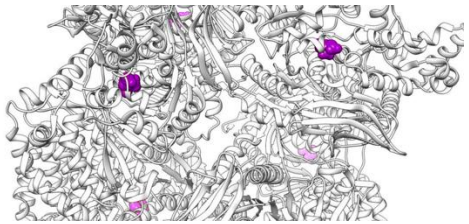
BIOENG-518:
Methods from disease models to therapy



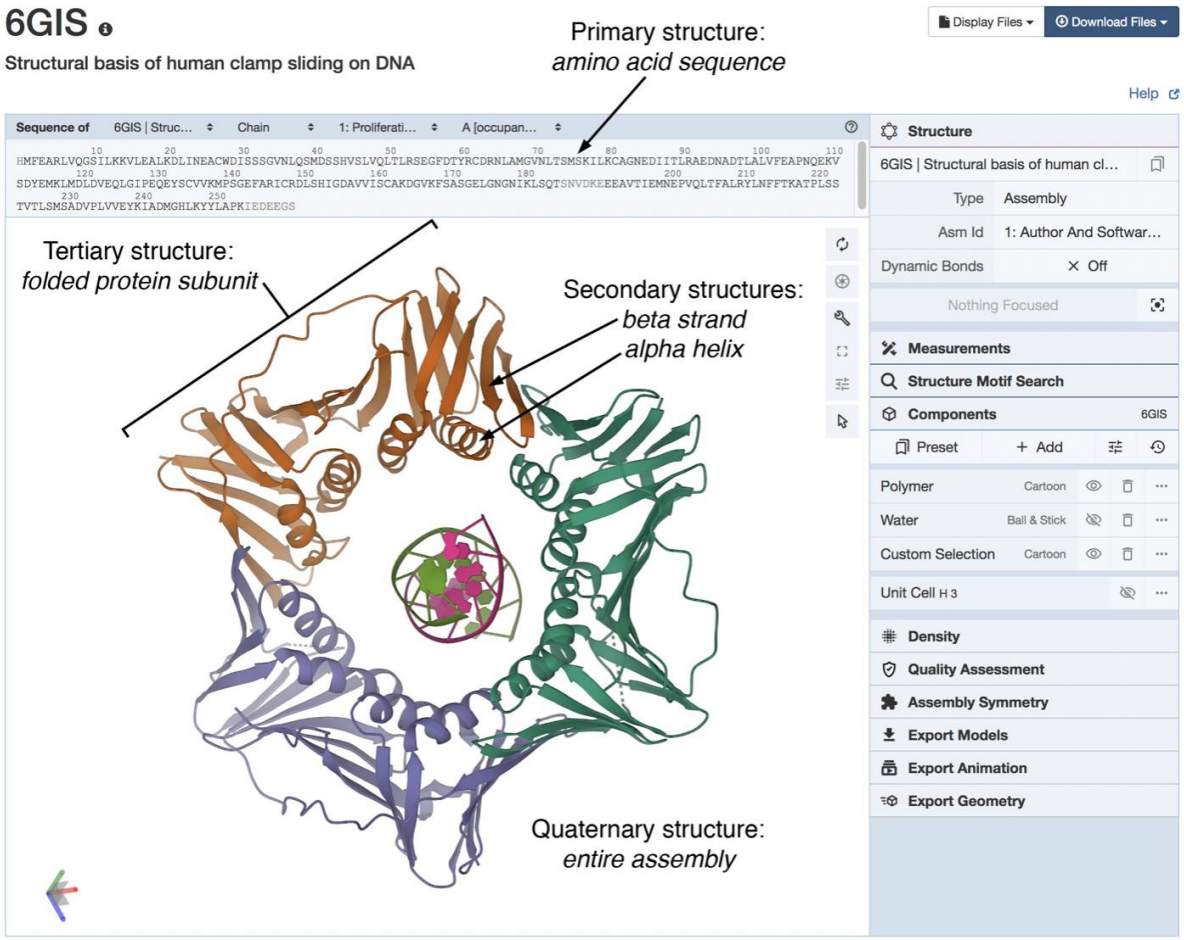
Biomolecular Integrative Structural Biology

Dr. Florence Pojer
Dr Kelvin Lau & Dr. Yoan Duhoo (practicals)

Spring Semester 2025

EPFL

Reminder of structure of proteins



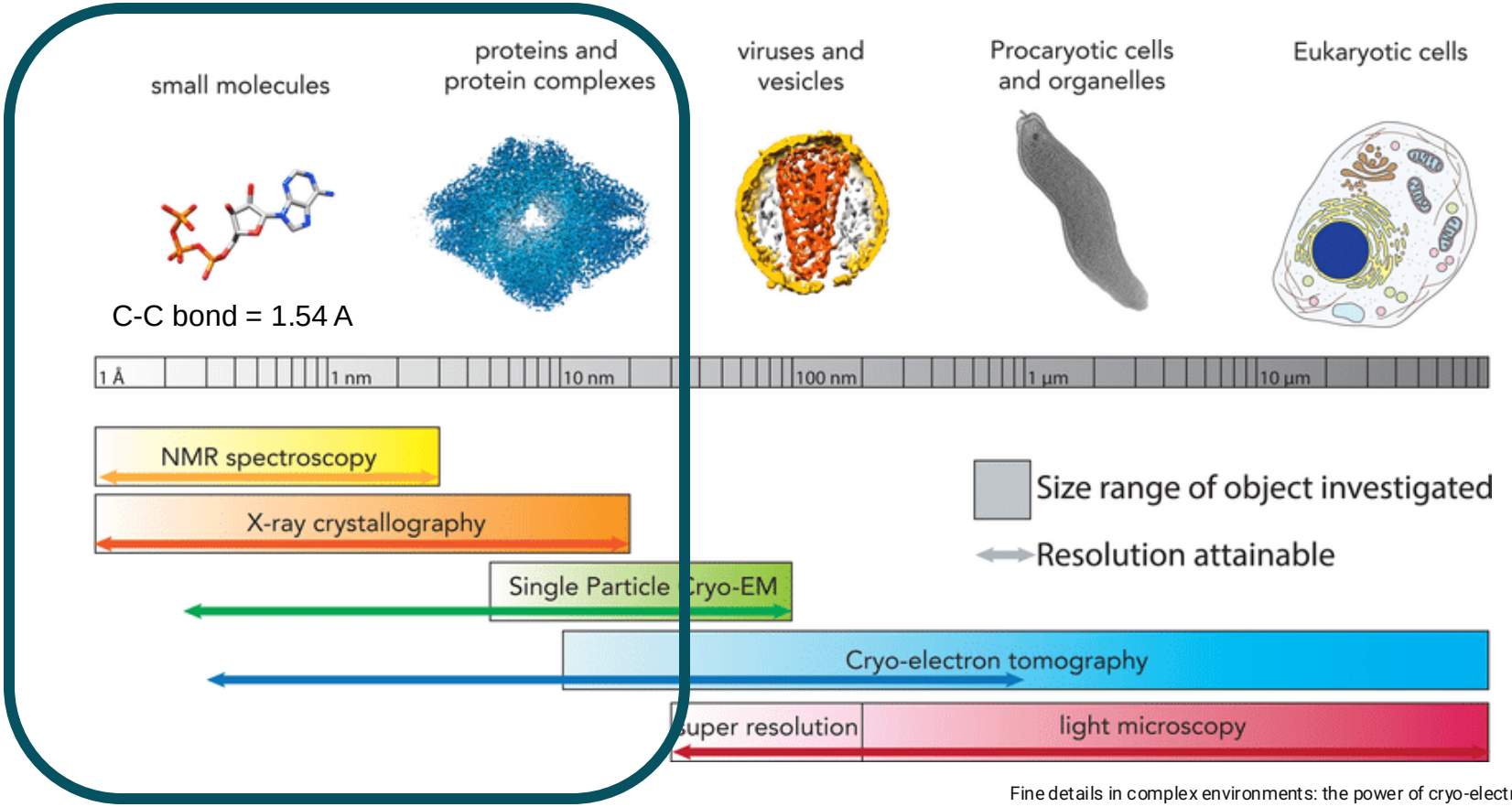
3D structure = function

Looking at protein structures in details reveal how they work

Objectives of this part of the course:

- Get an overview of the macromolecule structural techniques
- Good comprehension of their differences
- Visualization and interpretation of protein structures with open source software
- Hands-on practical at the facility on sample preparation for X-ray crystallography and cryoEM techniques, and data visualization.

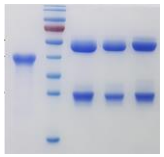
Tools for Structural characterization



Fine details in complex environments: the power of cryo-electron tomography

Experimental Methods

***In silico* models**
- Alphafold2
- Swissmodel
- Rosetta,..

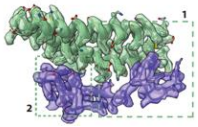


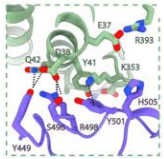
In 2024:

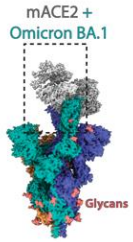
X-ray crystallography (9217 models)

NMR (284 models)

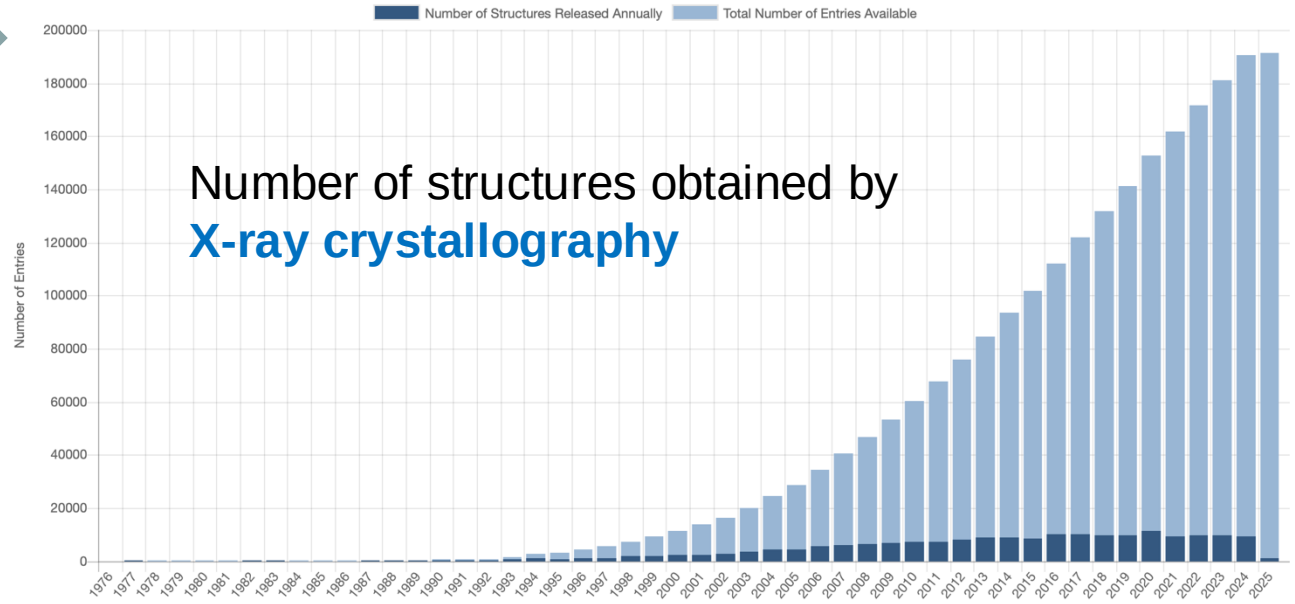
Single particle EM (5796 models)

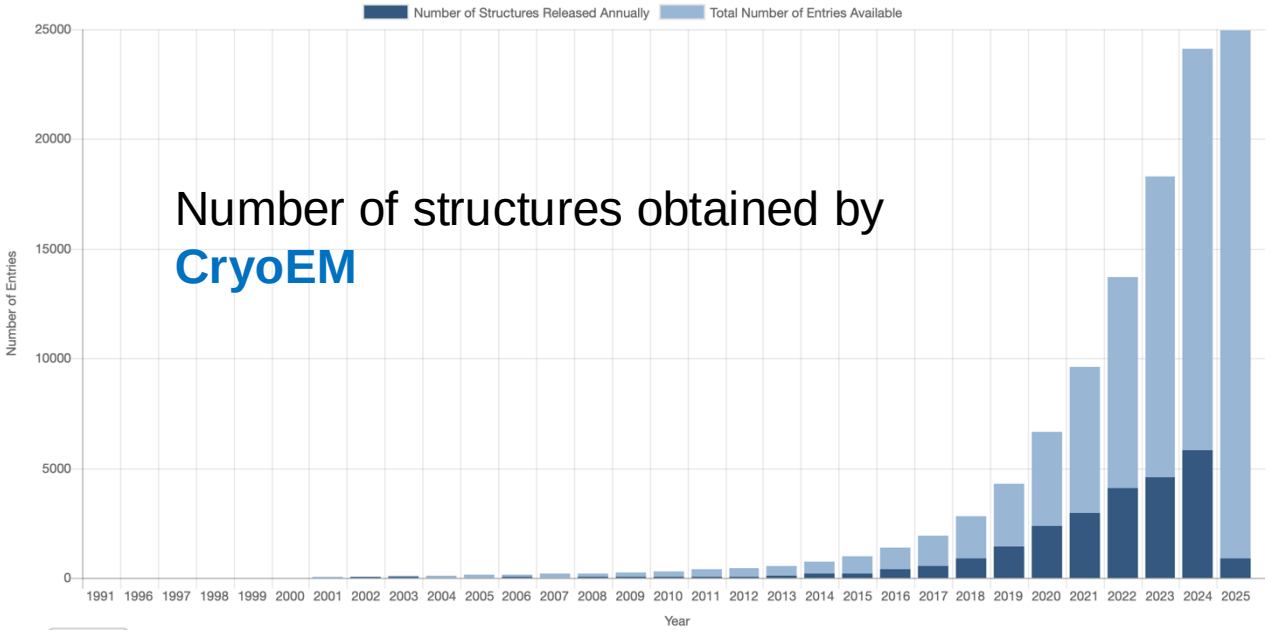




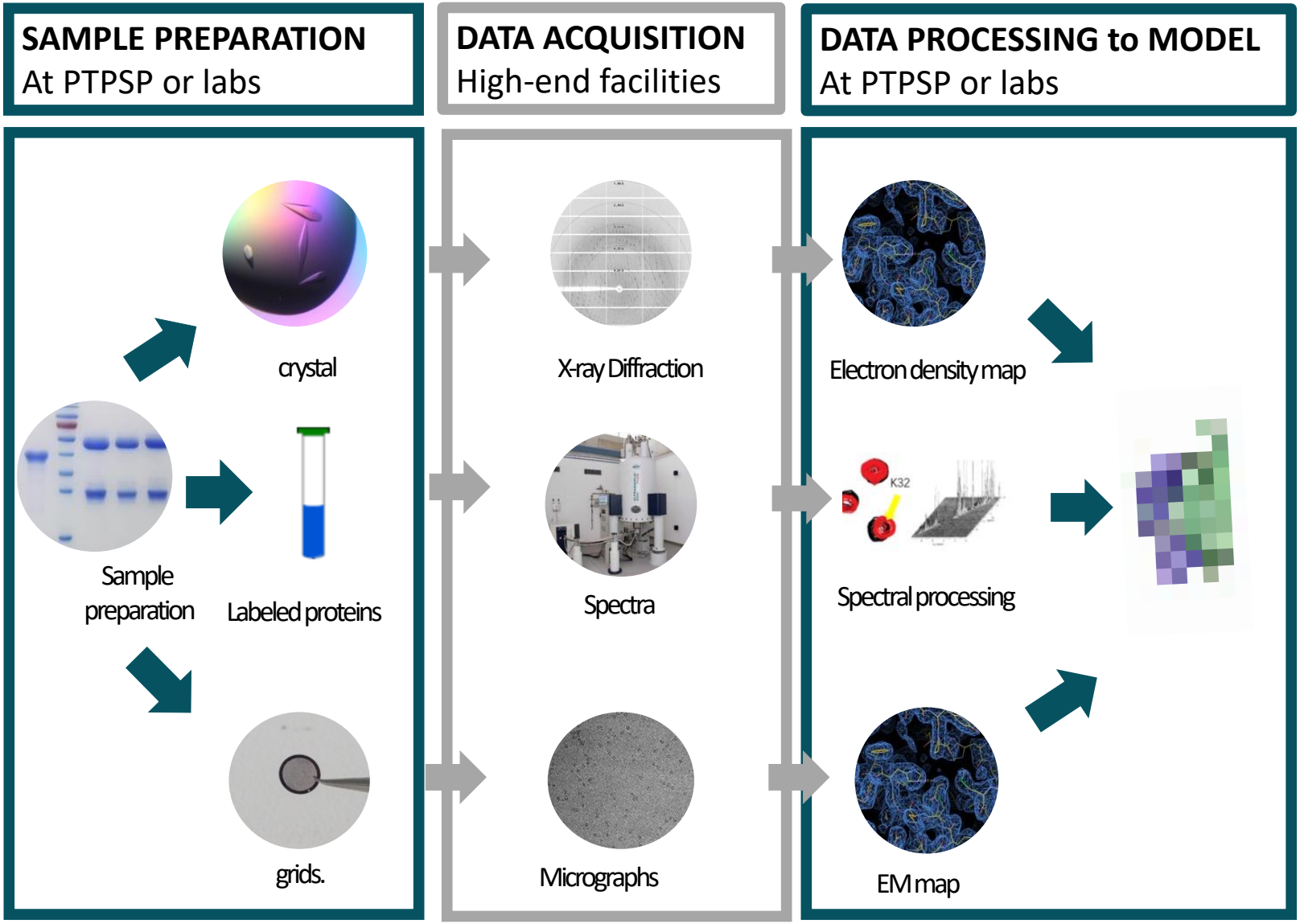


February 2025: 231 356 structures deposited in PDB database



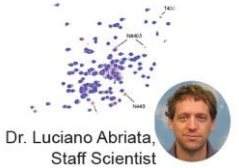


PDB database


SAMPLE PREPARATION
At PTPSP or labs
DATA ACQUISITION
High-end facilities
DATA PROCESSING to MODEL
At PTPSP or labs
crystal
Sample
preparation
Labeled proteins
grids.
X-ray Diffraction
Spectra
Micrographs
Electron density map
K32
Spectral processing
EM map

Close contacts to High-end facilities

Bio-NMR
@ NMR facility, EPFL

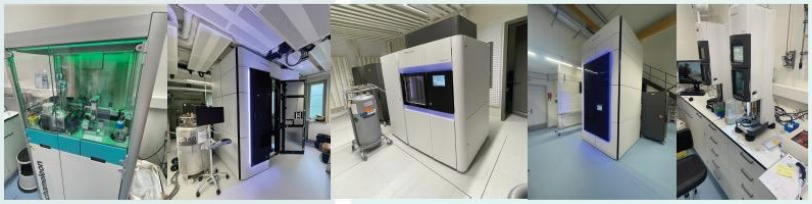


Dr. Luciano Abriata,
Staff Scientist

X-ray Crystallography
@ SLS PSI, Villigen

Dr. Kelvin Lau,
Staff Scientist

SPR, CryoEM
@ DCI, join EPFL,
UNIL & Uni Geneva



Dr. Yoan Duhoo,
Staff Scientist

Essential for success:

To know very well your protein of interested its context

Example of very useful website

1. UniProt

Very informative and up to date: Function, Names & Taxonomy, Subcellular location, Pathology & Biotech, Interaction, Structure, Family & Domains, Sequence, Cross-references.

Spike SARS-Cov-2: https://www.uniprot.org/uniprot/P0DTC2

Alphafold2 included now in UniProt

2. Protparam

Essential for biophysical parameters of your POI: Number of amino acids, Molecular weight, Theoretical PI, Amino acid composition, Atomic composition, Extinction coefficients.

https://web.expasy.org/protparam/

3. SWISSMODEL, I-TASSER and Alphafold2

Useful for quick 3D modelling of any proteins based on their AA sequence

https://swissmodel.expasy.org

https://zhanglab.ccmb.med.umich.edu/I-TASSER/

4 HHPRED:

Useful for unknown function/no structure POI. It is a sequence database searching and structure prediction

https://toolkit.tuebingen.mpg.de/tools/hhpred

Example of design of Spike SARS-Cov-2 construct

Uniprot essential info:

Molecule processing

Feature key	Position(s)	Description	Actions	Graphical view	Length
Signal peptide	1 – 12	UniRule annotation	Add BLAST		12
Chain (PRO_0000449646)	13 – 1273	Spike glycoprotein	Add BLAST		1261
Chain (PRO_0000449647)	13 – 685	Spike protein S1 UniRule annotation	Add BLAST		673
Chain (PRO_0000449648)	686 – 1273	Spike protein S2 UniRule annotation	Add BLAST		588
Chain (PRO_0000449649)	816 – 1273	Spike protein S2' UniRule annotation	Add BLAST		458

Topology

Feature key	Position(s)	Description	Actions	Graphical view	Length
Topological domain	13 – 1213	Extracellular UniRule annotation	Add BLAST		1201
Transmembrane	1214 – 1234	Helical UniRule annotation	Add BLAST		21
Topological domain	1235 – 1273	Cytoplasmic UniRule annotation	Add BLAST		39

Sites

Feature key	Position(s)	Description	Actions	Graphical view	Length
Site	685 – 686	Cleavage; by TMPRSS2 or furin UniRule annotation 2 Publications			2
Site	815 – 816	Cleavage; by host UniRule annotation			2

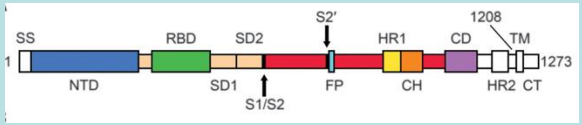
PTM databases

GlyConnect	2838, 256 N-Linked glycans (24 sites), 5 O-Linked glycans (3 sites) 2839, 38 N-Linked glycans (20 sites) 2840, 109 N-Linked glycans (22 sites)
GlyGen	P0DTC2, 24 sites, 172 N-linked glycans (22 sites), 4 O-linked glycans (2 sites)

Interaction

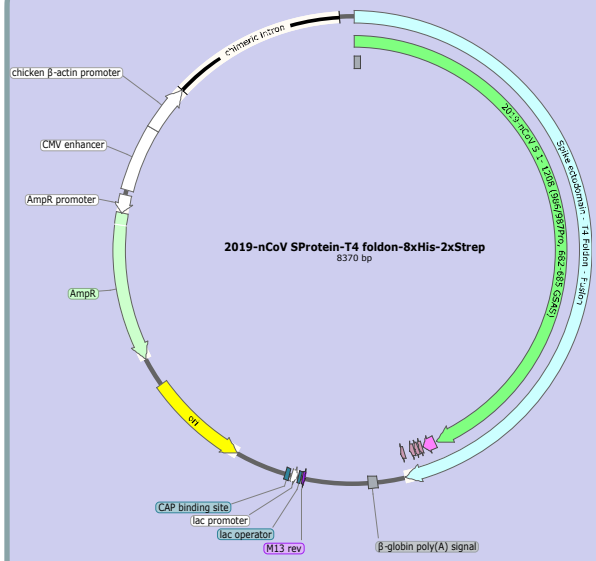
Construct was designed for cryo-EM, but very suitable also for serological assays

Good thinking



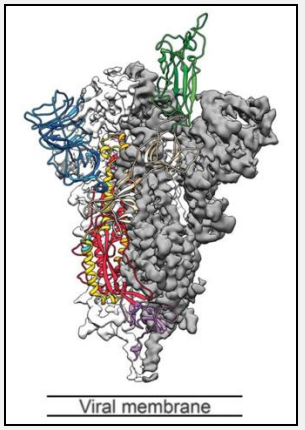
- Secreted natural Signal sequence kept
- Transmembrane domain removed
- T4 trimerization domain added
- Furin cleavage-site mutated
- Two stabilizing proline mutations added
- Purification tags added at C-term (twin-strep and His tags)

Vector Designed



Production in CHO cells
Purification by affinity tags

Great Results



Structure of Spike by CryoEM

Serological tests (CHUV)

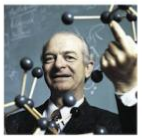
Wrapp et al., 2020

X-ray crystallography

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

See a complete list of winners at 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

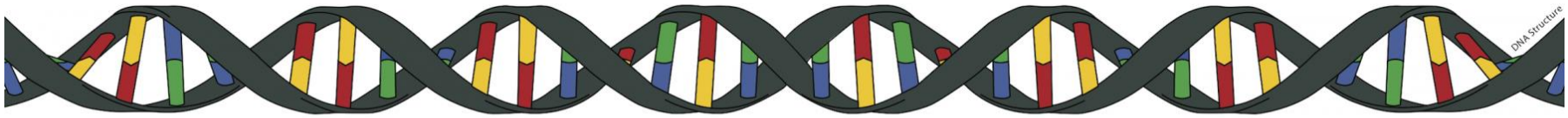
Herbert Hauptman & Jerome Karle
Direct mathematical methods of determining crystallized materials
1985

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



Max von Laue
First demonstrated X-ray diffraction through crystals
1914

Sir William H. & Sir William L. Bragg
First atomic crystal structure
1915

John Kendrew & Max Perutz
Hemoglobin: Transport protein, which led to the understanding of Sickle Cell Anemia
1962

Dorothy Hodgkin
Structures of cholesterol, penicillin, vitamin B12, and insulin
1964

William Lipscomb
The structure of boranes, illuminating problems of chemical bonding
1976

Johann Deisenhofer, Robert Huber & Hartmut Michel
First membrane protein that is essential to photosynthesis
1988

Peter Agre & Roderick MacKinnon
Discoveries concerning channels in cell membranes
2003

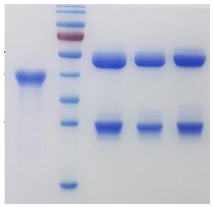
Roger Kornberg
Studies of the molecular basis of eukaryotic transcription
2006

Martin Karplus, Michael Levitt & Arieh Warshel
Development of sophisticated computer simulations for complex chemical processes
2013

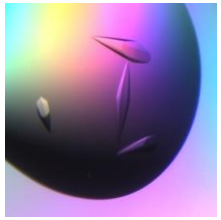
Additional Important Contributors to Crystallography

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

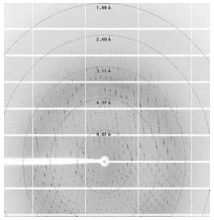


Sample preparation



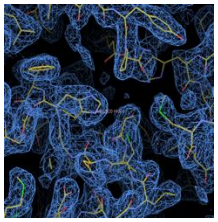
Crystallization

X-rays



Data collection

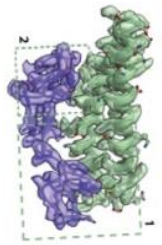
phases



Solving the structure-
Electron density map

fitting

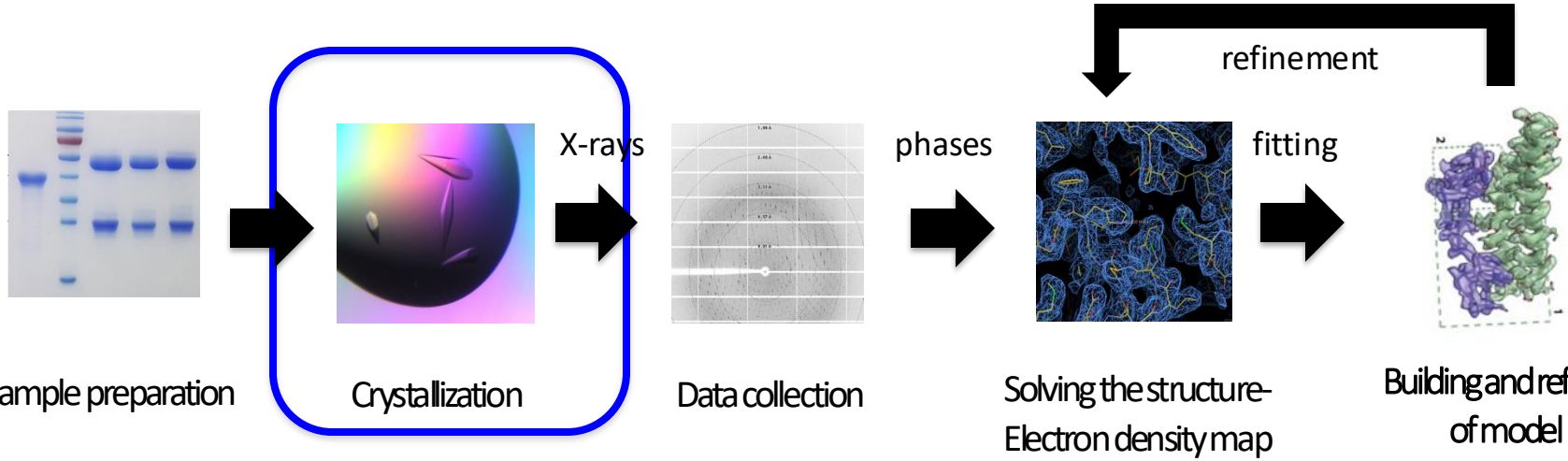
refinement



Building and refining
of model

Prerequisites:

- 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 300µl at 10mg/ml to screen around 600 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,…)

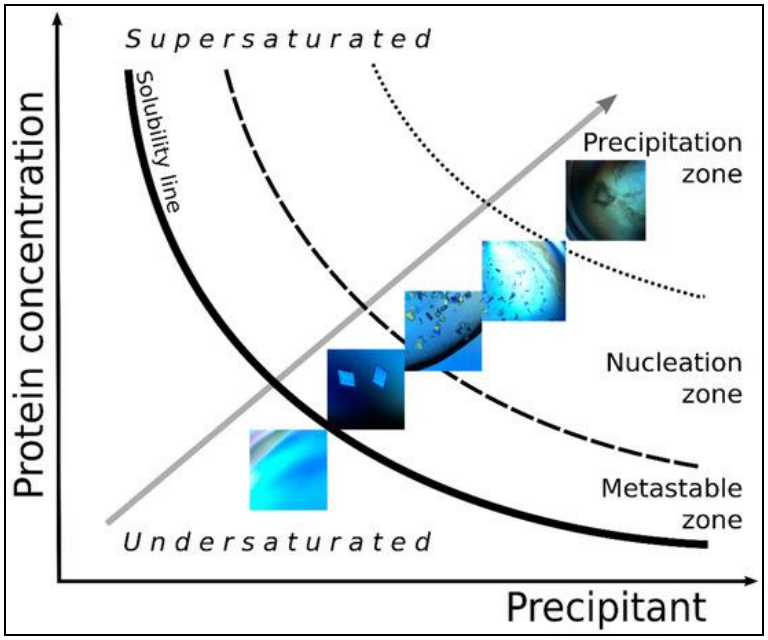
Checking of crystals formation under light microscope

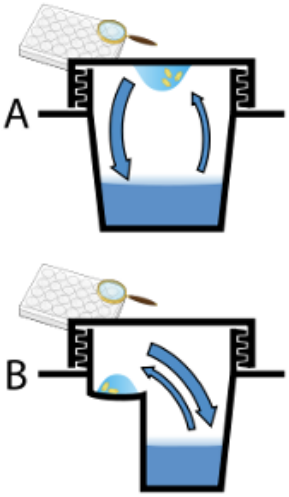
Crystals form in days to months; screening success rate is 0-10% depending on proteins

- Optimization of crystals to single crystals to obtain suitable diffraction pattern and improve resolution
- Cryoprotection for X-ray diffraction

Fish single crystal with loop and and add cryoprotectant (e.g. 25% glycerol)

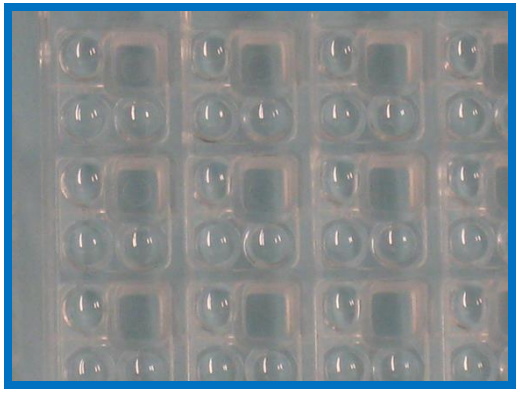
Vapor diffusion techniques





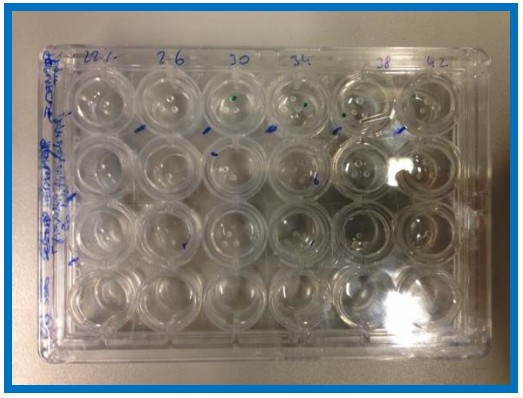
Hanging drops

Sitting drops

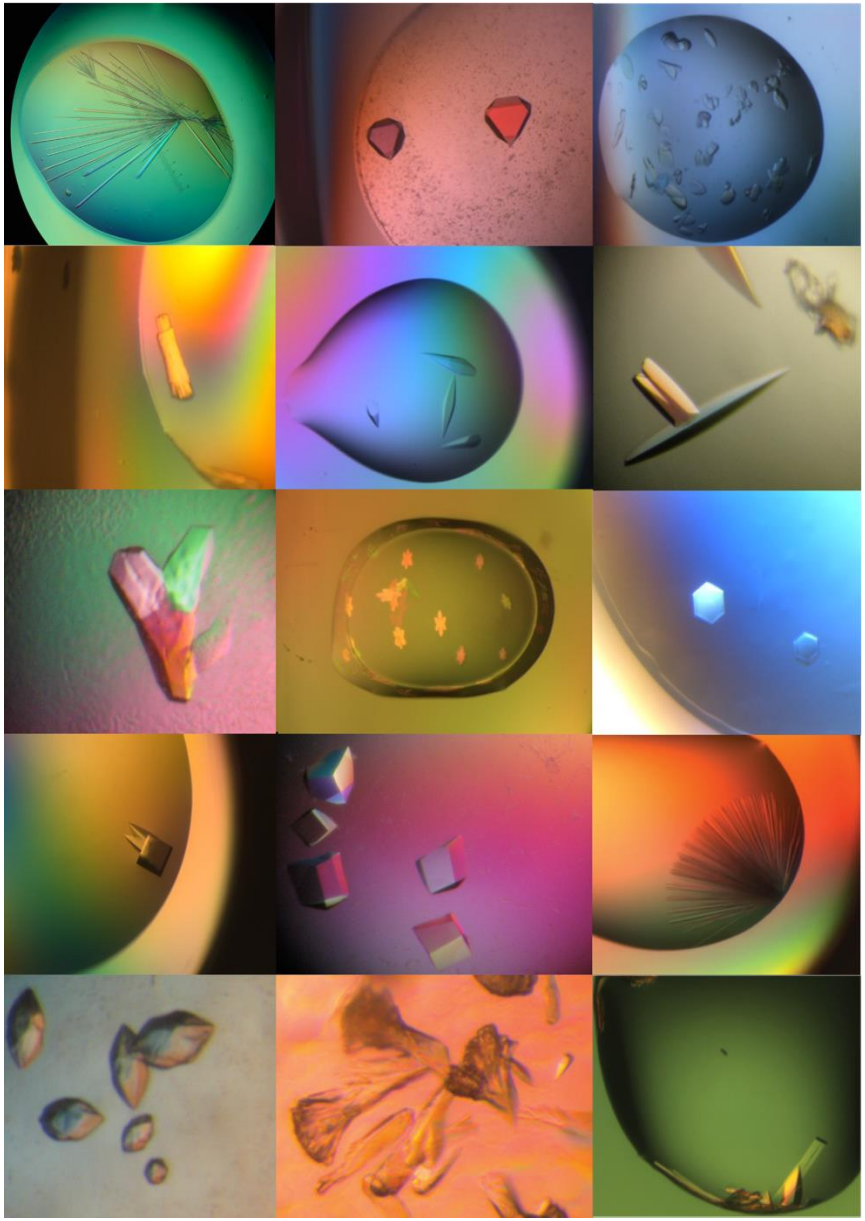


Sitting drops vapor diffusion
Mix of precipitants and protein
(total of 0.2ul to 1ul)

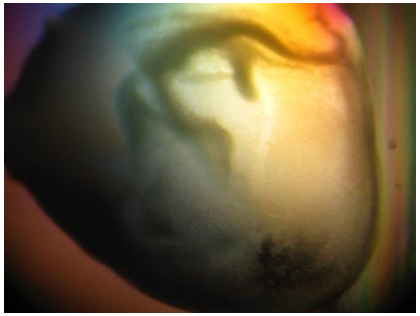
Mosquito (STP labtech)



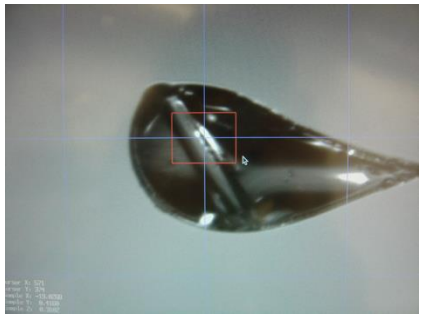
Hanging drops vapor diffusion
mix precipitants and protein
(total of 1ul to 4ul)



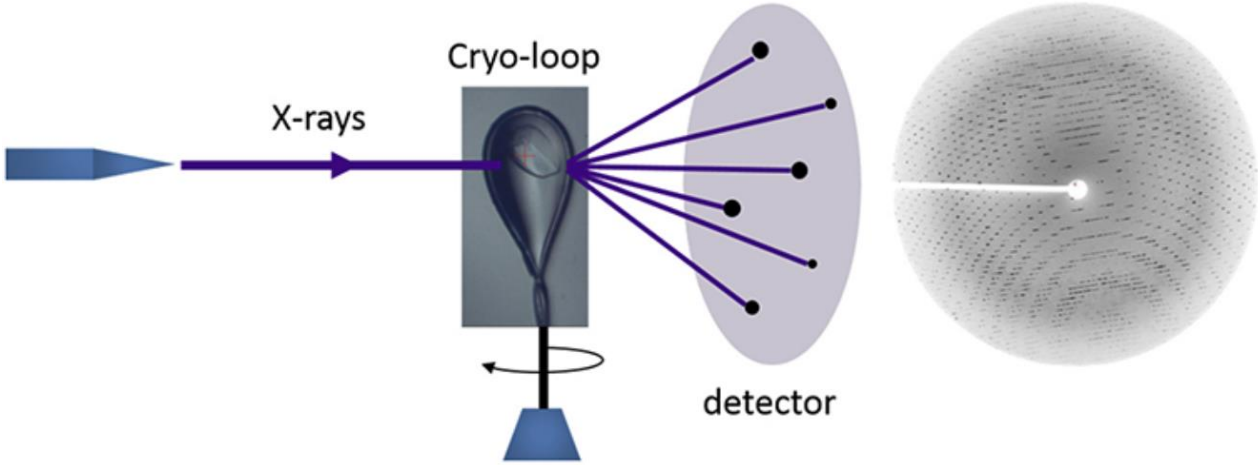
Crystals can have many shapes and sizes



Heavy protein precipitation

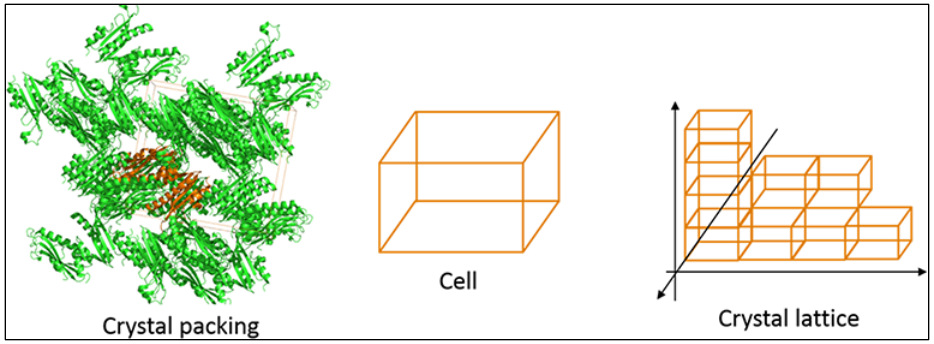


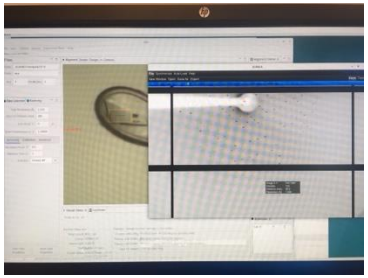
Crystal cryoprotected in a loop



Why using X-rays? Their wavelength is of the order of the angström and thus corresponds to the distance between two bound atoms.

Why using a crystal? 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.



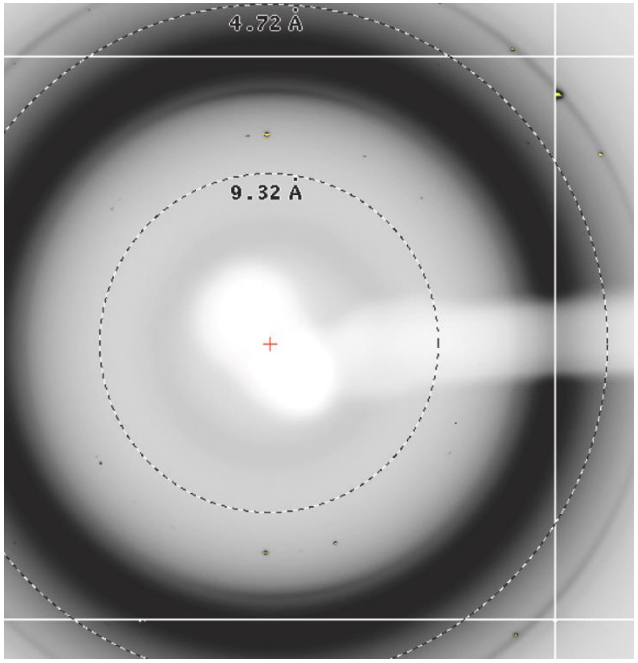


Data collected remotly from EPFL

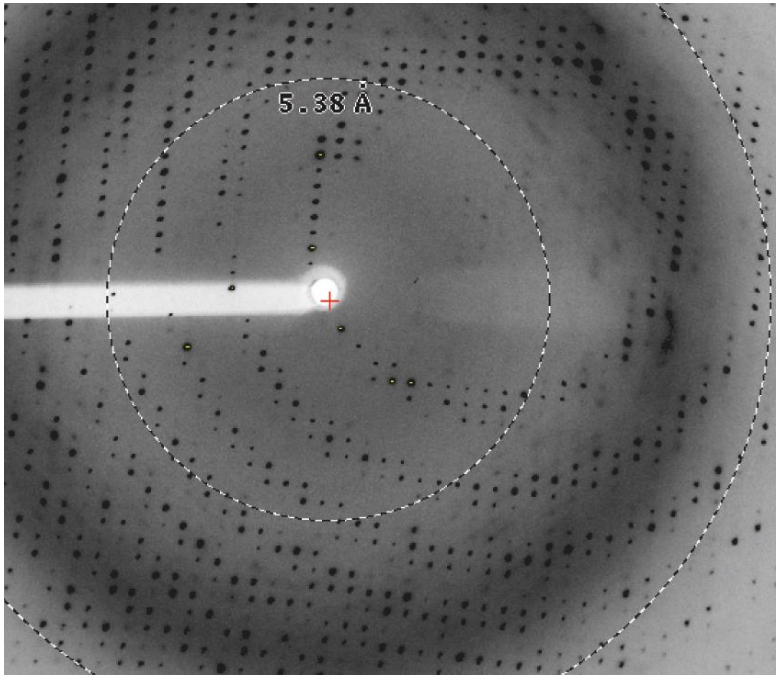
Data collected at Swiss Light Source (SLS-PSI)

Data collection: 2D diffraction images

- For a full data collection, images are collected by rotating the crystal (slices of 0.1° , total of 360° for crystal with low symmetry)
- The total number of images to collect depends on the symmetry of molecules in the crystal
- A full data collection takes only 3 minutes



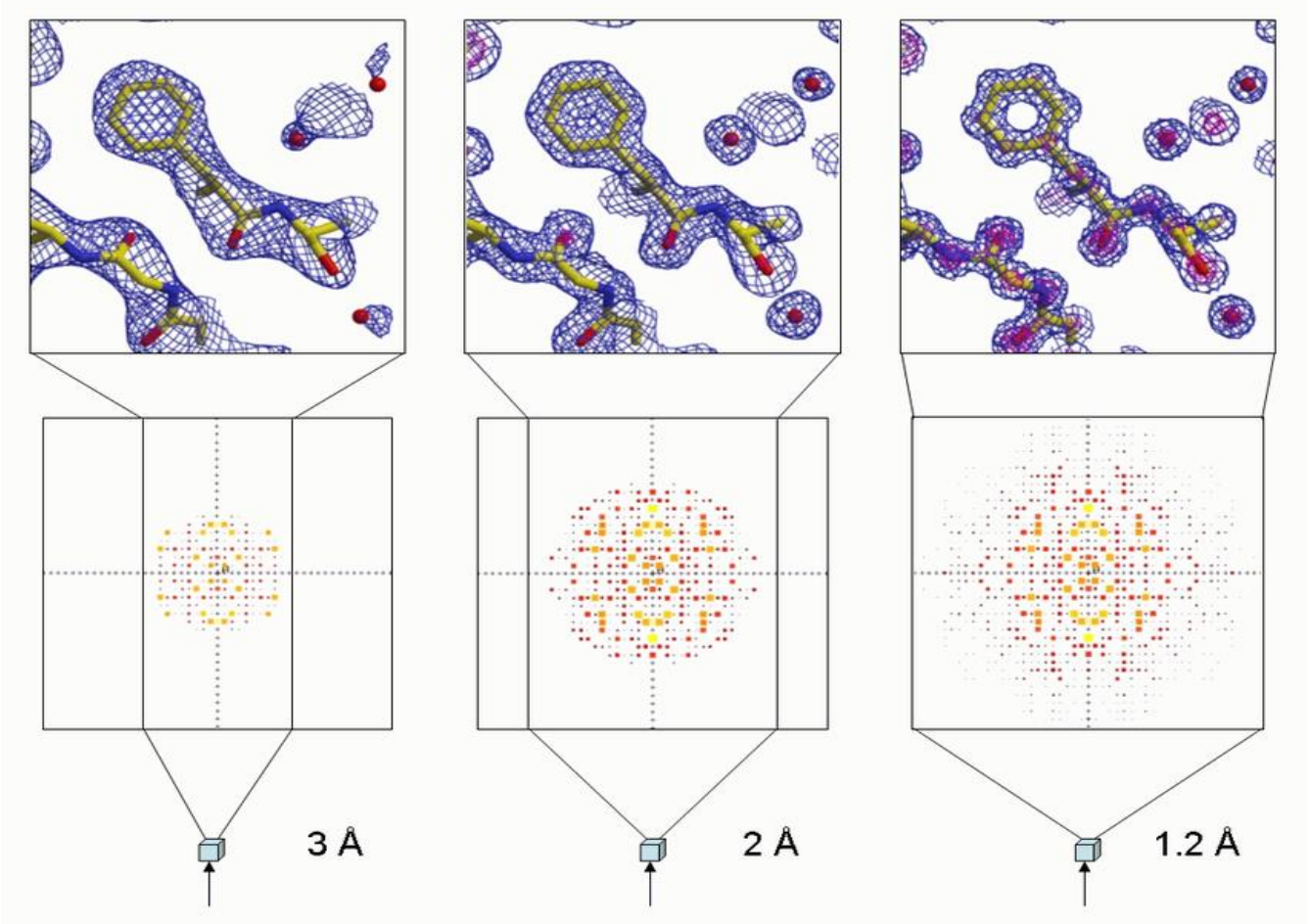
X-ray diffraction of crystal of salt!



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

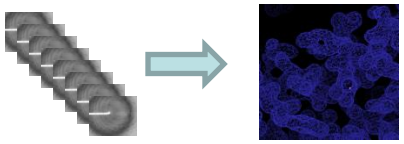
Electron density maps in function of resolution

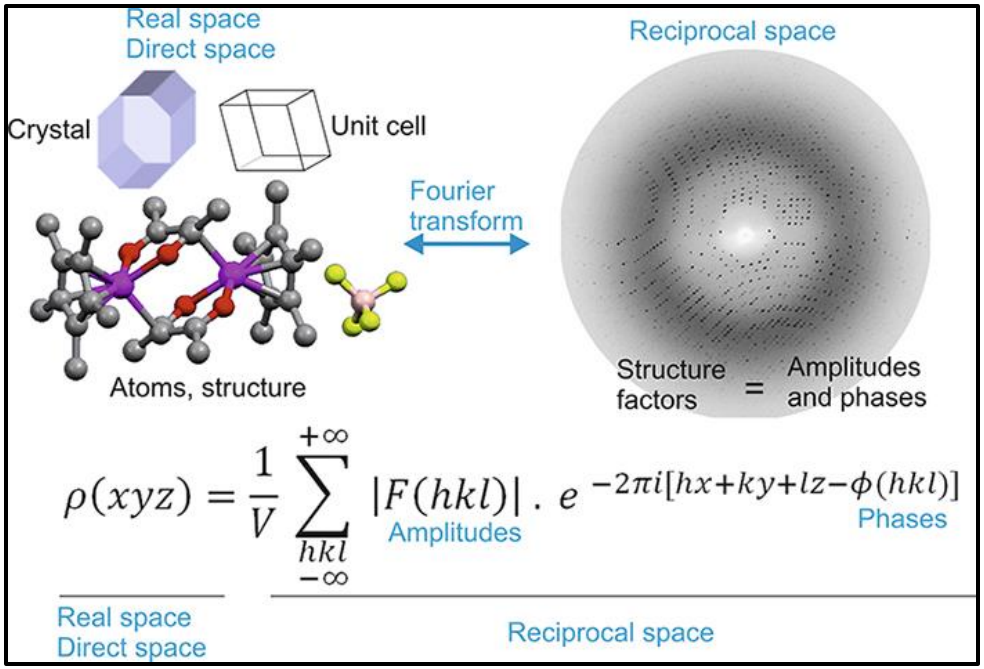
movie



http://www.ruppweb.org/Xray/resolution.html

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

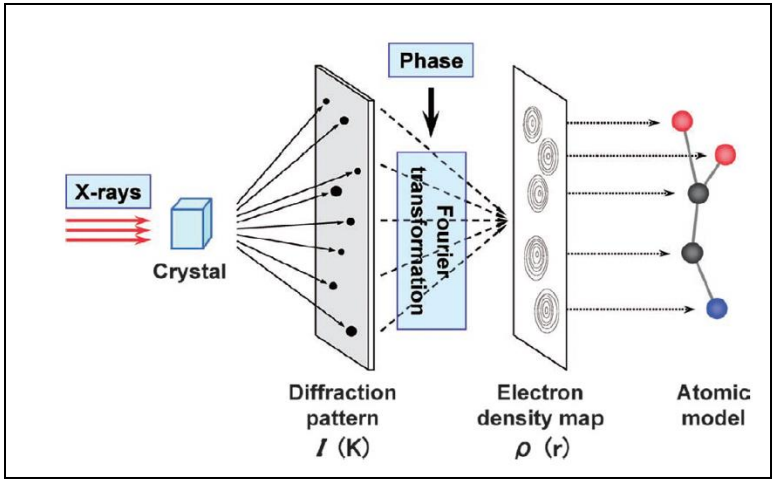




Software, such as XDS and mosflm, converts 2D diffraction images in an unique reflection file, that is necessary to generate a first electron density map. Data processing steps include spots finding and autoindexing; Integration (measurement of spot intensities); scaling and merging into one reflection file.

The phases are still missing and can be estimated by different methods.

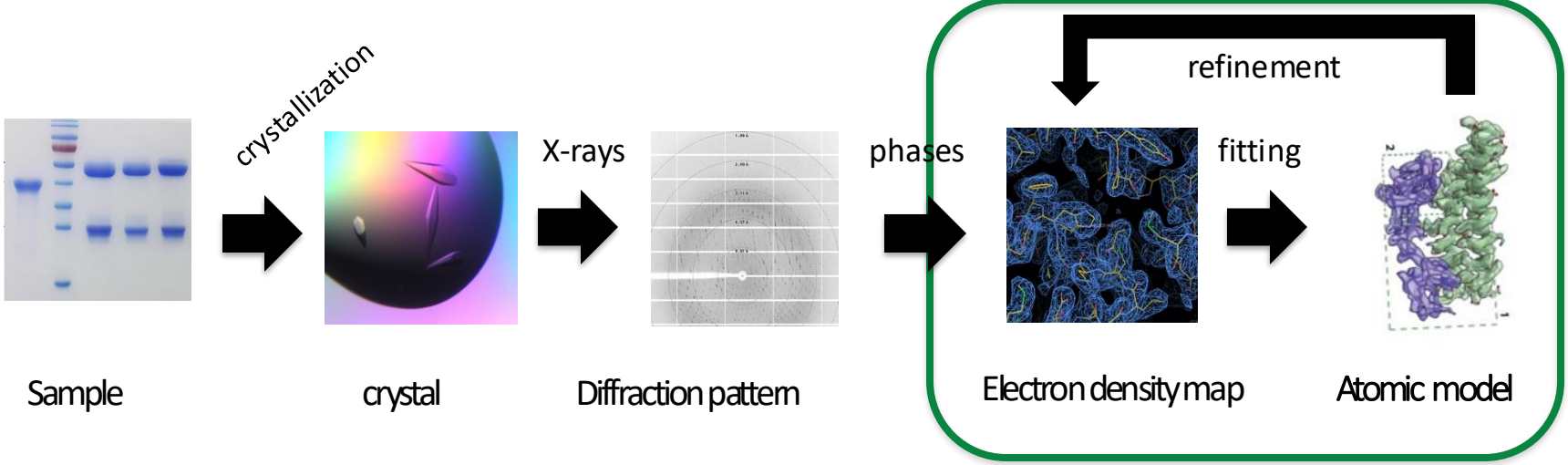
Methods to provide estimates of phases



Methods to solve X-ray structures:

- **Molecular replacement (MR):** it relies upon the existence of a previously solved structure which is similar to the unknown one. MR tries to find the model which fits best experimental intensities by rotation and translation. Modeling software, such as Alphafold2, are usefull, and now even integrated in X-ray 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

https://www.rigaku.com/newsletters/mabu/june2018/rigaku.journal_01.pdf



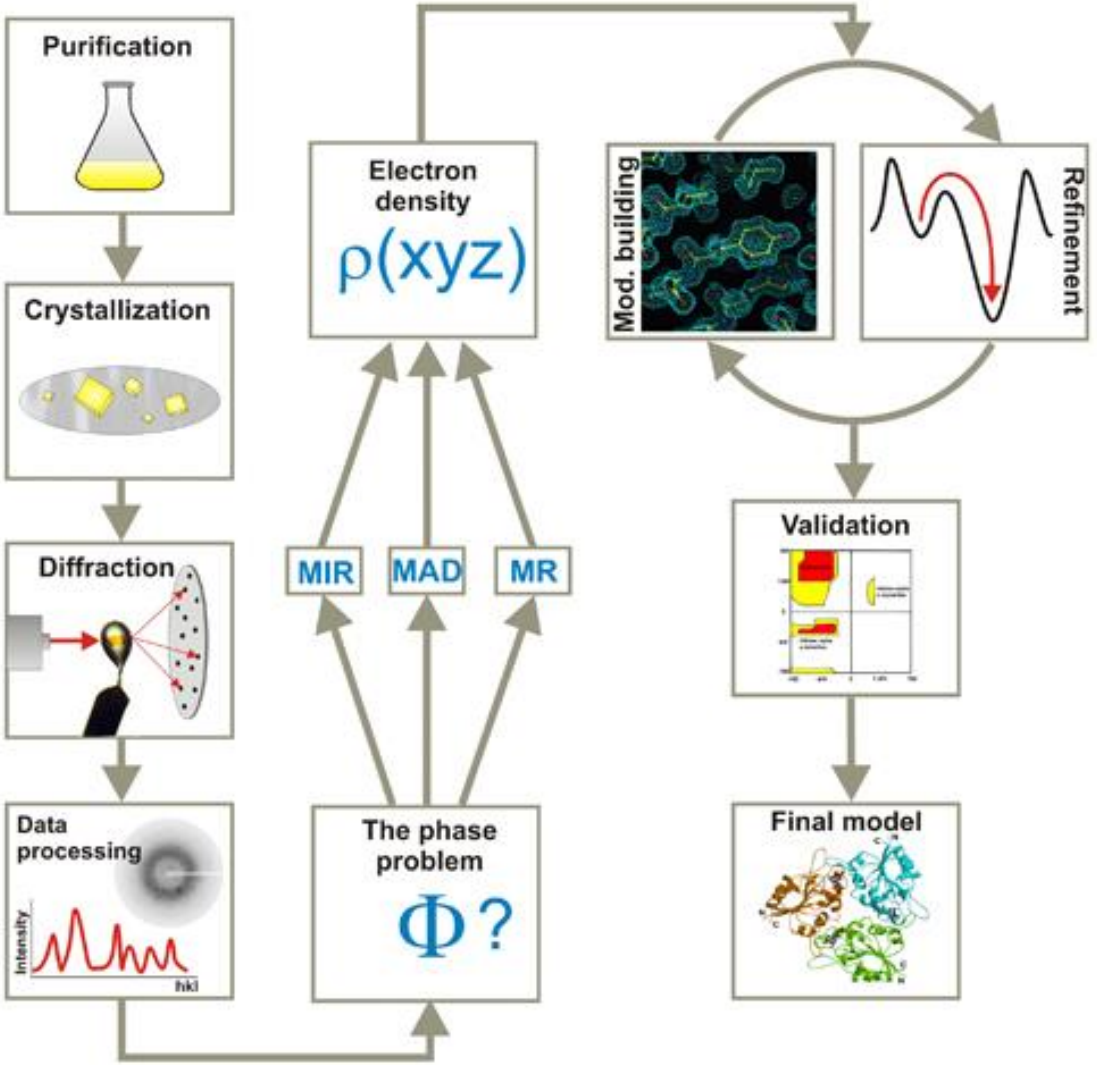
Building and refining of the initial model to fit the data:

- **Iterative process**: **Manual building of model** into electron density maps and **refinement**: Phenix, CCP4i, Coot
- **Validation** of the 3D model
- **Interpretation** of 3D model and **deposition into PDB database**

Higher is the resolution, more accurately can the atoms be placed into the election density map. The full process can take hours to days or months.

CCP4 Phenix PyMOL

Summary of X-ray structure determination pipeline



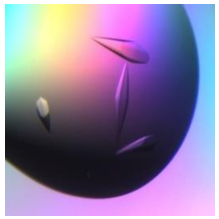
Once an initial model is placed in electron density, some additional steps (**manual building** of the detailed model, **refinement** and **validation**) are carried out to obtain the final model

https://www.xtal.iqfr.csic.es/Cristalografia/parte_07-en.html

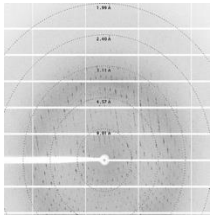
SAMPLE PREPARATION

DATA ACQUISITION

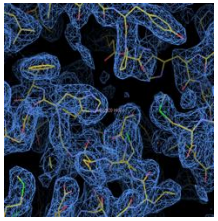
MODEL BUILDING AND REFINING



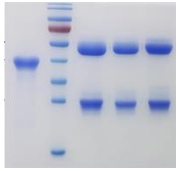
crystal



X-ray Diffraction



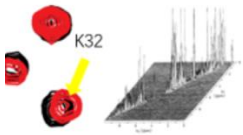
Electron density map



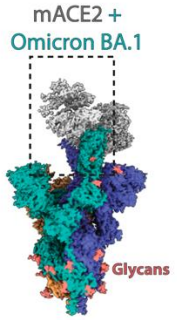
Sample preparation

Labeled proteins

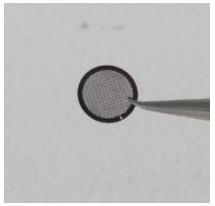
Data acquisition



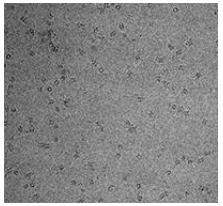
Spectral processing



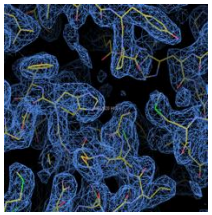
High resolution 3D model



Sample prep.

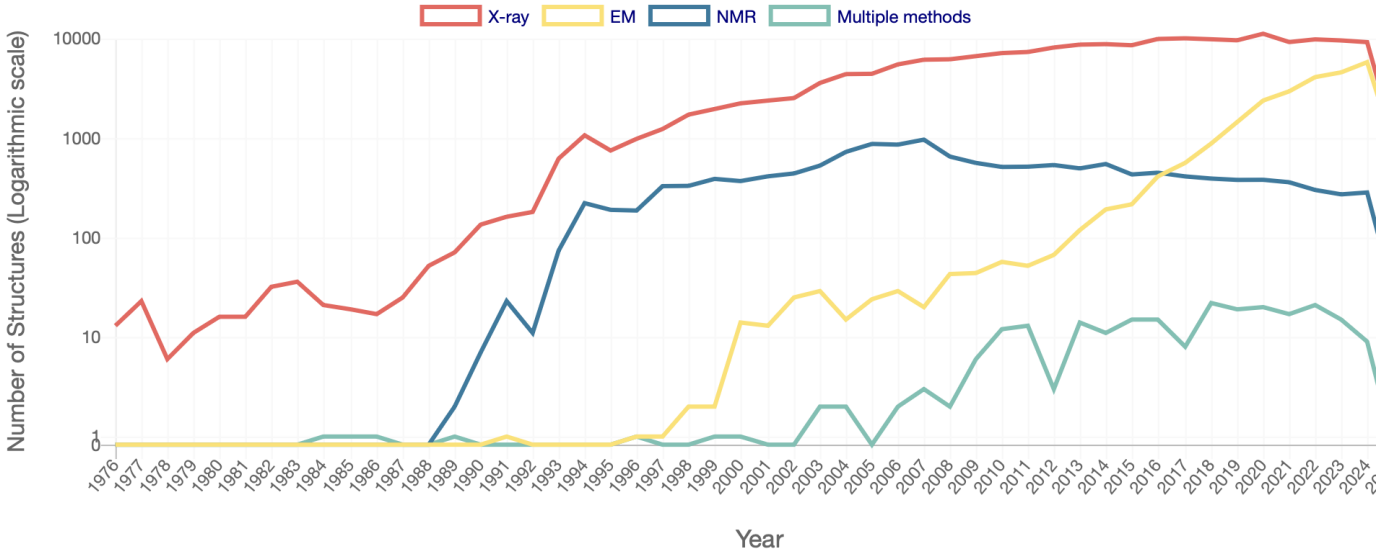


Micrographs



EM map

Single Particle (SPR) Electron Microscopy (EM)



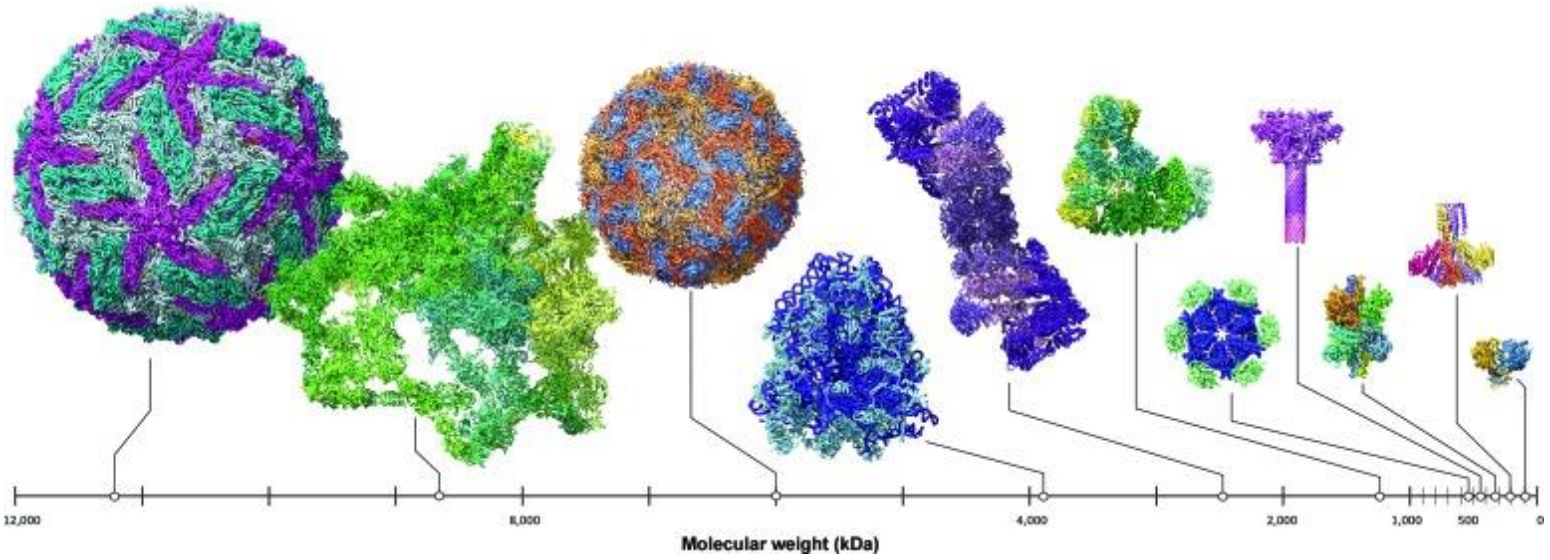
Year	X-Ray	NMR	EM
2025	1,082	29	876
2024	9,217	284	5,796
2023	9,588	272	4,578
2022	9,827	301	4,104
2021	9,238	360	2,951
2020	11,196	381	2,387
2019	9,619	380	1,451
2018	9,853	392	882
2017	10,069	412	564
2016	9,923	449	412

Resolution revolution in cryoEM

I. The Advanced in streamlining many of the steps in the cryo-EM workflow : Direct electron detectors, new movie-processing methods, new classification methods separate images of different structures.

II. Applied to a range of proteins and protein complexes of broad general interest

III. The creation of national facilities that provide access to the latest cryo-EM technology



Sriram Subramaniam (2019) The cryo-EM revolution: fueling the next phase

DUBOCHET CENTER FOR IMAGING
LAUSANNE

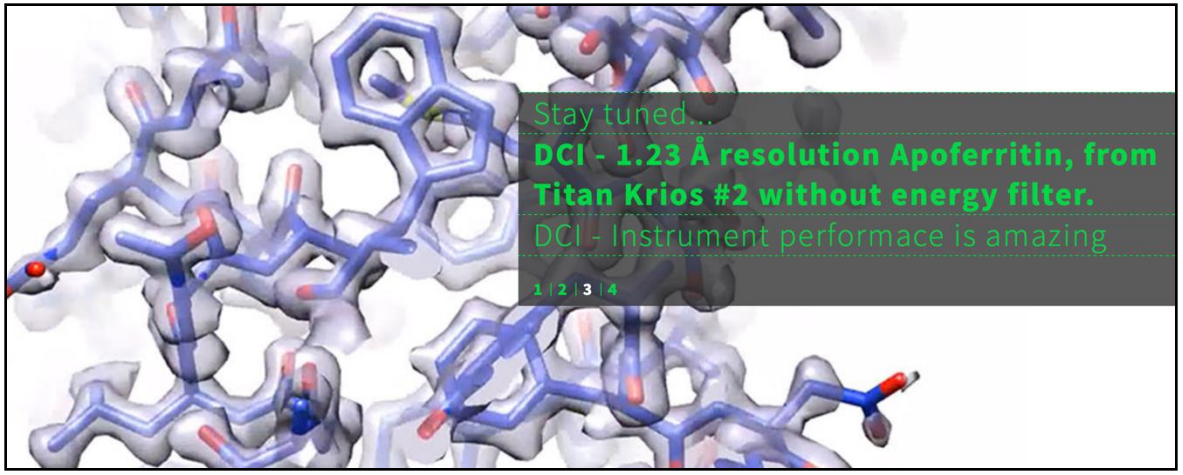
Dubochet Center for Imaging
UNIL-EPFL-UNIGE
High-end EM microscopes
and expertise

Instruments at the DCI Lausanne

Titan Krios (300kV), E-CFEG, SelectrisX, Falcon4

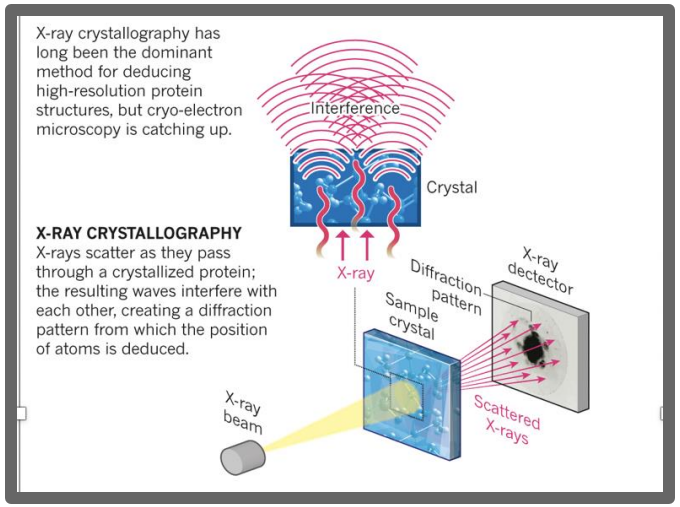
Titan Krios (300kV), E-CFEG, Falcon4

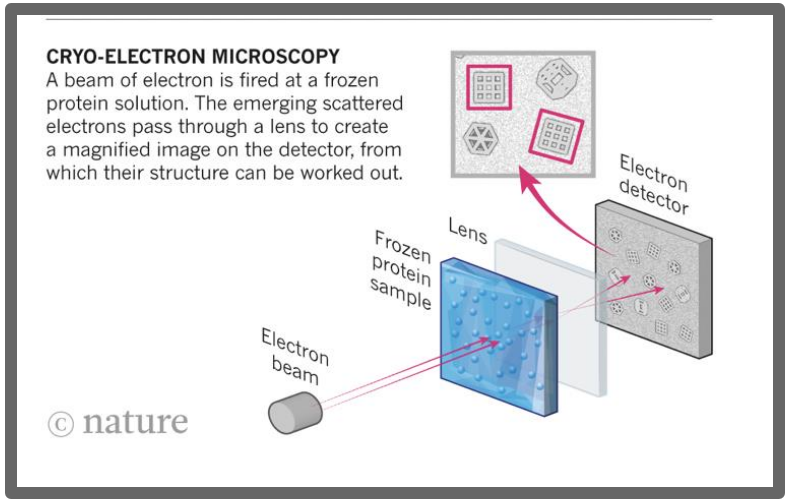
Glacios (200kV), X-FEG, Falcon4

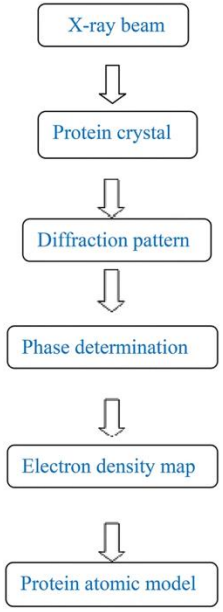


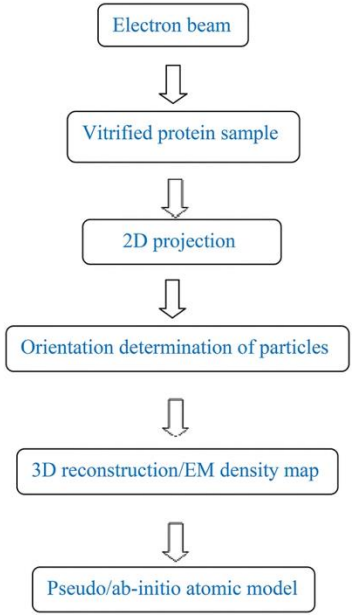
Jacques Dubochet in front of the first Titan Krios of the Dubochet Center for Imaging, June 4, 2021

X-ray beam versus Electron beam





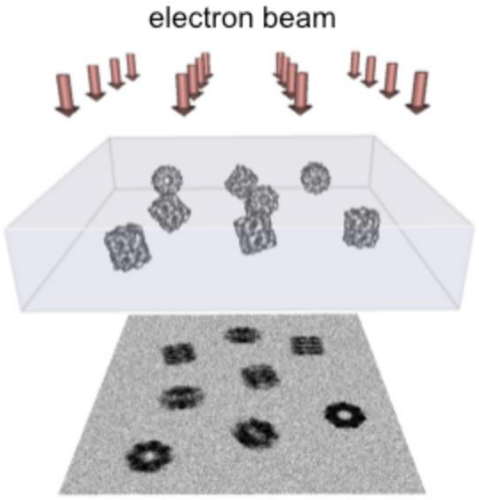


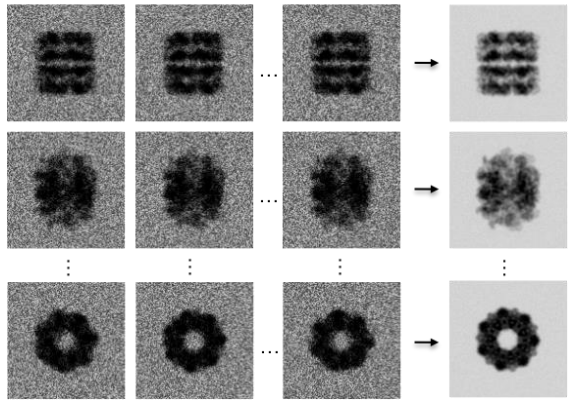


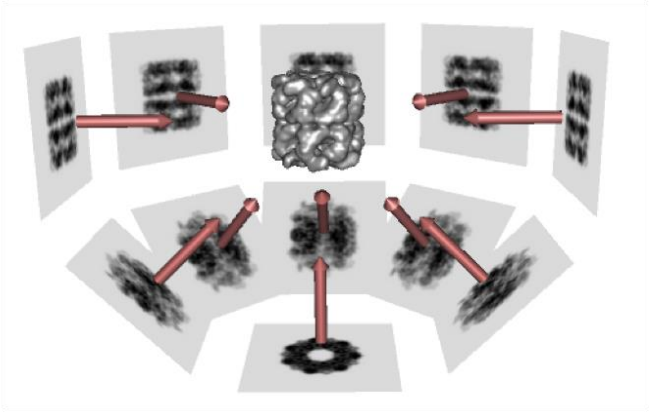
Ewen Callaway (2015) The revolution will not be crystallized: a new method sweeps through structural biology

General principle of Single Particle Reconstruction

2D projection of a 3D object are recorded with an electron microscopes and aligned to generate a 3D reconstruction



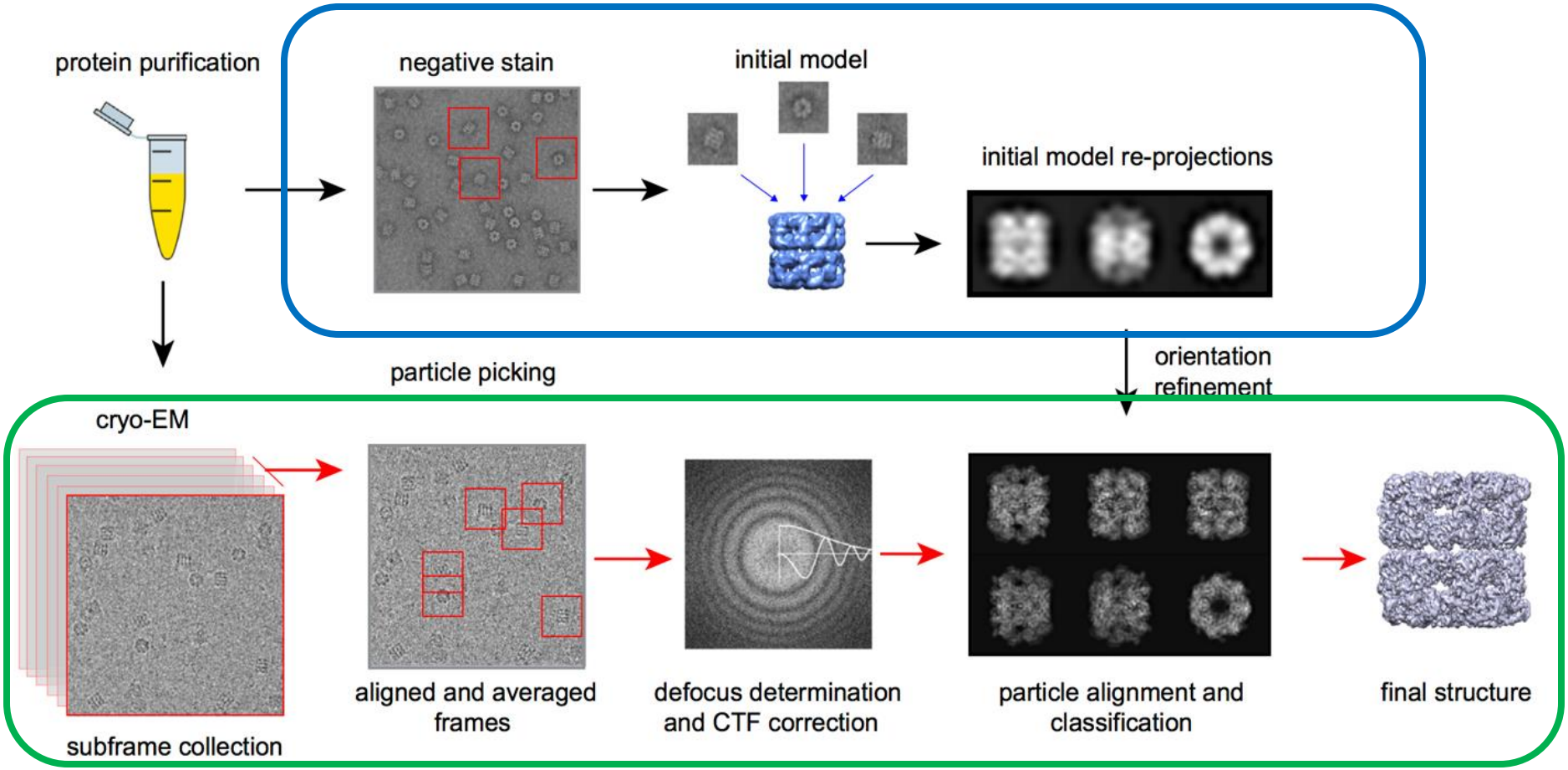




http://people.csail.mit.edu/gdp/cryoem

Pipeline for Single Particle Reconstruction

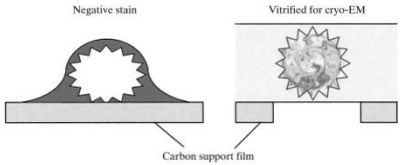
Negative stain: fast, cheap, for quality control of your prep



Cryo-EM: time consuming, for high resolution structures

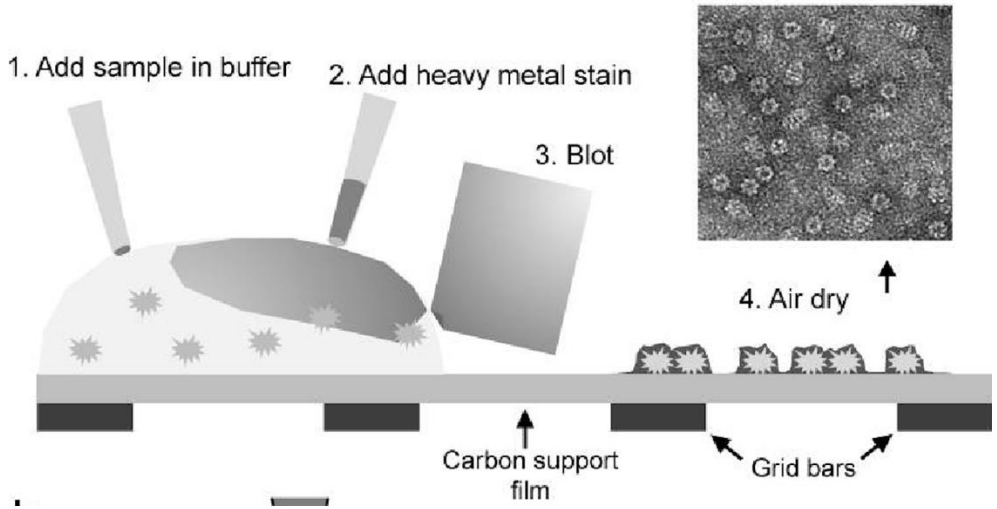
M. Carroni and HR. Saibil (2016). Cryo electron microscopy to determine the structure of macromolecular complexes

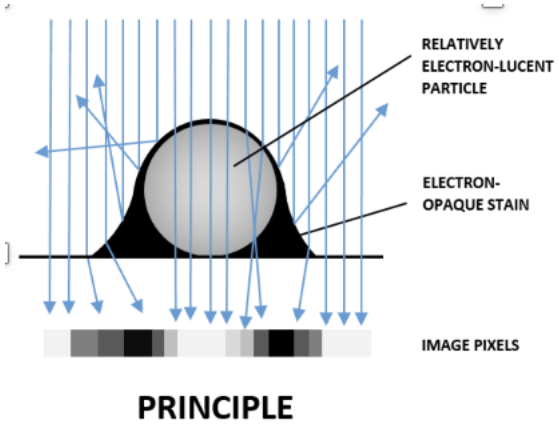
Negative stain screens as quick test to quality control your prep



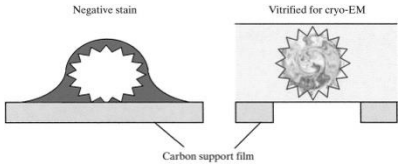
- A quick and cheap approach with high contrast and at RT
- Few uls at ugs/ml protein concentration
- Macromolecules are embedded in heavy atoms (e.g. uranyl acetate)
- Give Low resolution 3D reconstruction
- Good to check sample stability and homogeneity
- Good also to check sample concentration

Cons: Need to screen again for cryo conditions to obtain atomic structures.

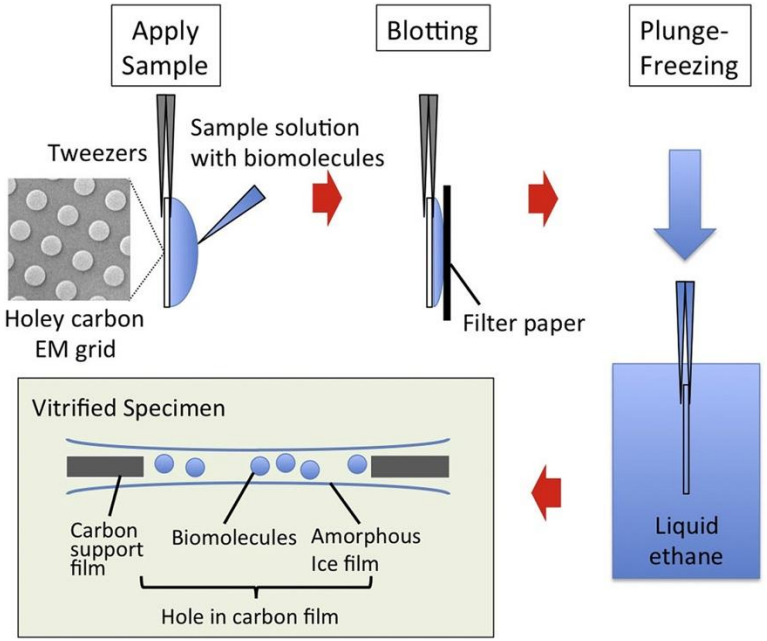


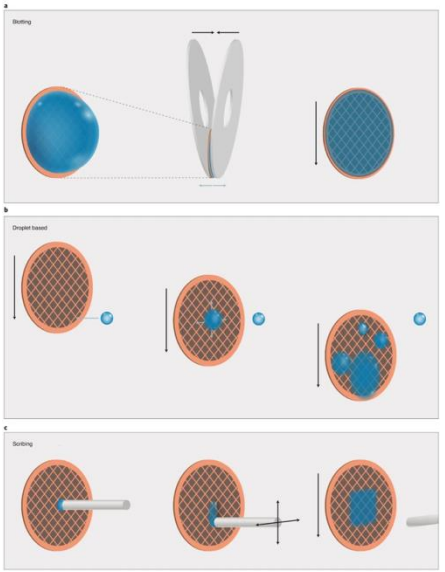


Cryo screens for atomic resolution structures



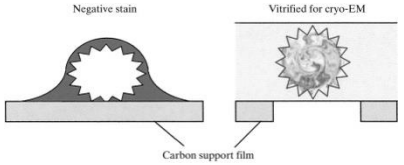
- More time consuming to find the best conditions
- Few uls protein at around 1-5 mg/ml
- Macromolecules are embedded in vitreous ice (Nobel Prize, Jacques Dubochet, 2017)
- Sample is kept in native conditions by plunge freezing (in liquid ethane, -160° C)
- Allow to get atomic resolution 3D reconstruction (best today at DCI; 1.23A for apoferritin)





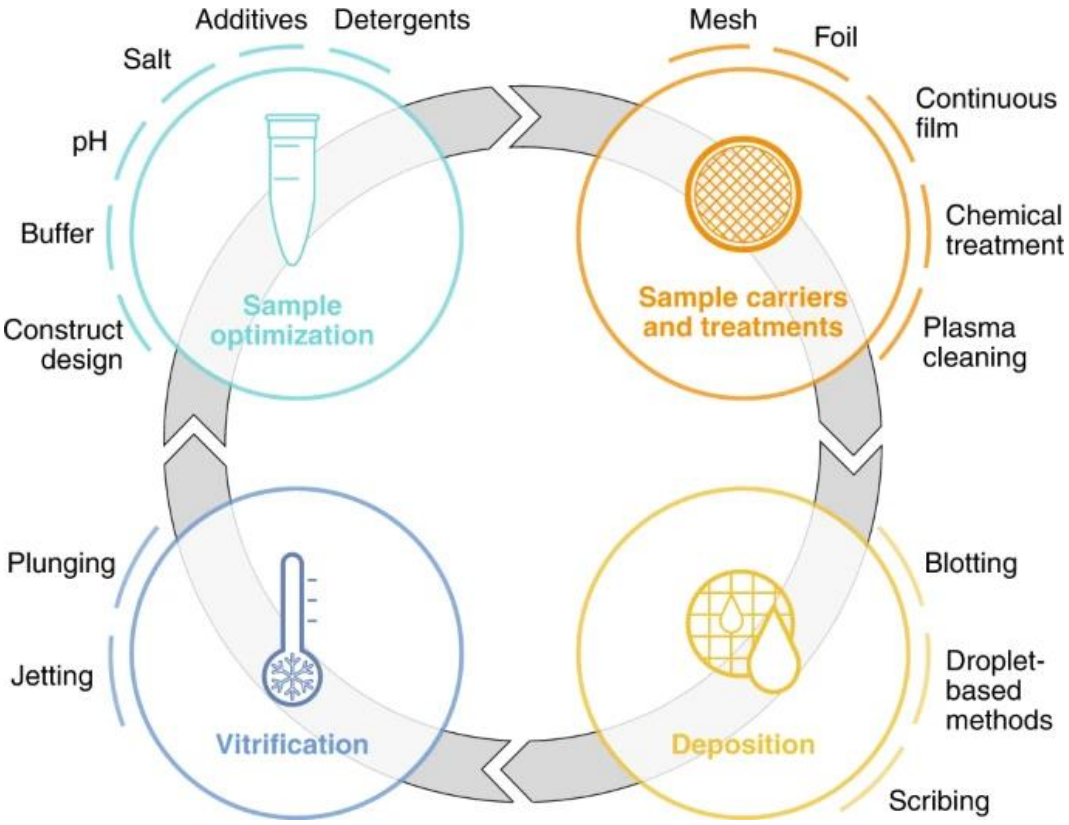
Kazuyoshi Murata & MatthiasWolf: Cryo-electron microscopy for structural analysis of dynamic biological macromolecules

Optimization of cryo conditions- screening



The problems encountered in practice:

- preferential orientation of particles may occur within thin films
- unexpectedly low numbers of particles may be found within holes
- particles may disintegrate within thin aqueous films
- unexplained aggregation of sample material may be observed.

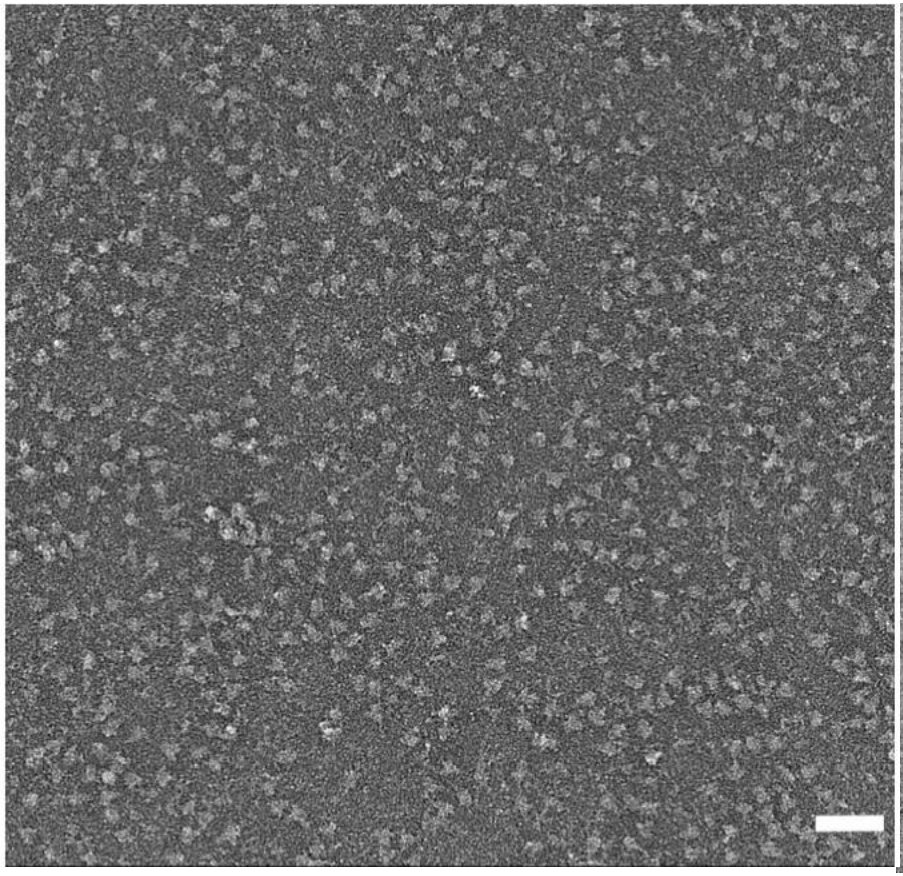


https://www.nature.com/articles/s41592-021-01130-6

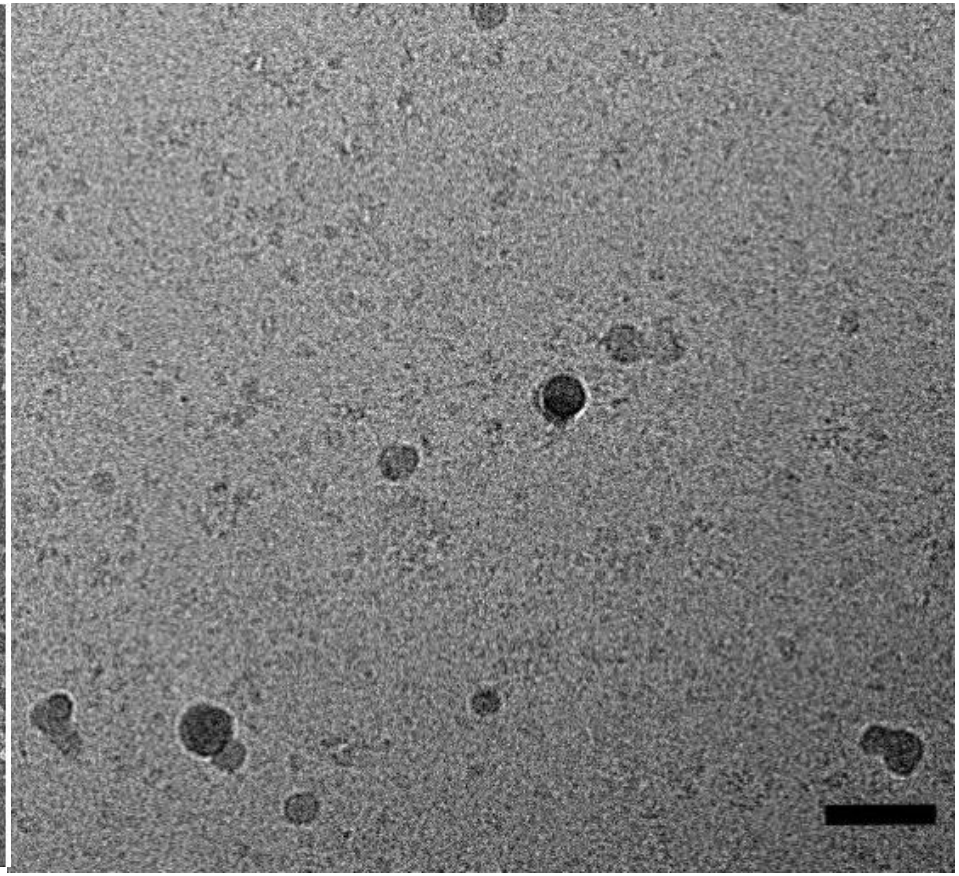
Cryoem101.org

Very nice resource with videos

EM micrographs of Spike mixed with purified ACE2

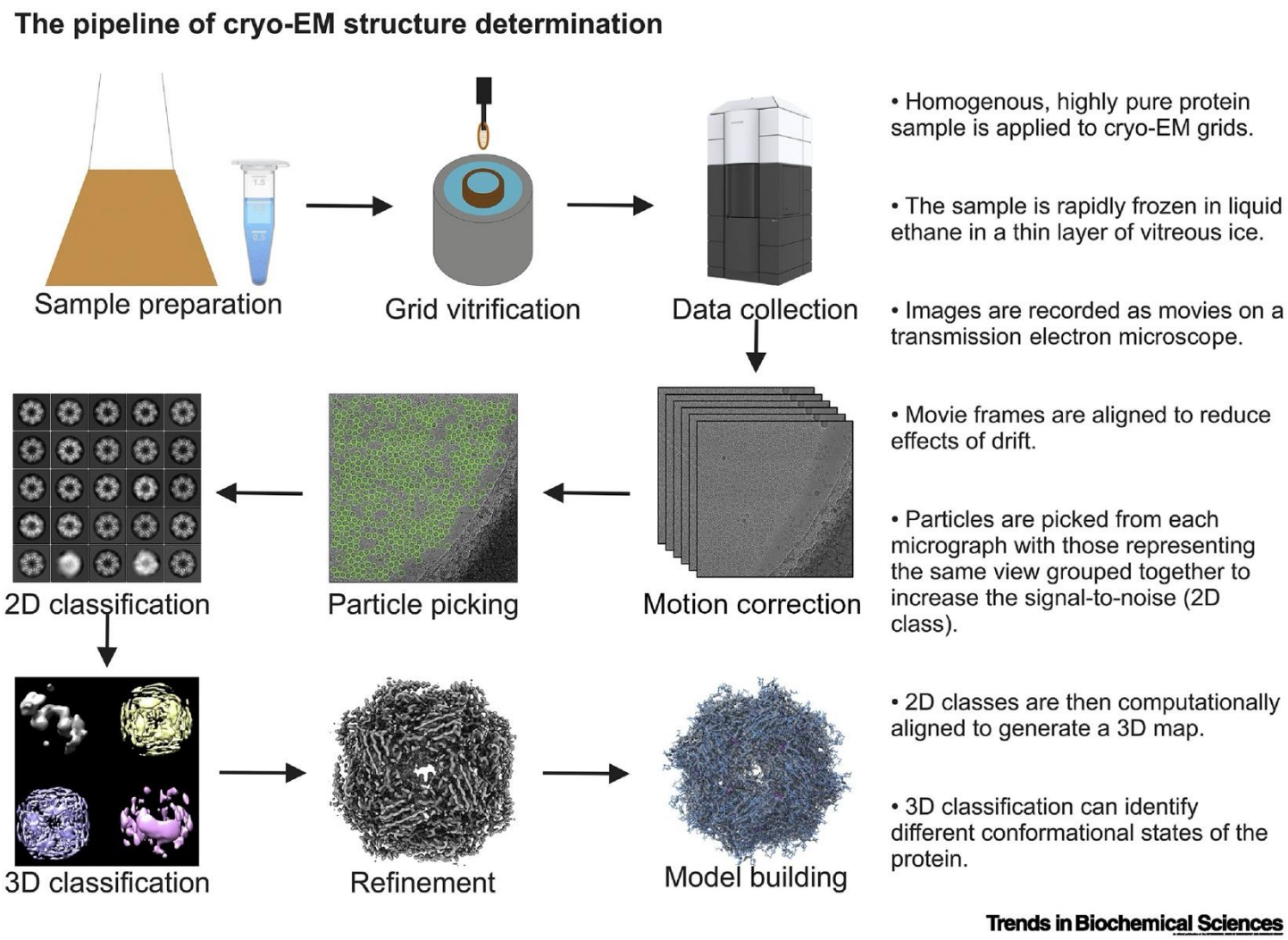


Negative stain



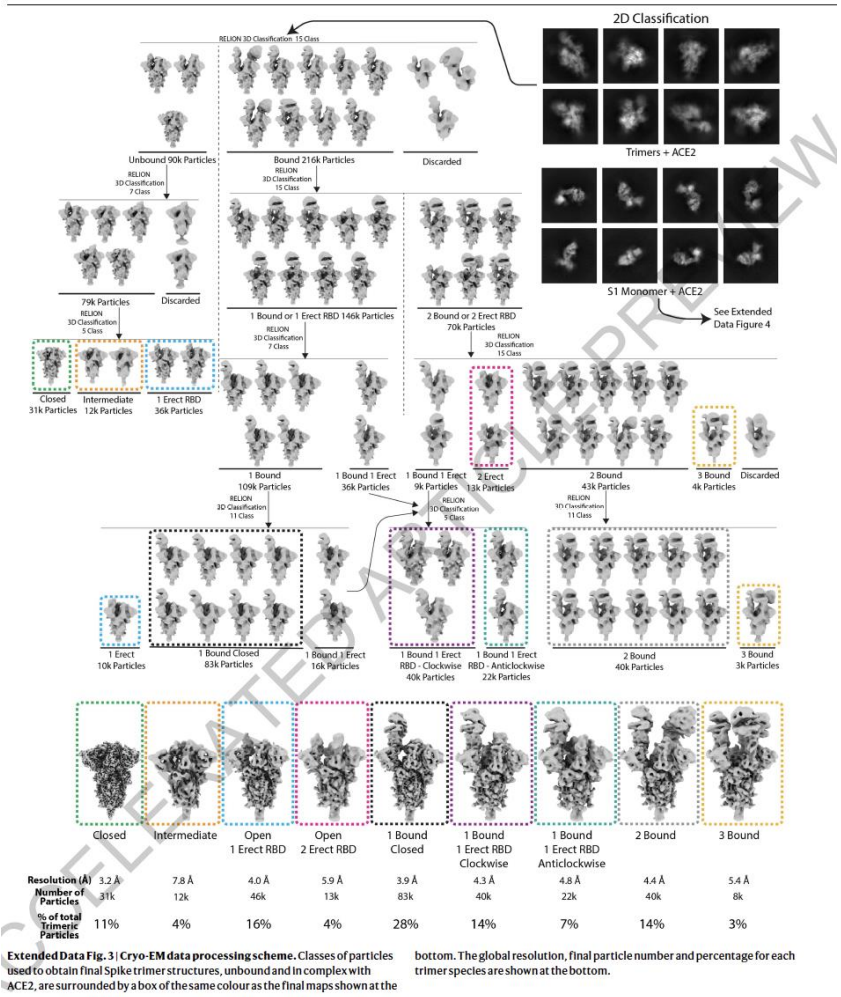
Cryo condition

Data collection, particle selection, reconstruction to model building, refinement and validation

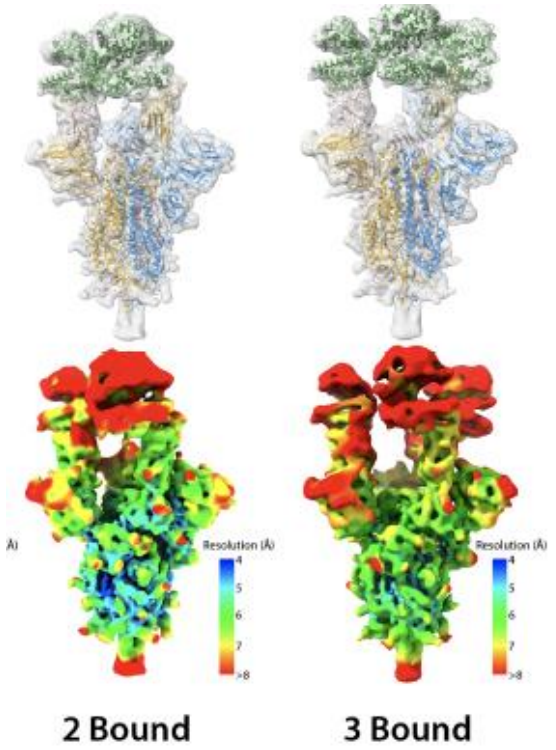


https://leaderna.com/technology/

A lot of information obtained on a single EM grid! Integrative Structural Biology is essential



Extended Data Fig. 3 | Cryo-EM data processing scheme. Classes of particles used to obtain final Spike trimer structures, unbound and in complex with ACE2, are surrounded by a box of the same colour as the final maps shown at the bottom. The global resolution, final particle number and percentage for each trimer species are shown at the bottom.



Benton et al (2020). Receptor binding and priming of the spike protein of SARS-CoV-2 for membrane fusion

Advice: How to decide the methods

To consider:

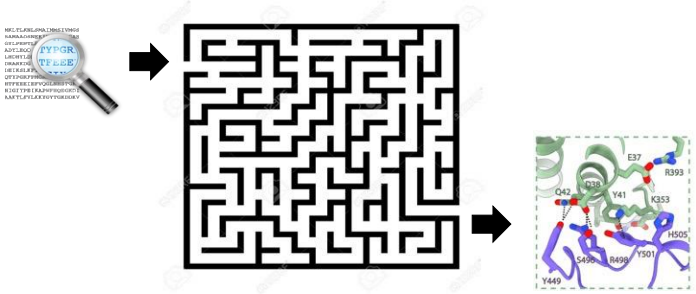
- **The Biological question**: nature and size of proteins/complexes and type of results wished

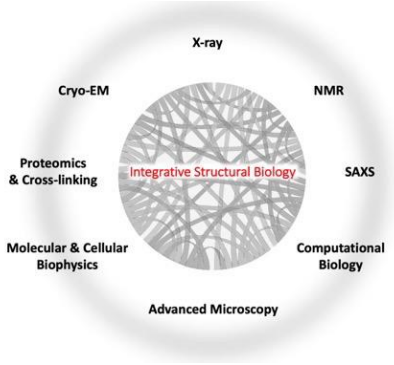
  Pros/cons for each method. **Each project is evaluated carefully**.

  Roughly: SPR cryoEM for proteins > 60kDa & big complexes & difficult to produce; NMR for small flexible proteins < 40kDa; crystallography for drug target & protein any size; but with limiting factor being the formation of crystals

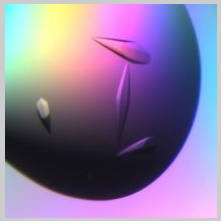
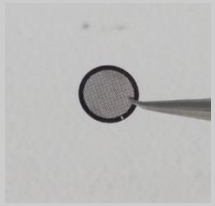
- **Time and funding**

- **Expertise and access to technology**





- ü Methods complement each other
- ü Often performed in parallel

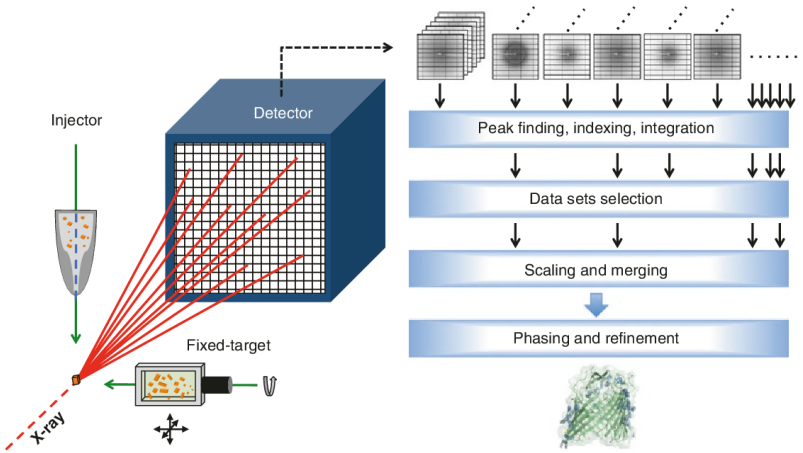
Techniques	PROS	CONS
X-ray crystallography 	✓ Provide very detailed atomic information ✓ Easy to perform ✓ Not expensive ✓ Software free and user friendly ✓ No size limitation ✓ Synchrotron facilities around the world	✓ Need to form crystals ✓ High protein quantity ✓ Difficult for membrane proteins ✓ Difficult for flexible domains
BioNMR 	✓ Small flexible proteins ✓ In solution ✓ Info on dynamics ✓ Info on ligand binding	✓ Not for big complex. (samples<40kDa) ✓ Low through-put ✓ High expertise ✓ High protein quantity, labeled ✓ Expensive
Single-particle EM 	✓ Big complex, membrane proteins ✓ Less sample needed (5-10 times less than crystallography) ✓ Achieve atomic resolution	✓ Still challenging for small proteins <60kDa ✓ High expertise ✓ Low Throughput ✓ High-end equipment ✓ Expensive

RECAP of techniques with videos on PDB101

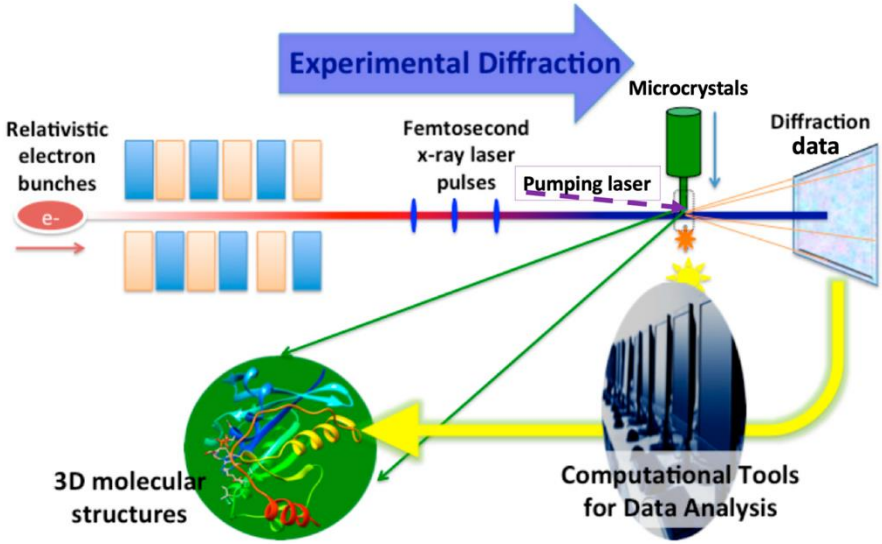
https://pdb101.rcsb.org/learn/guide-to-understanding-pdb-data/methods-for-determining-structure

Other techniques on the X-ray side

SSX: Serial Synchrotron X-ray crystallography



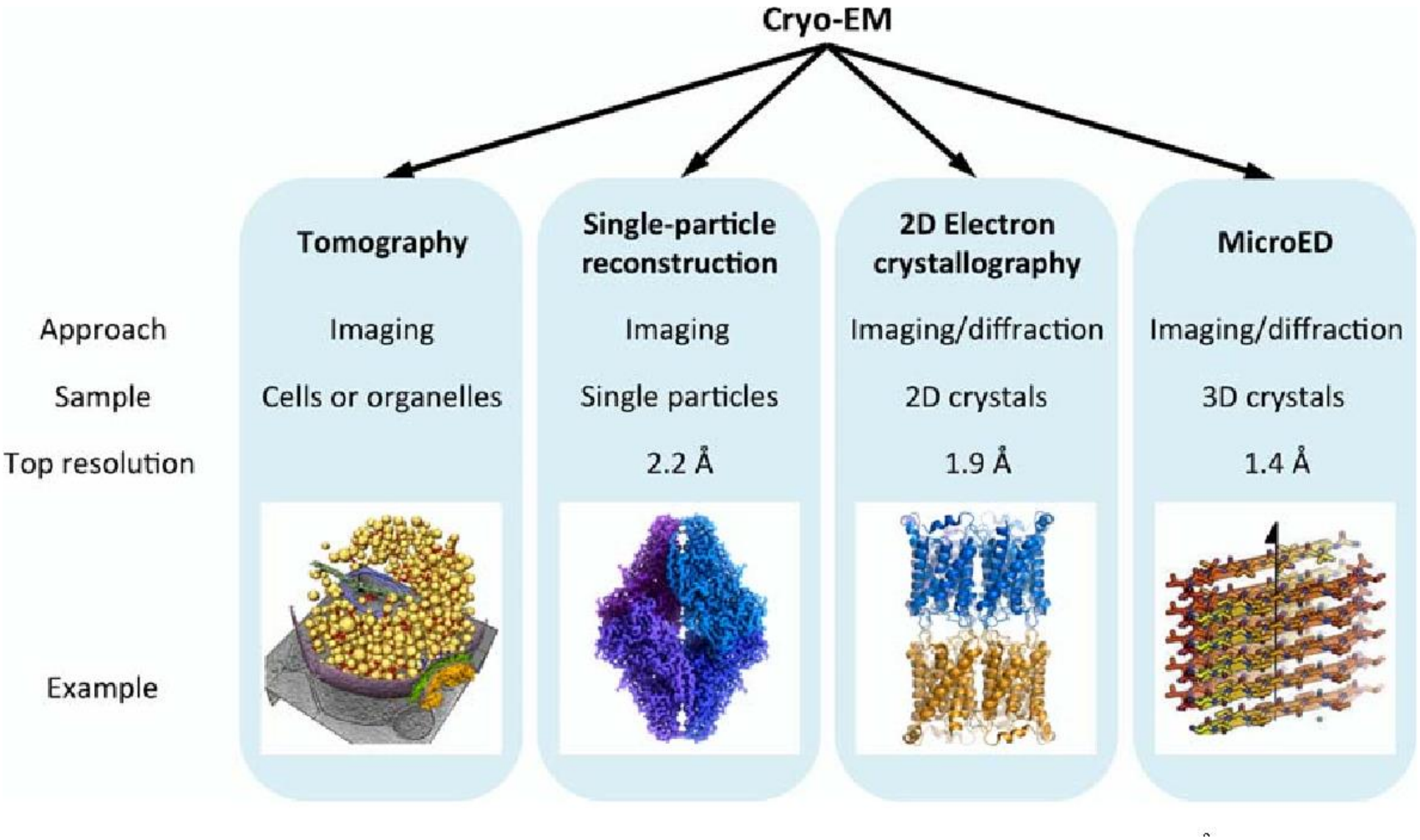
XFEL: X-Ray Free-Electron Laser



Kay Diederichs, Meitian Wang (2017) Serial Synchrotron X-Ray Crystallography (SSX)

Lui et al (2019) The XFEL Protein Crystallography: Developments and Perspectives

Other techniques on the cryoEM side

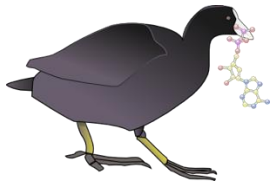


Shian Liu (2007) Atomic resolution structure determination by the cryo-EM method MicroED.

Practicals at PTPSP

Dr. Kelvin Lau and Dr. Yoan Duhoo

- Crystallization screens and look at and analyze X-ray data with Coot software



--> COOT Xray and EM model building: https://phenix-online.org/download/other.html

- CryoEM sample preparation and Visualization of EM data files with ChimeraX

--> ChimeraX: https://www.cgl.ucsf.edu/chimerax/

Warning:

Models and maps are deposited in PDB database without being peer-reviewed, nor being curated by the database, thus always important to check the data before stating a project