on the PTPSP website
- Private lab tours can be arranged upon request

# Experimental techniques for atomic structural studies



# Sample quality is key to form crystals and for a good diffraction



# Success of X-ray crystallography

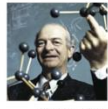
Highlights of the Many Nobel Prizes **Awarded to Crystallographers**

See a complete list of winners at [iucr.org/people/nobel-prize](http://iucr.org/people/nobel-prize)



**Wilhelm Röntgen**  
Discovery of X-rays

1901



**Linus Pauling**  
Alpha-helical structure of proteins, nature of chemical bonds

1954



**Francis Crick, James Watson & Maurice Wilkins**  
Created DNA model: double-helical structure for biological information storage

1962



**Herbert Hauptman & Jerome Karle**  
Direct mathematical methods of determining crystallized materials

1985



**Clifford Shull & Bertram Brockhouse**  
Electron diffraction and neutron diffraction

1994



**Venki Ramakrishnan, Tom Steitz & Ada Yonath**  
Studies of the structure and function of the ribosome

2009



**Dan Shechtman**  
Discovery of quasicrystals

2011



**Jean-Pierre Sauvage, J. Fraser Stoddart & Ben Feringa**  
Design and synthesis of molecular machines

2016



1914



**Max von Laue**  
First demonstrated X-ray diffraction through crystals



1915



**Sir William H. & Sir William L. Bragg**  
First atomic crystal structure



1962



**John Kendrew & Max Perutz**  
Hemoglobin transport protein, which led to the understanding of Sickle Cell Anemia



1964



**Dorothy Hodgkin**  
Structures of cholesterol, penicillin, vitamin B12, and insulin



1976



**William Lipscomb**  
The structure of boranes, illuminating problems of chemical bonding



1988



**Johann Deisenhofer, Robert Huber & Hartmut Michel**  
First membrane protein that is essential to photosynthesis



2003



**Peter Agre & Roderick MacKinnon**  
Discoveries concerning channels in cell membranes



2006



**Roger Kornberg**  
Studies of the molecular basis of eukaryotic transcription



2013



**Martin Karplus, Michael Levitt & Arieh Warshel**  
Development of sophisticated computer simulations for complex chemical processes



## Additional Important Contributors to Crystallography



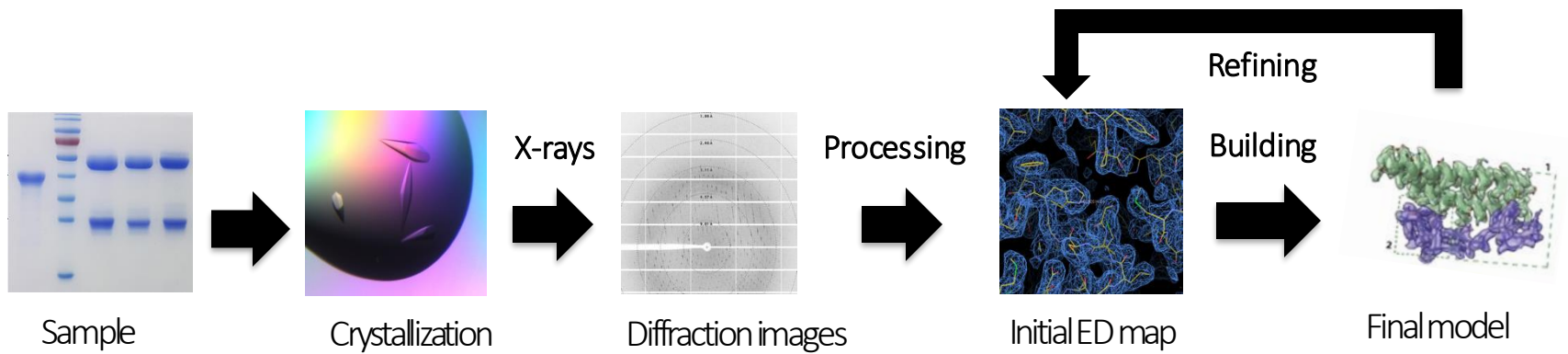
**Arthur Patterson**  
The Patterson Function (equation) gives a map of the vectors between atoms



**David Harker**  
Applied Patterson's map to identify planes and sections on different axes in molecular structures

PDB database In Oct 2025:

198 063 models obtained by X-ray crystallography out of 242 874 models (cryoEM: 29 460 models)

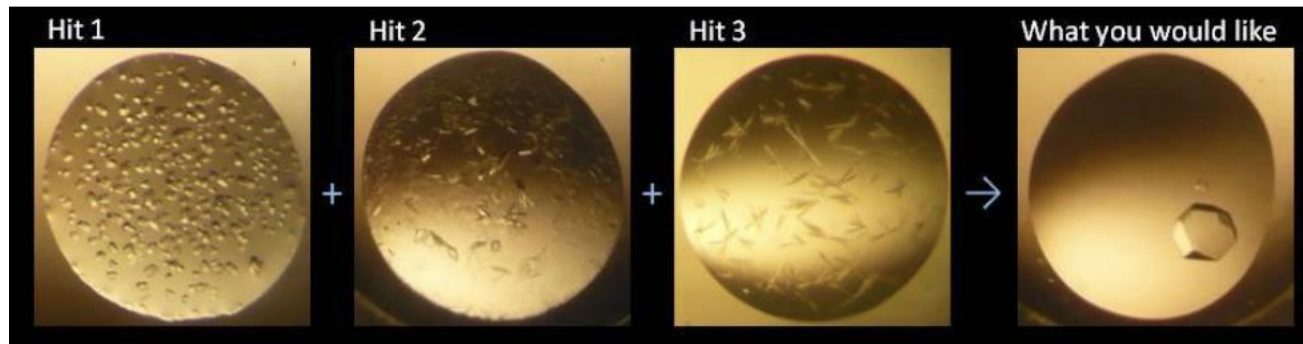


## Prerequisites in sample preparation:

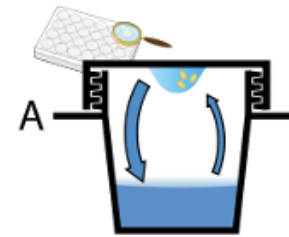
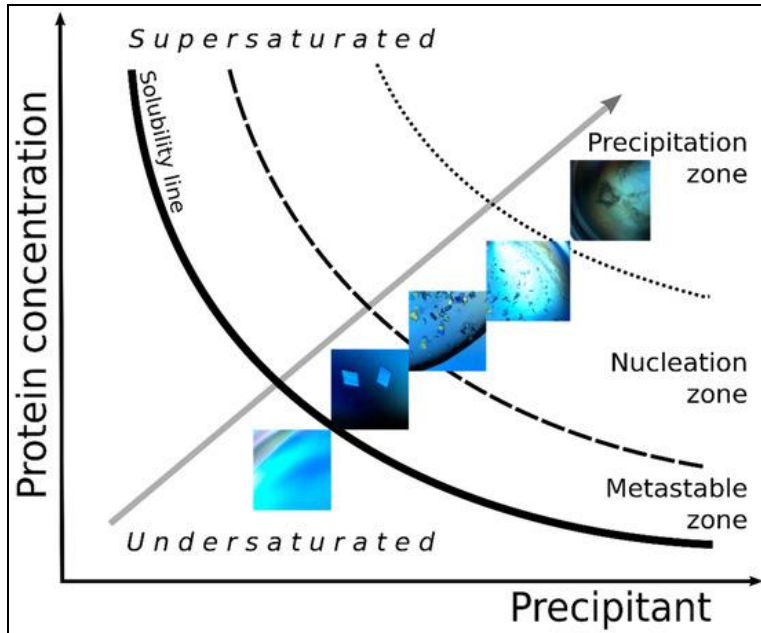
- Essential to **design suitable construct**, to obtain stable and homogeneous protein or complex (e.g.; add binding partners, co-factors, small molecules Or remove flexible domains Or focus on certain domains)
- High-purity and quantity** of your protein of interest or complex: Minimum two steps purification (e.g. affinity followed by size exclusion chromatography)
- Need around 120µl at 10mg/ml to screen around 400 conditions
- No phosphate buffer, to reduce formation of salt crystals

## Steps for crystallization:

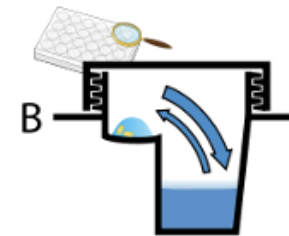
- **Screening** of crystallization conditions using commercial kits with robots (variation of salt, precipitant, pH, temperature, protein concentration,...)
- Incubation of plates at 18C or 4C
- Checking of **crystals formation under light microscope**  
Crystals form in days to months; screening success rate is 0-10% depending on proteins
- **Optimization of crystals** to single crystals to obtain suitable diffraction pattern and improve resolution



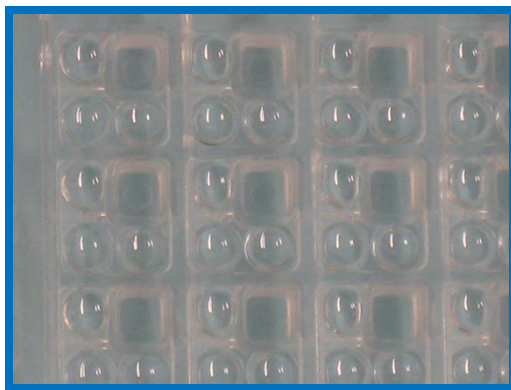
# Vapor diffusion techniques



Hanging drops



Sitting drops

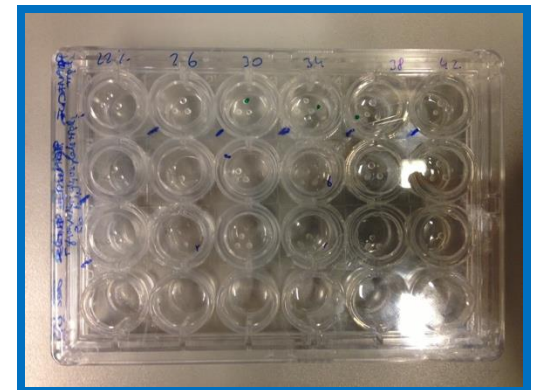


**Sitting drops vapor diffusion**

Mix of precipitants and protein  
(total of 0.2ul to 1ul)

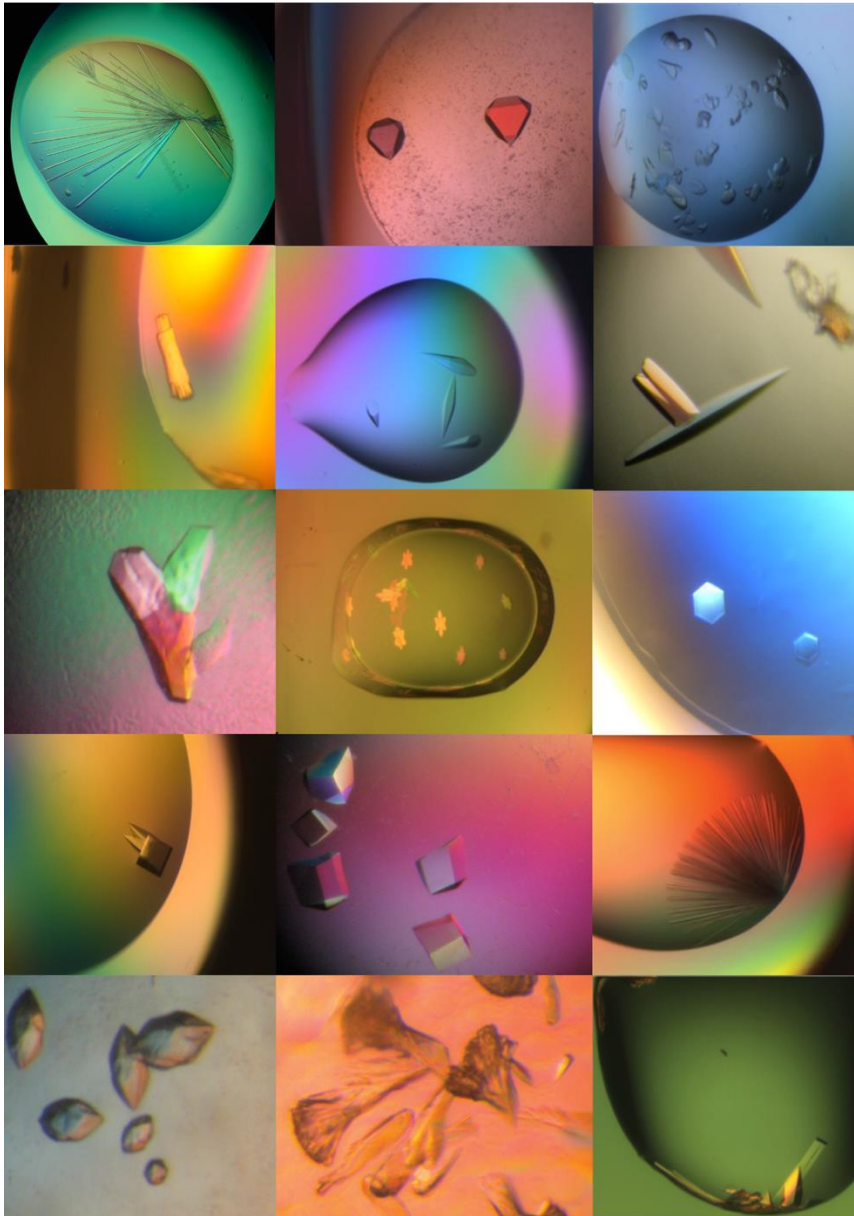


Mosquito (STP labtech)



**Hanging drops vapor diffusion**

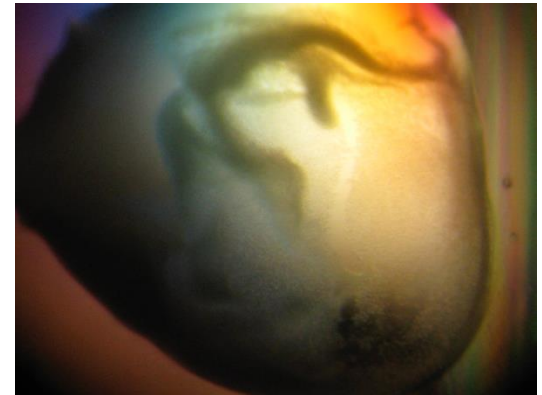
mix precipitants and protein  
(total of 1ul to 4ul)



Crystals can have many shapes and sizes

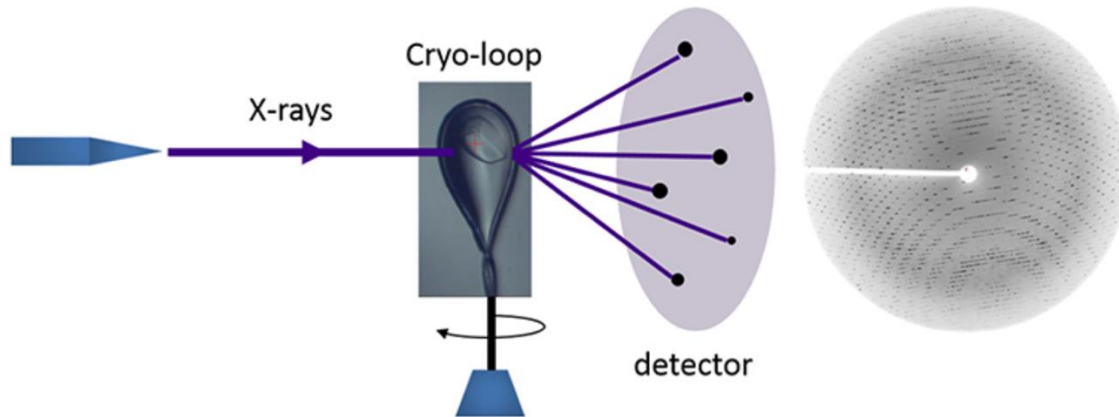


Observation under microscope  
(Automated instrument on market also)



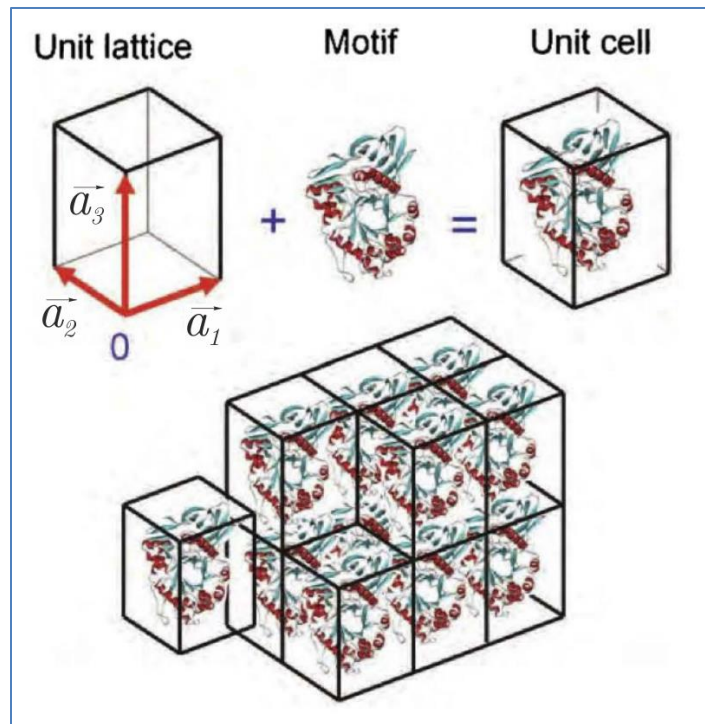
Heavy protein precipitation

# Why X-rays?



**Why X-rays?** Their wavelength is of the order of the angström and are comparable to the distances between atoms. This wavelength allows the X-rays to interact with the crystal's atomic structure, producing diffraction patterns that provide information about the precise arrangement and spacing of the atoms.

# Why forming crystals?

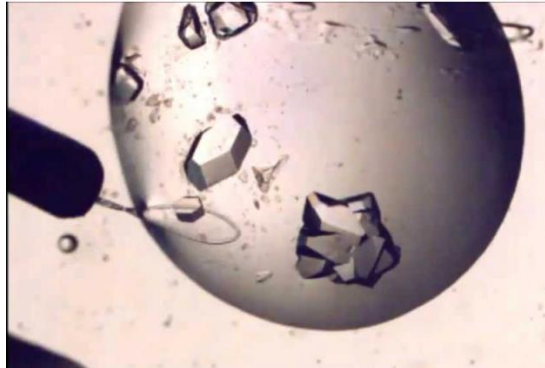


**Why forming crystals?** A crystal arranges huge numbers of molecules in the same orientation, so that scattered X-ray waves can add up in phase and raise the signal to a measurable level.

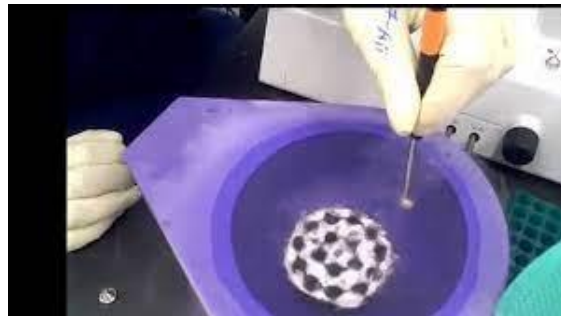
A crystal acts as an amplifier.

# Preparation of crystals prior data collection

## Cryoprotecting and Freezing



Add 25% glycerol (or other protectants) to drop before fishing



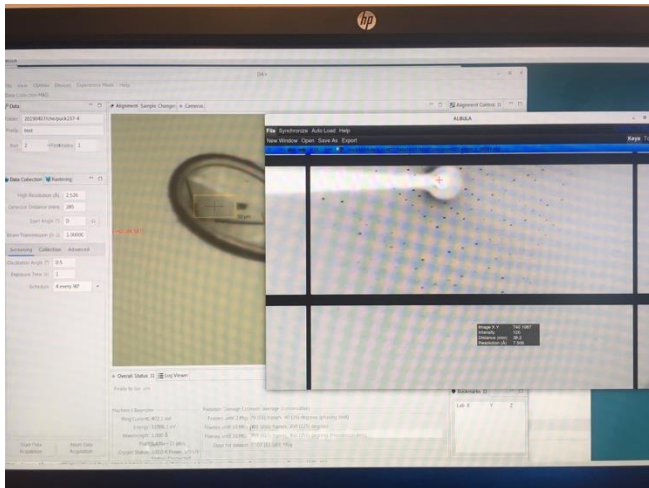
16 crystals per puck- 7 pucks per dewar-  
Total of 112 crystals per dewar

# Data collection at synchrotrons



Data collected at Swiss Light Source (SLS-PSI, Villigen) or ESRF (Grenoble)

Beamtime is free for Academic labs: application once per year for beamtime via an academic proposal

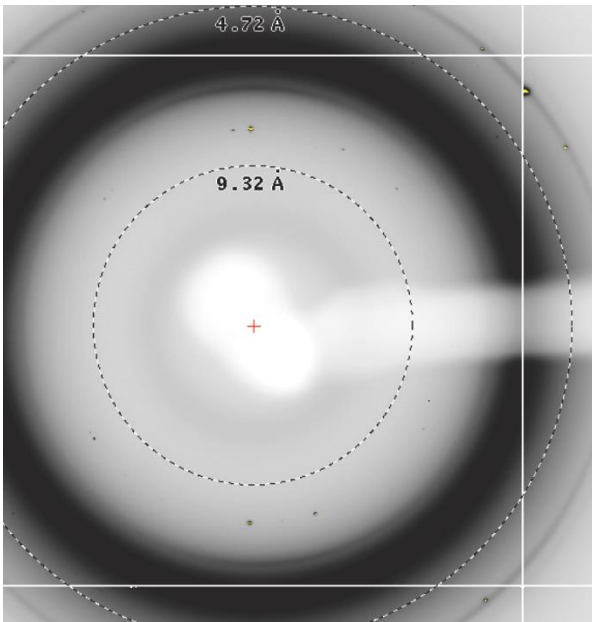


Data are collected automatically and often also pre-processed

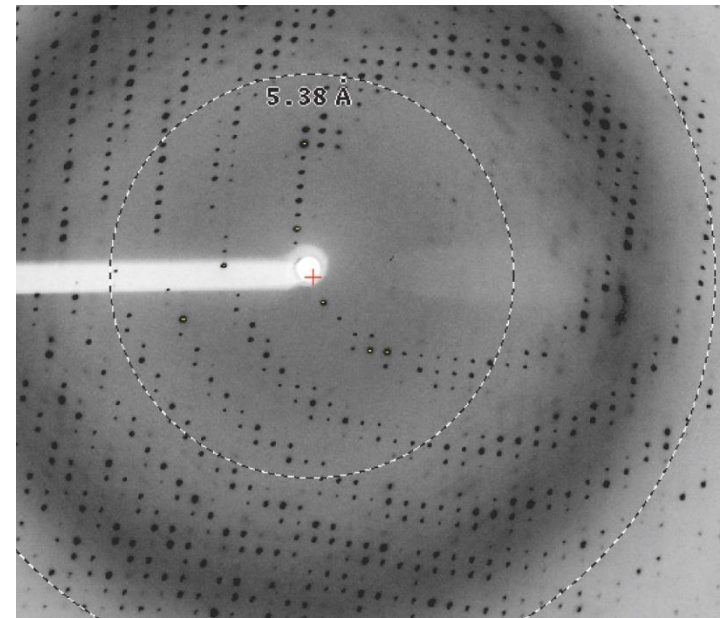
Data can also be collected remotely

# Data collection at synchrotrons

- For a full data collection, images are collected by rotating the crystal (slices of  $0.1\text{-}0.2^\circ$  for a total of  $360^\circ$  for crystal with low symmetry)
- The total number of images to collect depends on the symmetry of molecules in the crystal
- A complete data collection takes only 2-3 minutes (no strategy needed)
- High thought-put method



X-ray diffraction of crystal of salt!

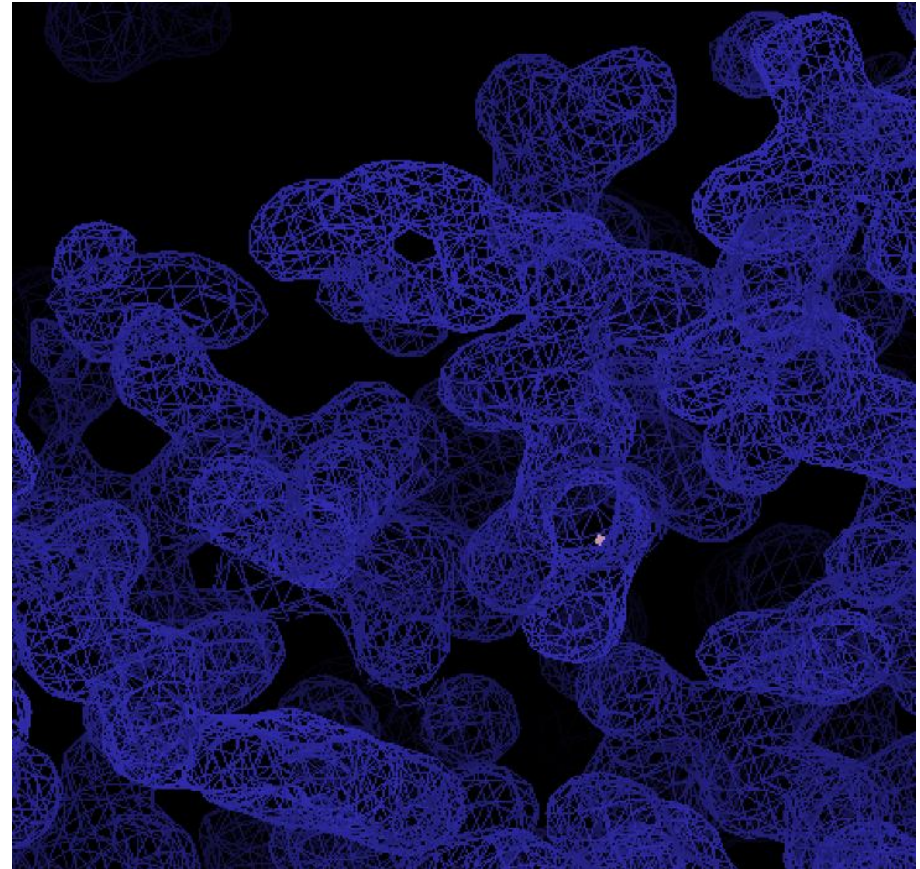
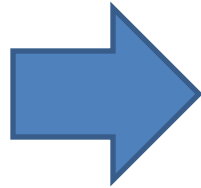
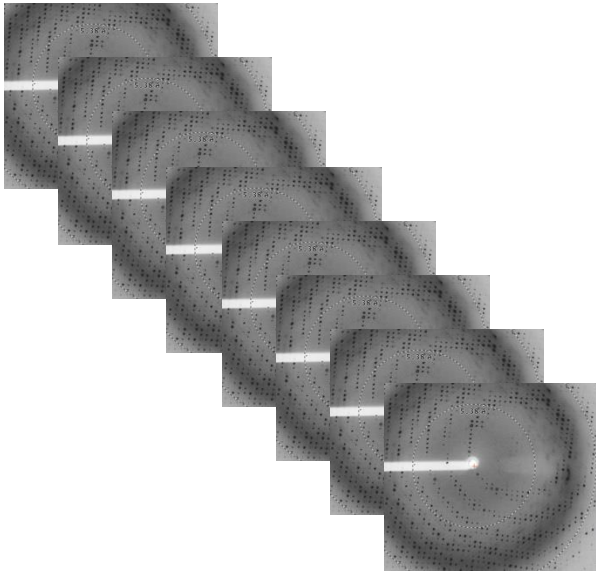


X-ray diffraction of protein crystal (one 2D image)

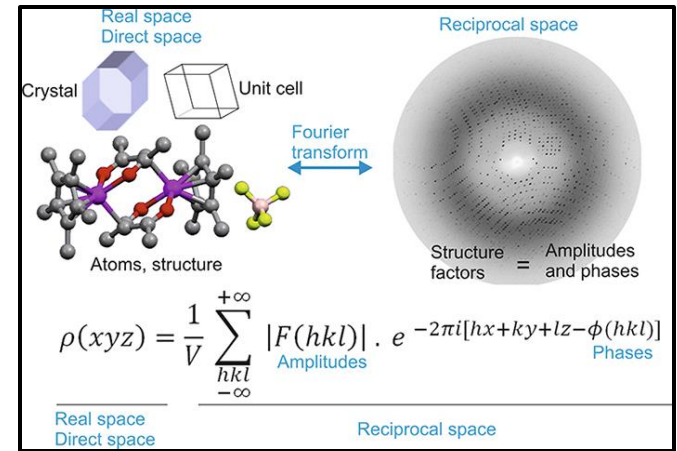
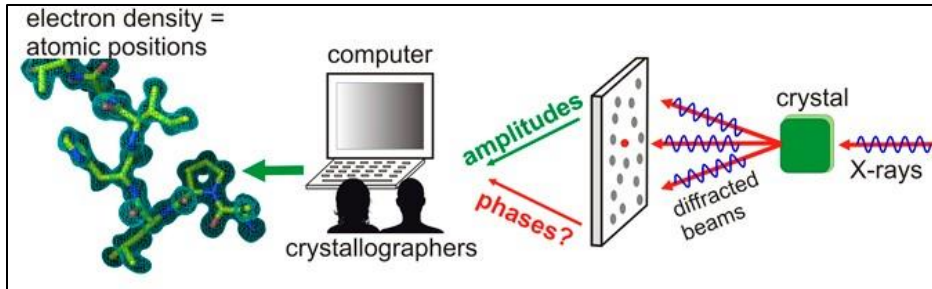
ID	Protein	design	Type	Res. (corner)	OSC	Wavelength	Images	Plot	Resolution	Completeness	Rmerge	a	b	c	Autoprocessing	SAD
#5 OSC 2025-09-05 16:35:37	design-dPTE-wt_p2_custom-H10-3	design	Res. (corner)	3.16 Å (2.26 Å)	0.9677 Å	9313		P 21 21 2	Overall: 62.11-2.56 Inner: 62.11-8.31 Outer: 2.81-2.56	79.2% 93.0% 99.5%	32.8 18.7 45.3	62.41 Å a 90 *	69.32 Å β 90 *	139.76 Å c Y 90 *	Autoprocessing	
#5 OSC 2025-09-05 16:30:31	design-dPTE-26pMH-p1-custom-F	design	Res. (corner)	1.78 Å (1.36 Å)	0.9677 Å	3166		C 1 2 1	Overall: 45.97-1.76 Inner: 45.97-6.82 Outer: 1.82-1.76	99.0% 99.6% 99.3%	18.9 4.7 421.2	444.66 Å a 90 *	54.05 Å b 100.58 *	84.79 Å c Y 90 *	Autoprocessing	
#5 OSC 2025-09-05 16:24:39	design-dPTE-26pMH-p1-custom-F	design	Res. (corner)	1.81 Å (1.38 Å)	0.9677 Å	8227		C 1 2 1	Overall: 73.68-1.63 Inner: 73.68-8.95 Outer: 1.66-1.63	95.9% 89.2% 87.0%	9.5 3.7 122.8	444.18 Å a 90 *	52.55 Å b 100.61 *	85.06 Å c Y 90 *	Autoprocessing	SAD
#2 Mesh 2025-09-05 16:22:04	mesh-design-2-dPTE-26pMH-p1-C	design	Res. (corner)	2.00 Å (1.50 Å)	0.9677 Å	550										
#5 OSC 2025-09-05 16:17:06	design-dPTE-26pMH-p1-GS083-G	design	Res. (corner)	1.17 Å (0.98 Å)	0.9677 Å	3198		C 1 2 1	Overall: 42.62-0.98 Inner: 42.62-5.27 Outer: 1.00-0.98	75.1% 97.6% 8.4%	8.2 7.2 6.9	15.63 Å a 90 *	170.48 Å b 129.68 *	14.05 Å c Y 90 *	Autoprocessing	SAD
#5 OSC 2025-09-05 16:11:38	design-dPTE-26pMH-p1-custom-H	design	Res. (corner)	1.70 Å (1.31 Å)	0.9677 Å	4777		C 1 2 1	Overall: 72.38-1.93 Inner: 72.38-10.59 Outer: 1.97-1.93	98.0% 99.1% 91.6%	13.4 4.1 137.1	442.24 Å a 90 *	57.90 Å b 100.87 *	84.41 Å c Y 90 *	Autoprocessing	
#5 OSC 2025-09-05 16:05:26	design-dPTE-wt_p1_custom-F3-2	design	Res. (corner)	2.26 Å (1.66 Å)	0.9677 Å	8072		C 1 2 1	Overall: 72.99-2.15 Inner: 72.99-9.12 Outer: 2.21-2.15	96.7% 99.6% 90.9%	17.9 5.9 83.8	85.03 Å a 90 *	52.89 Å b 90.35 *	218.98 Å c Y 90 *	Autoprocessing	
#5 OSC 2025-09-05 15:59:54	design-dPTE-wt_p1_custom-G7_1	design	Res. (corner)	2.39 Å (1.75 Å)	0.9677 Å	8092		C 1 2 1	Overall: 109.30-2.21 Inner: 109.30-9.10 Outer: 2.27-2.21	96.1% 90.4% 87.4%	22.2 4.8 108.3	85.44 Å a 90 *	52.69 Å b 90.54 *	218.82 Å c Y 90 *	Autoprocessing	
#5 OSC 2025-09-05 15:54:30	design-dPTE-wt_p3_custom-G10-2	design	Res. (corner)	1.43 Å (1.14 Å)	0.9677 Å	2988		P 21 21 21	Overall: 46.90-1.67 Inner: 46.90-6.48 Outer: 1.73-1.67	99.7% 99.7% 97.9%	21.1 8.2 215.9	63.36 Å a 90 *	69.76 Å b 90 *	153.76 Å c Y 90 *	Autoprocessing	
#2 Mesh 2025-09-05 15:49:35	mesh-design-dPTE-wt_p1-GS083-I	design	Res. (corner)	2.00 Å (1.50 Å)	0.9677 Å	1734										

Gui at ESRF beamline:  
Highly automated and easy to look and download collected data

# Data Processing and phasing: from 2D Images to first electron density map

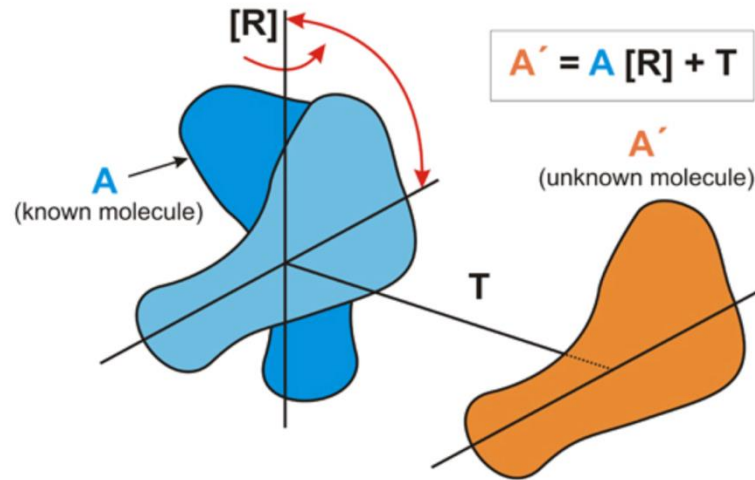


# Data Processing and phasing: from 2D Images to first electron density map



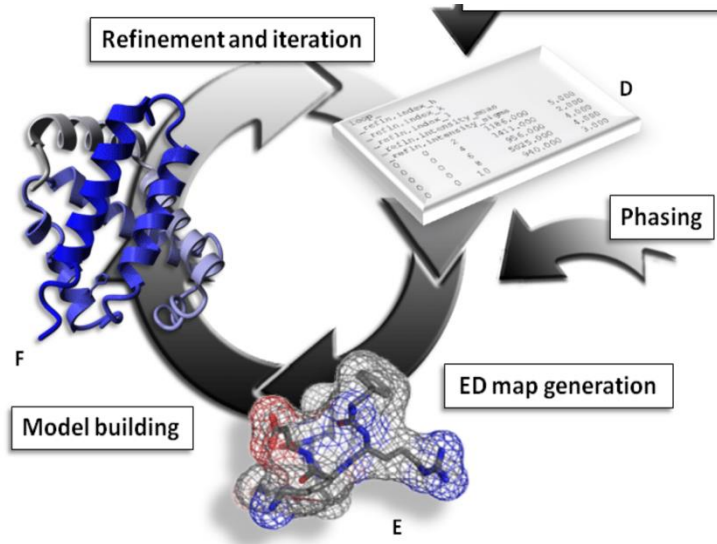
- Software, such as XDS and Mosflm, converts 2D diffraction images in a unique reflection file, to generate a first electron density map.
- Data processing steps include spots finding and autoindexing; Integration (measurement of spot intensities); scaling and merging into one reflection file.
- The phases are still missing and can be estimated by Molecular Replacement or other methods.

# Methods to provide estimates of phases



- **Molecular replacement (MR)**: uses a known, similar model to determine an unknown one by finding the best fit through rotation and translation against experimental data. Tools like AlphaFold now assist in this process and are integrated into X-ray crystallography software.
- **Multi-wavelength Anomalous Dispersion (MAD) or SAD**: Heavy atoms are electron dense and give rise to measurable differences in the intensities of the spots in the diffraction pattern
- **Native SAD**: take use of natural sulfur atoms present in proteins

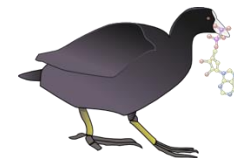
# Building and refining of the initial model



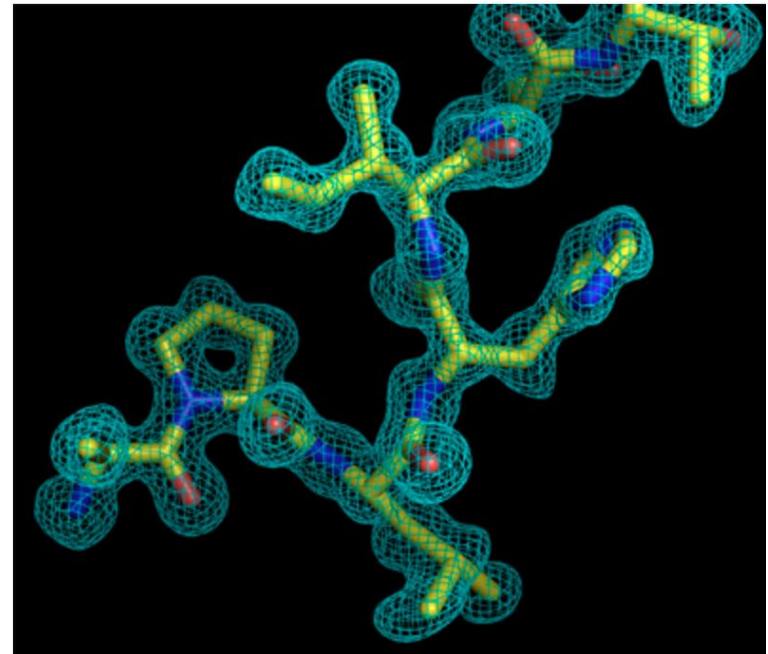
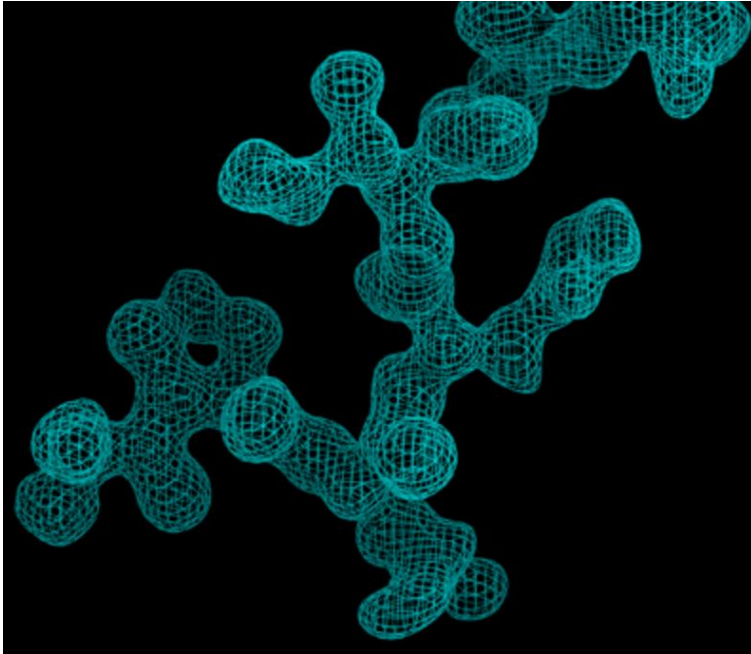
The initial model is usually suboptimal so the model has to be refined. Cycle of manual model building, refinement and new map generation is iterated until the model is considered correct

**Phenix**

**CCP4**



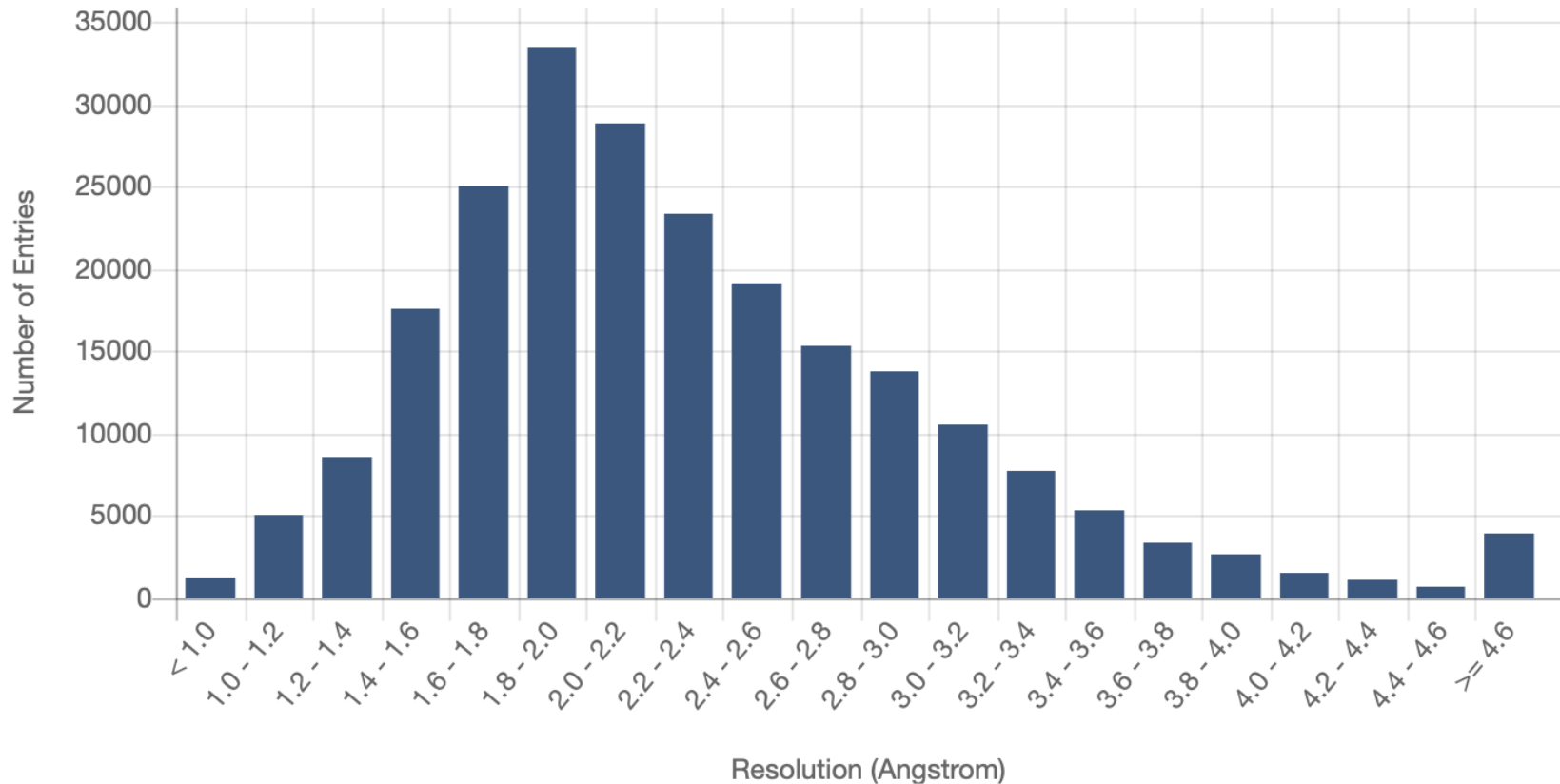
# Goal: the model should fit well the data



- The process can take hours to days or months depending on initial quality of the map and of the model
- Each cycle of manual model building and refinement contributes to the progressive optimization of the electron density map
- Importance of resolution in model building

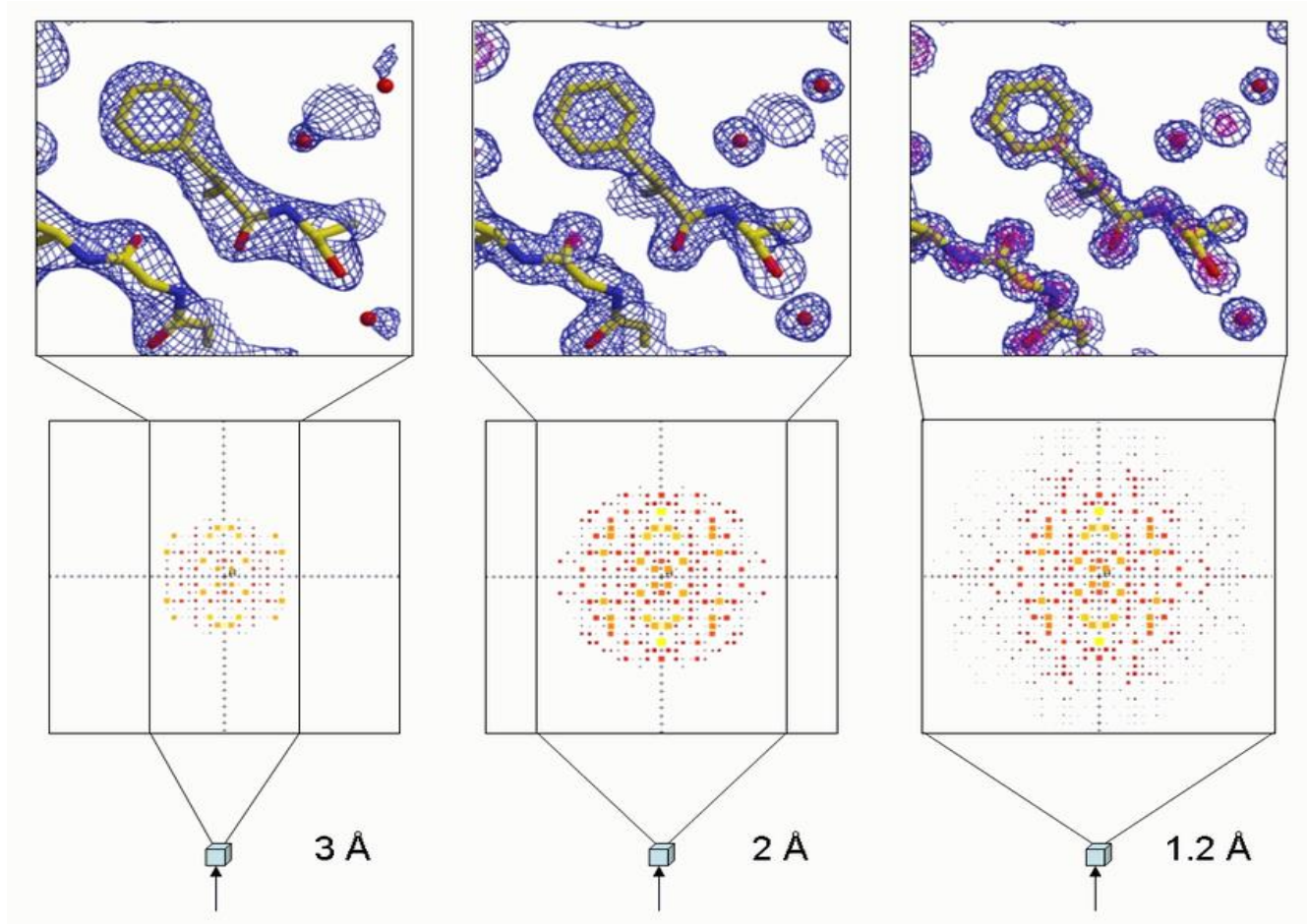
# PDB Statistics:

## PDB Data Distribution by Resolution

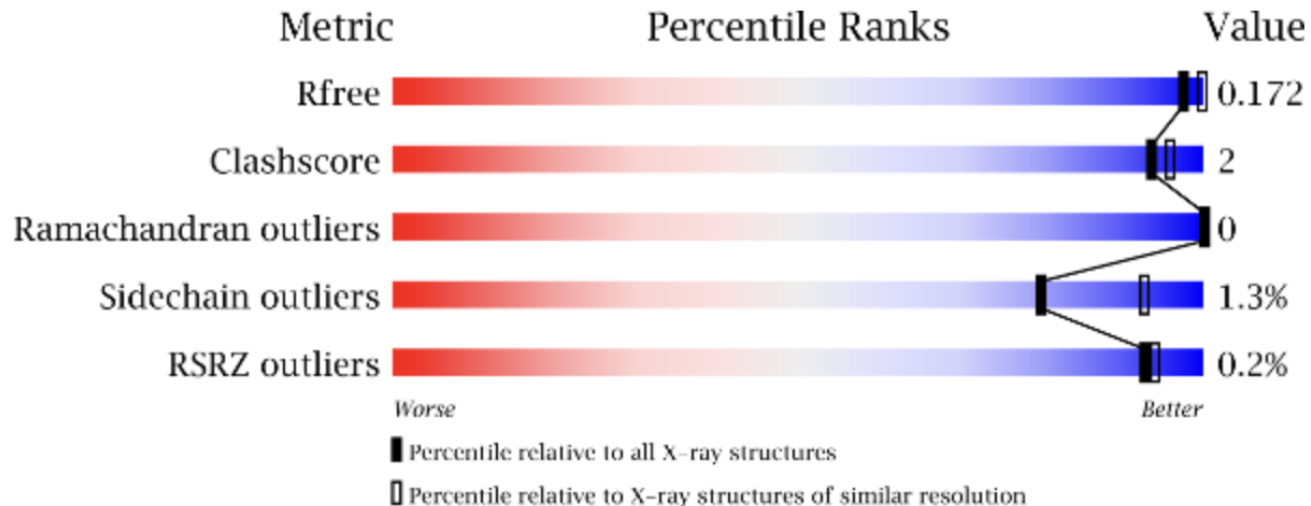


# Electron density maps in function of resolution

[movie](#)



# Validation of final model with metrics (can be very tedious)



*An example slider graphic for a relatively good structure.*

<b>Rfree</b>	This is a measure of the fit of the model to a small subset of the experimental data, which was not used in model refinement (Brünger, 1992). The value quoted is from a recalculation by the DCC program (Yang et al., 2016). Further information can be found in the <i>Data refinement and statistics</i> section of the report, <a href="#">as described below</a> .
<b>Clashscore</b>	This score is derived from the number of pairs of atoms in the model that are unusually close to each other. It is calculated by MolProbity (Chen et al., 2010) and expressed as the number of such clashes per thousand atoms. Further information can be found in the <i>Close contacts</i> section of the report, <a href="#">as described below</a> .
<b>Ramachandran outliers</b>	A residue is considered to be a <a href="#">Ramachandran plot</a> outlier if the combination of its $\phi$ and $\psi$ torsion angles is unusual, as assessed by MolProbity (Chen et al., 2010). The Ramachandran outlier score for an entry is calculated as the percentage of Ramachandran outliers with respect to the total number of residues in the entry for which the outlier assessment is available. Further information can be found in the <i>Torsion angles, Protein backbone</i> section of the report, <a href="#">as described below</a> .
<b>Sidechain outliers</b>	Protein sidechains mostly adopt certain (combinations of) preferred torsion angle values (called rotamers or rotameric conformers), much like their backbone torsion angles (as assessed in the Ramachandran analysis). MolProbity considers the sidechain conformation of a residue to be an outlier if its set of torsion angles is not similar to any preferred combination. The sidechain outlier score is calculated as the percentage of residues with an unusual sidechain conformation with respect to the total number of residues for which the assessment is available.

# Interpretation and making sense of observations

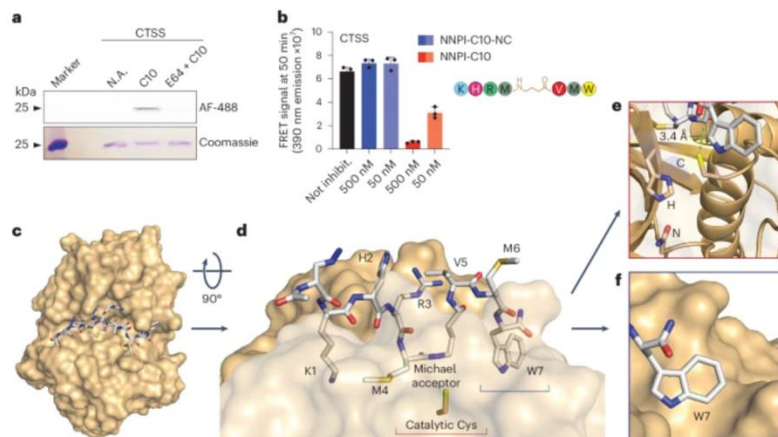
## Deposition in database and publication

> Nat Chem Biol. 2024 Sep;20(9):1188–1198. doi: 10.1038/s41589-024-01627-z.  
Epub 2024 May 29.

### Antibody–peptide conjugates deliver covalent inhibitors blocking oncogenic cathepsins

Aaron Petruzzella<sup>1,2</sup>, Marine Bruand<sup>1,2</sup>, Albert Santamaria-Martínez<sup>1,2</sup>,  
Natalya Katanayeva<sup>1,2</sup>, Luc Reymond<sup>3</sup>, Sarah Wehrle<sup>4</sup>, Sandrine Georgeon<sup>4</sup>, Damla Inel<sup>5,6</sup>,  
Floris J van Dalen<sup>7,8</sup>, David Viertl<sup>9,10</sup>, Kelvin Lau<sup>11</sup>, Florence Pojer<sup>11</sup>,  
Margret Schottelius<sup>9,12</sup>, Vincent Zoete<sup>2,5,6</sup>, Martijn Verdoes<sup>7,8</sup>, Caroline Arber<sup>5,6</sup>,  
Bruno E Correia<sup>13</sup>, Elisa Oricchio<sup>14,15</sup>

**Fig. 2: Covalent binding mode of NNPI-C10.**



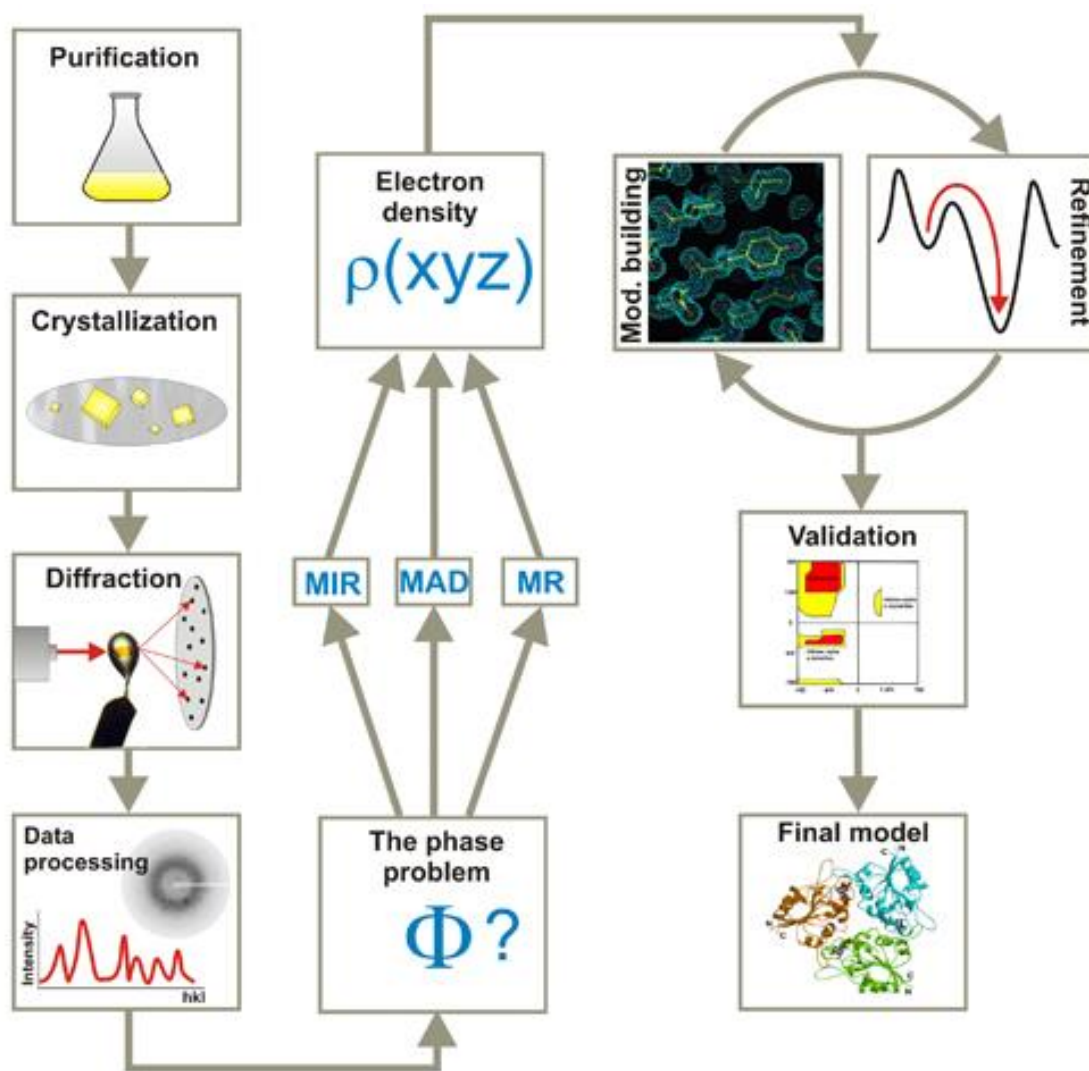
**Supplementary Table 1. X-ray data collection and refinement statistics.**

	Cathepsin S mutant + NNPI-C10	Cathepsin S WT + NNPI-C10
<b>Data collection</b>		
Space group	<i>P</i> 42 21 2	<i>P</i> 21 21 2
<i>a, b, c</i> (Å)	87.71, 87.71, 69.44	44.63, 129.22, 37.67
$\alpha, \beta, \gamma$ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	43.85–1.73 (1.82–1.73)	42.19–1.56 (1.64–1.56)
<i>R</i> <sub>mess</sub>	0.12 (2.71)	0.143 (1.46)
Mean <i>I</i> / $\sigma$ <i>I</i>	13.1 (0.74)	8.67 (1.31)
Completeness (%)	99.9 (99.3)	99.8 (99.7)
Multiplicity	8.5 (8.1)	5.7 (5.4)
CC1/2	0.99 (0.324)	0.997 (0.665)
<b>Refinement</b>		
Resolution (Å)	43.85–1.73	42.19–1.56
Total reflections	460653	344424
<i>R</i> <sub>work</sub> <sup>#</sup>	0.1761 (0.3881)	0.1708 (0.3365)
<i>R</i> <sub>free</sub> <sup>#</sup>	0.2059 (0.4062)	0.2251 (0.3663)
Number of non-hydrogen atoms	1928	1946
macromolecules	1781	1775
ligands	82	61
solvent	65	110
protein residues	226	226
RMS deviations (bonds) <sup>#</sup>	0.008	0.006
RMS deviations (angles) <sup>#</sup>	0.90	0.81
Ramachandran favored (%) <sup>#</sup>	97.73	97.73
Ramachandran allowed (%) <sup>#</sup>	2.27	2.13
Ramachandran outliers (%) <sup>#</sup>	0	0
Rotamer outliers (%) <sup>#</sup>	1.59	2.13
Clashscore <sup>#</sup>	4.10	4.00
Average B-factor <sup>#</sup>	38.85	27.96
macromolecules	38.08	26.79
ligands	55.38	59.07
solvent	39.11	29.56
PDB	<b>8PI3</b>	<b>8RND</b>

## Data availability

The atomic models related to this study have been deposited to the PDB under accession codes [8PI3](#) and [8RND](#). [Source data](#) are provided with this paper.

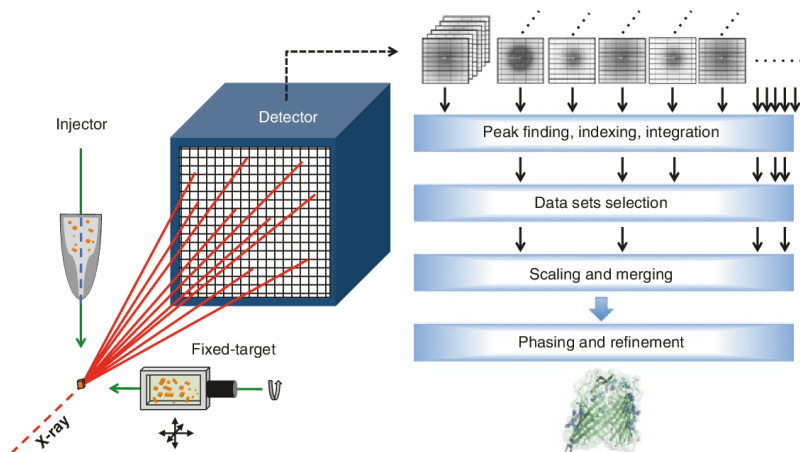
# Summary Pipeline of X-ray structure determination



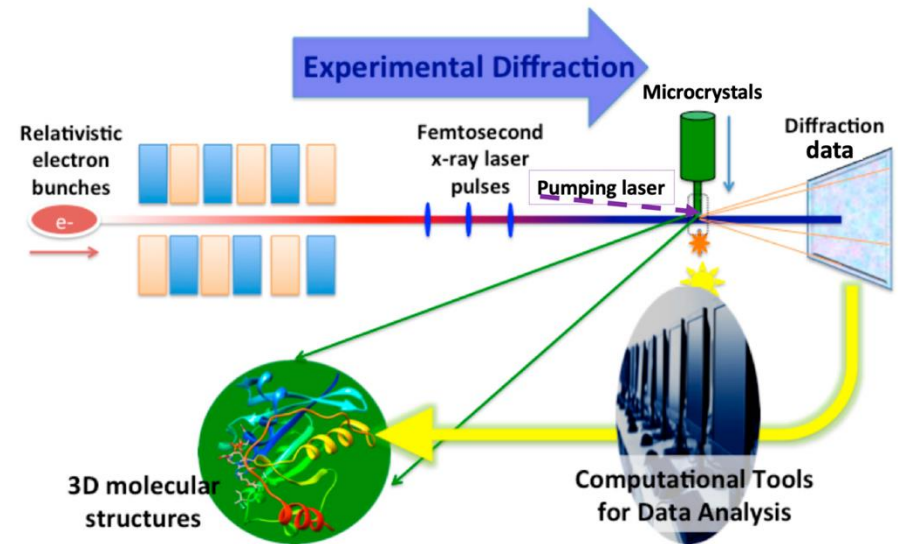
Once an initial model is placed in electron density, some additional steps (**manual building** of the detailed model, **refinement** and **validation**) are carried out to obtain the final model

# Other techniques on the X-ray side

## SSX: Serial Synchrotron X-ray crystallography



## XFEL: X-Ray Free-Electron Laser



# Time to do some Hands-on work

- **Data collection and processing**

Look at diffraction pattern of different crystals; and see how to go from 2D images to the data file (called reflection file) with XDS software

- **Solving the structure by Molecular Replacement**

Using Phaser in Phenix suite software

- **Manual model building and refinement**

Using coot and phenix-refine

- **Validation of model and deposition**

In coot and PDB database