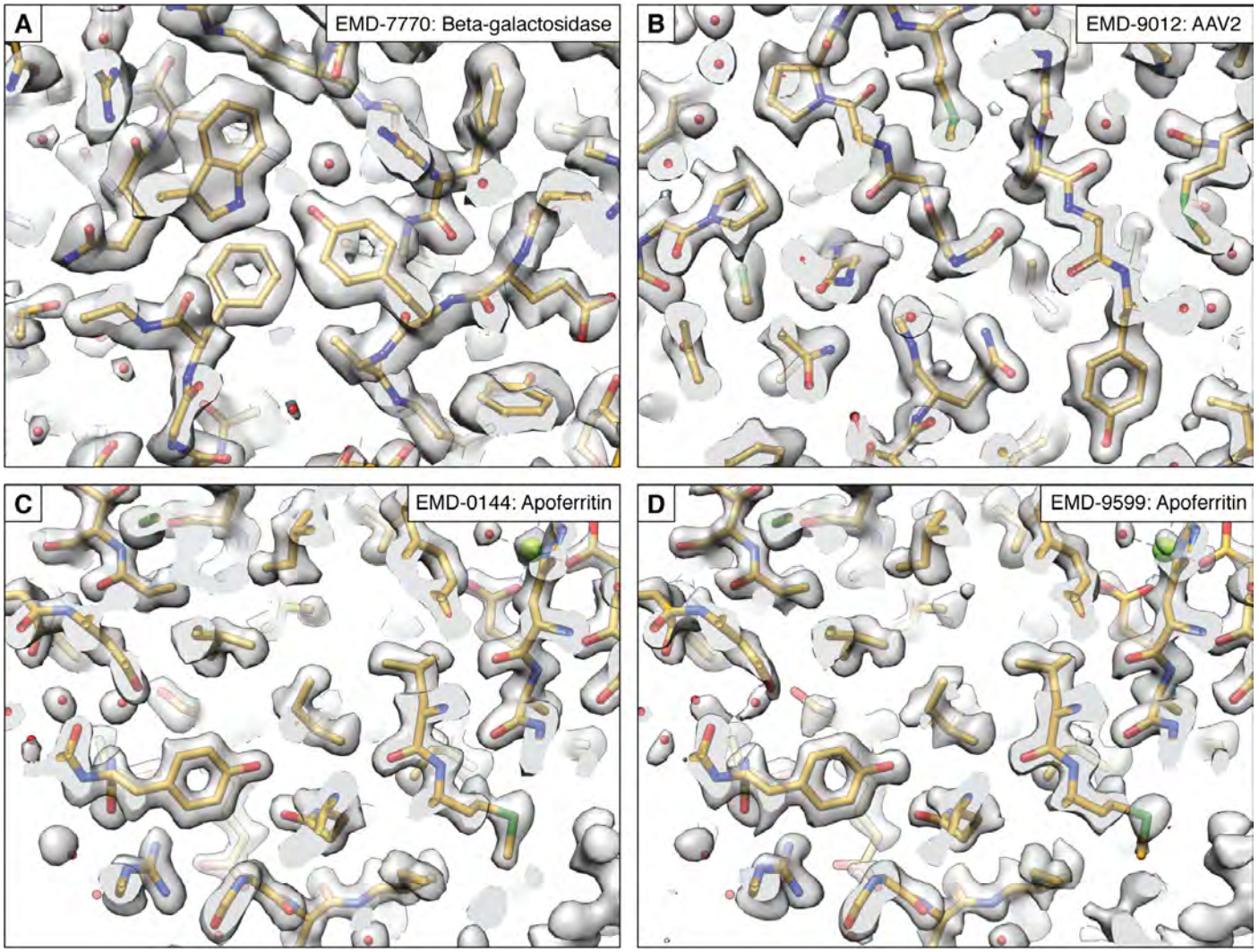
Part II

FROM SAMPLES TO STRUCTURES



(Dmitry Lyumkis, JBC, 2019)

Prélude

- An X-ray or EM structure is the Signal/Noise (S/N)-enhanced image of homogeneous specimen.
- Cryo-EM (especially single-particle analysis and microED) allows protein structural determination to sub-2Å resolution.
- For many, it is still challenging to **"routinely"** obtain high-resolution data and structures.
- Many aspects of protein preparations are EM specimen preparation still depend on empirical experience; thus difficult to master.

How do we start? 4 aspects.

Core knowledge

Biochemistry & Sample preparation

Data collection

Processing & Data analysis

Outline

- **Core knowledge (Pt I & else)**
- Evaluation of protein biochemistry
- EM sample preparation
  (Single particles & Micro-/2-D crystals)
- Data collection
- Image processing and analysis

Resources: Local (Ottawa)

uOttawa

Main Campus

Vitrobot (?)

Gatan 626 cryo holder (?)

FEI Tecnai G2
120 kV

Faculty of Medicine

JEOL JEM-1400+ (?)
120 kV

Resources: Local (National Centre)

McGill: Facility for Electron Microscopy Research (FEMR)

Vitrobot

Gatan 626 cryo holder

FEI Tecnai G2
120 kV

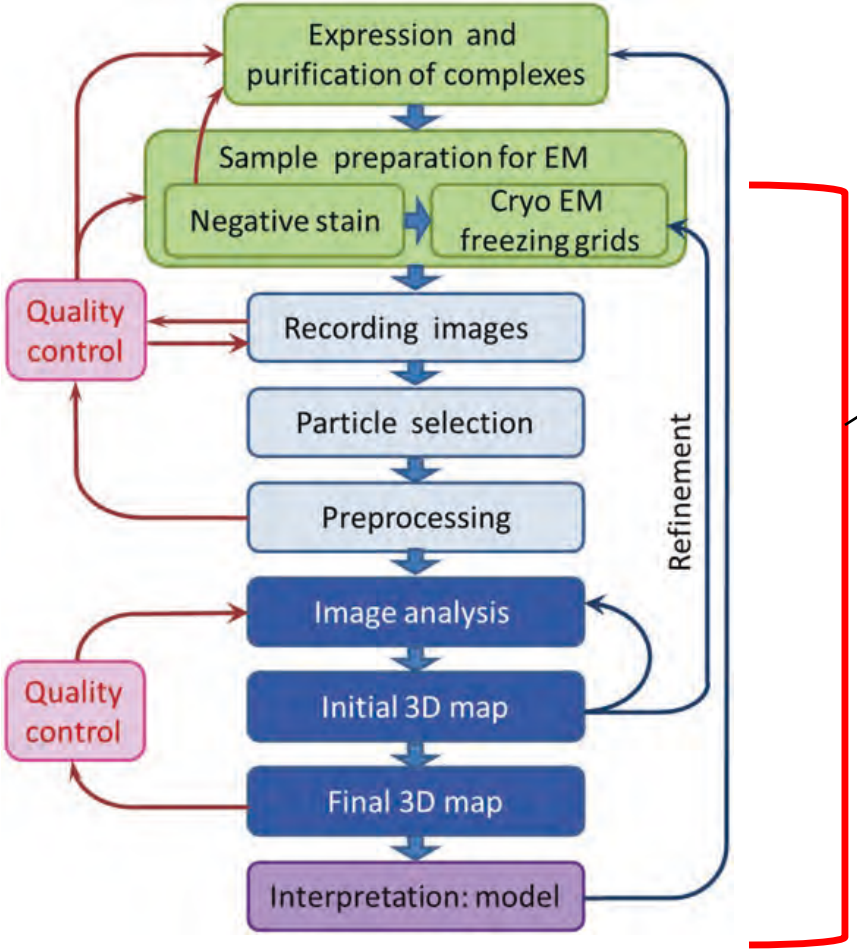
FEI F20
200 kV

FEI Titan Krios
300 kV

Resources: North America

https://commonfund.nih.gov/CryoEM

Cryo-EM Workflow

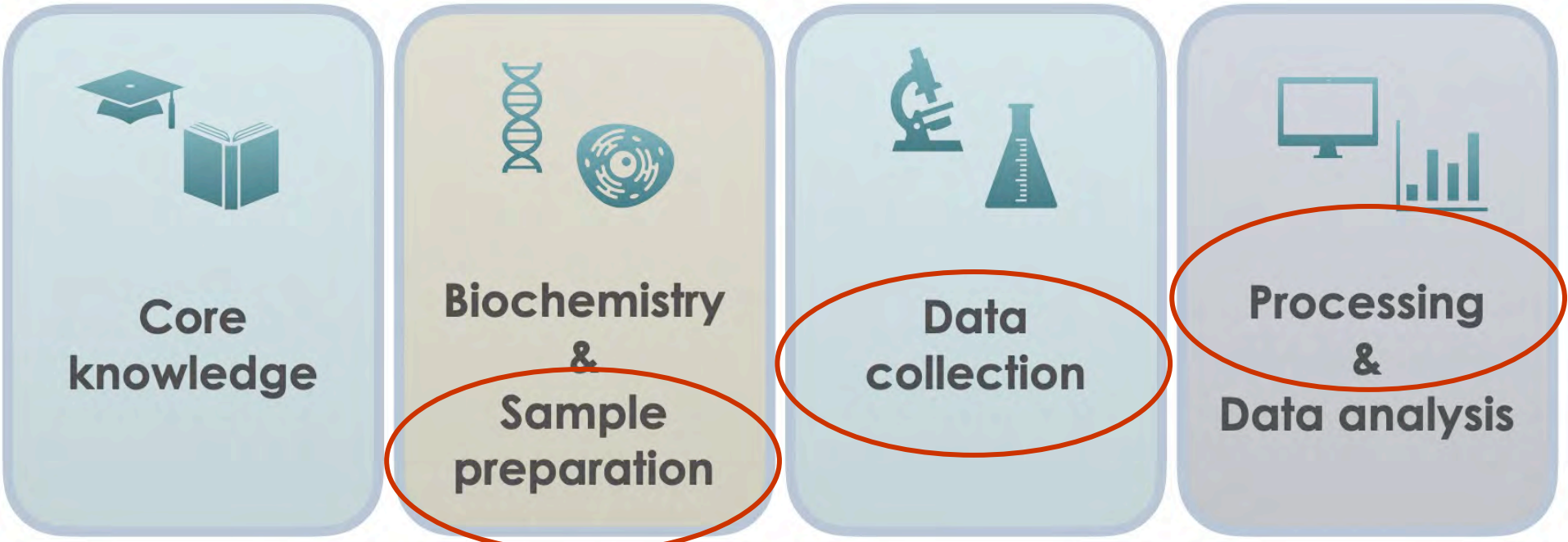


Theoretically, 2 Å within hours!!

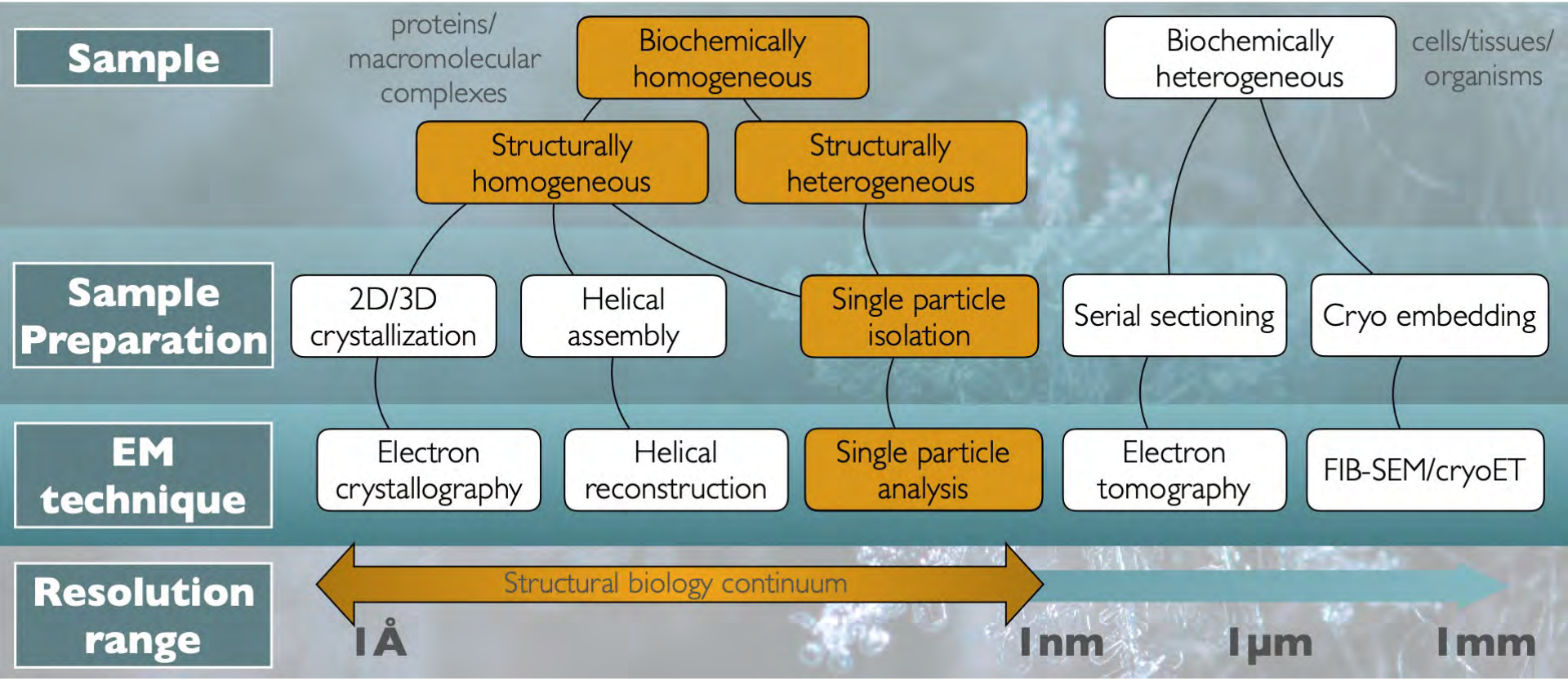
Fig. 1 Workflow of EM structural analysis. In *green* is the experimental part of structural analysis. The computational part is shown in *light* and *dark blue*; the initial steps of processing are shown in *light blue*. They include image frame alignment, CTF correction, normalisation, and filtering. The subsequent steps—alignment, statistical analysis, determination of particle orientations, and initial three-dimensional reconstruction (3D)—are shown in *dark blue*. The final step (*light purple*) is the interpretation of the maps obtained

(Costa et al, Meth Mol Biol, 2017)

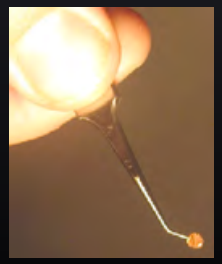
Bottlenecks

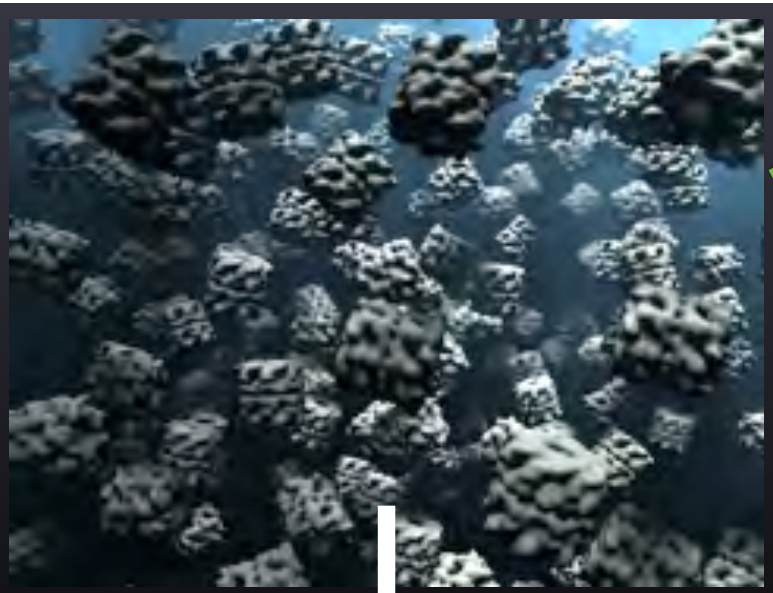


Bottleneck: Sample Preparation

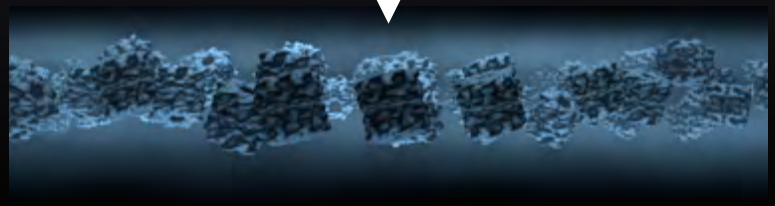


Bottleneck: Sample Preparation





~3µl



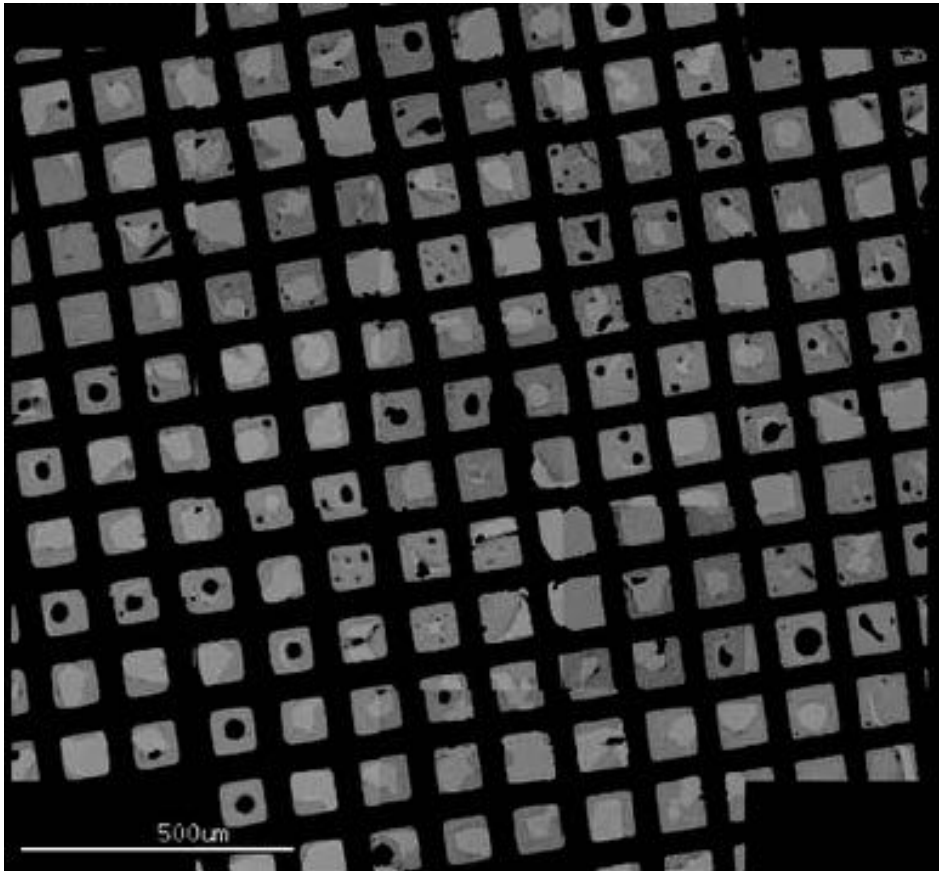
<3nl

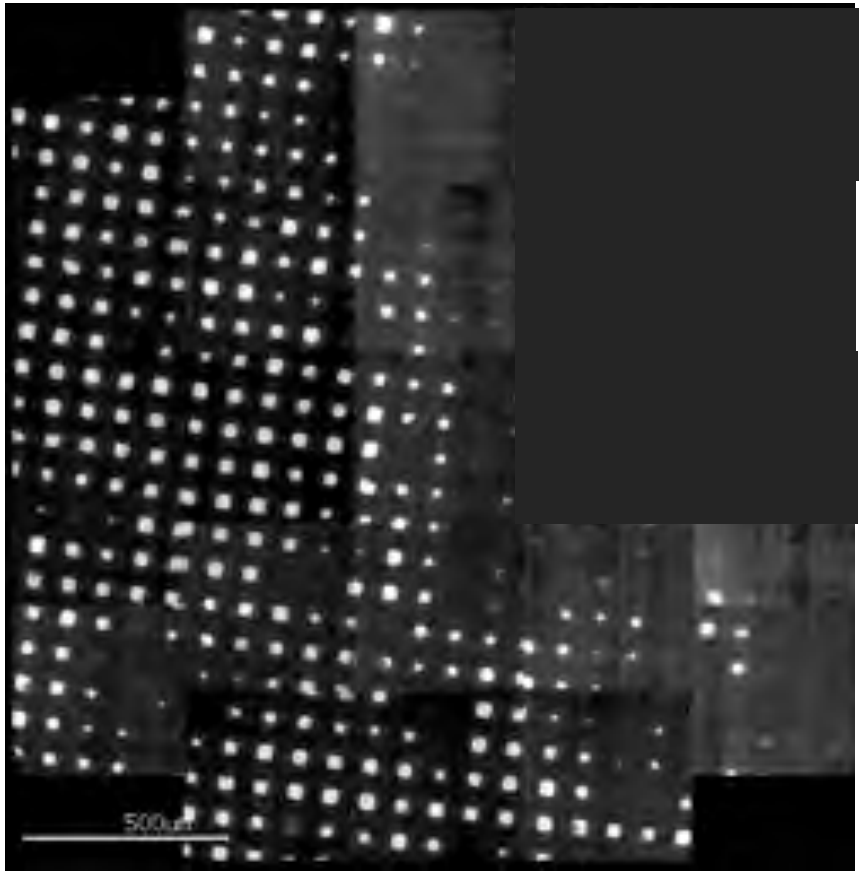
>99.999%

<0.001%

>100,000 potential imaging targets; most of them are not usable.

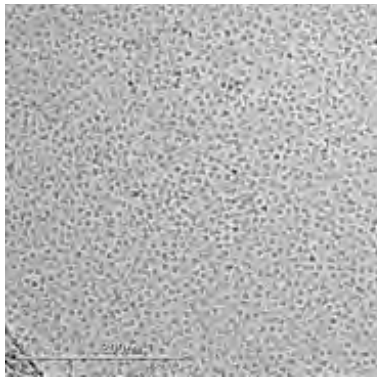
Bottleneck: Data Collection



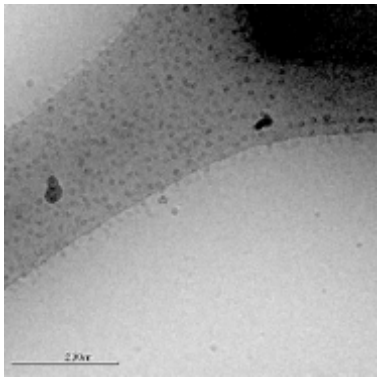


Bottleneck: Data Collection

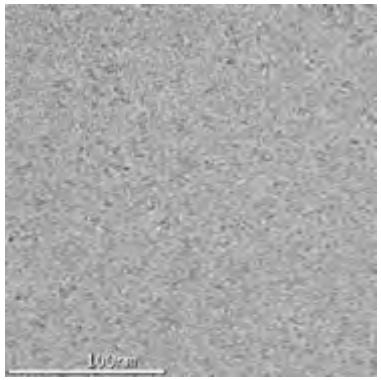
Aggregating in ice



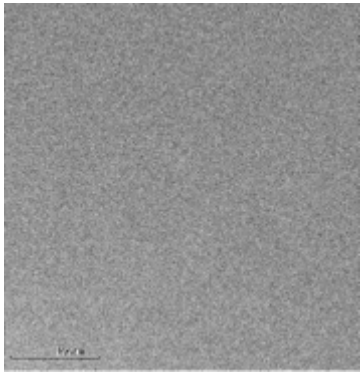
Preferred orientation



Particles not going into holes

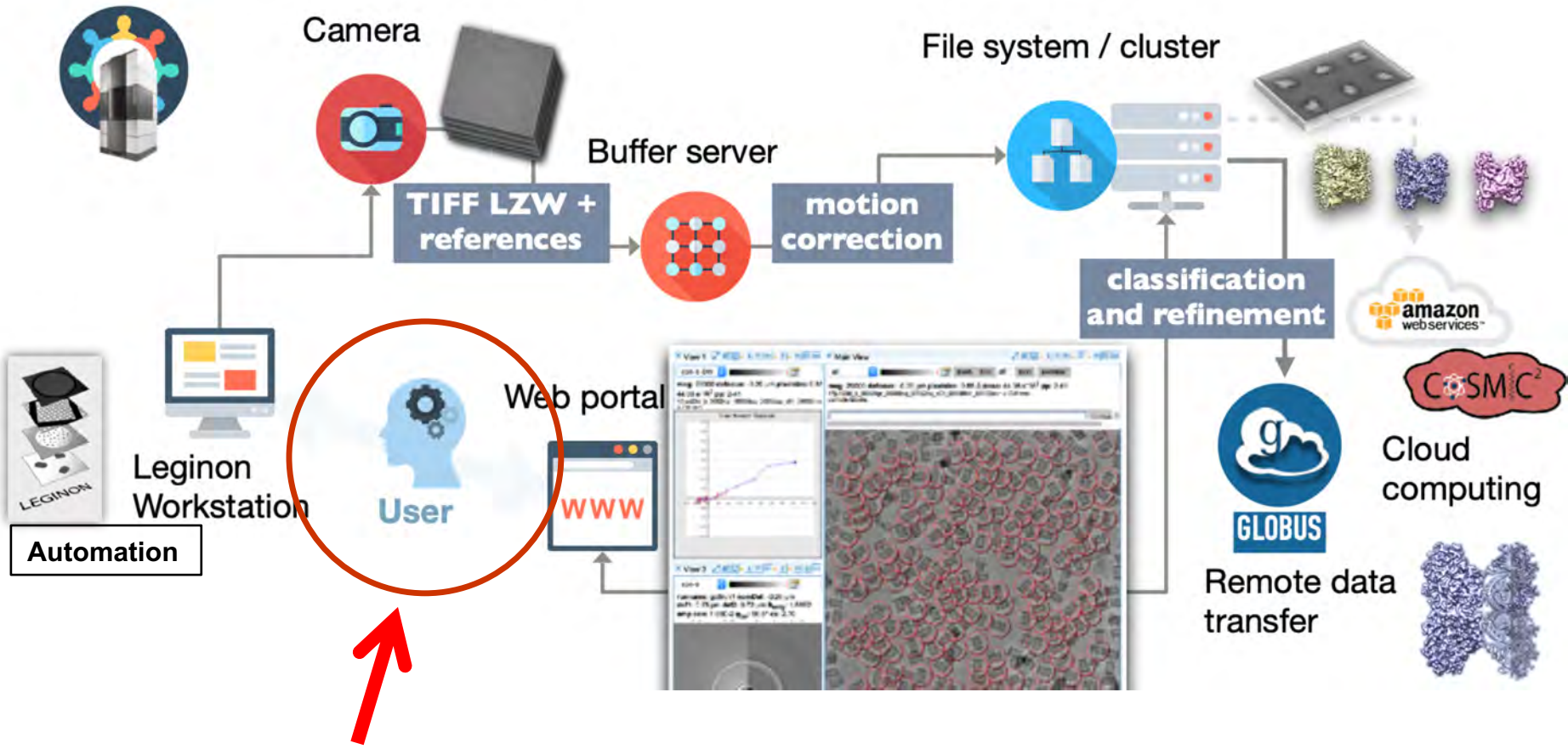


Rejecting 90% of particles



Particles disappearing in ice

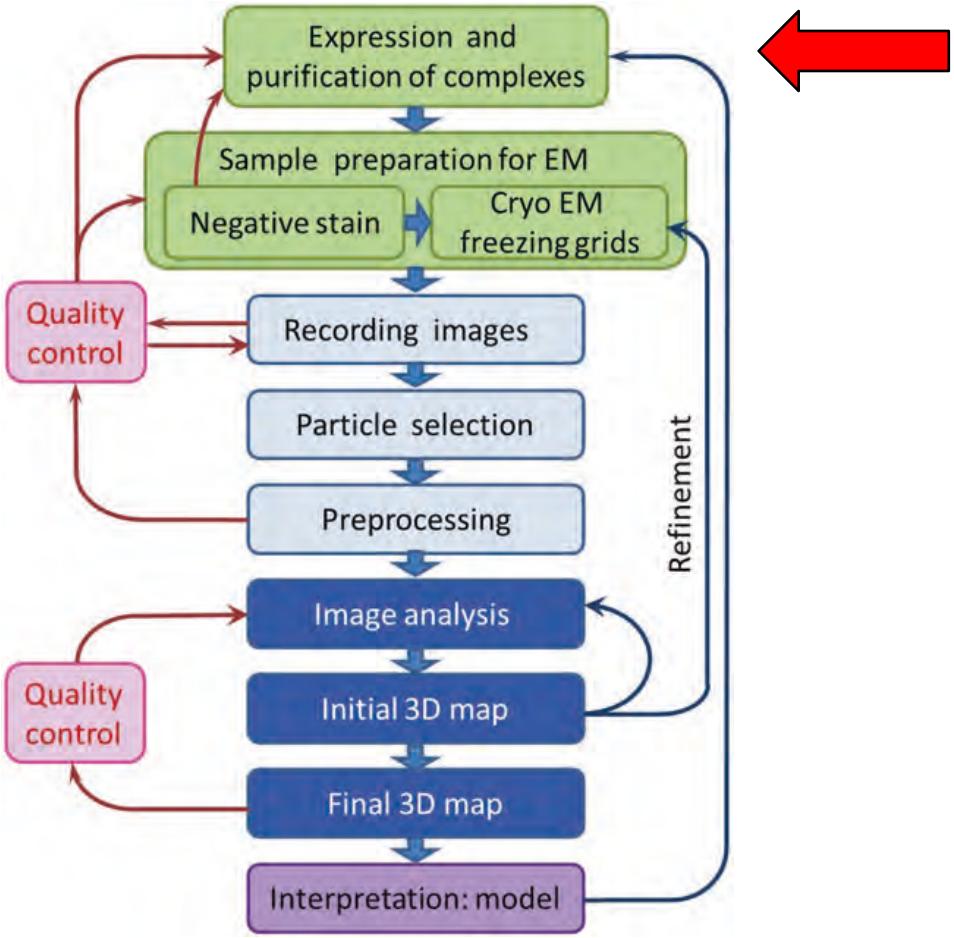
Bottleneck: "Pre-"Processing



Outline

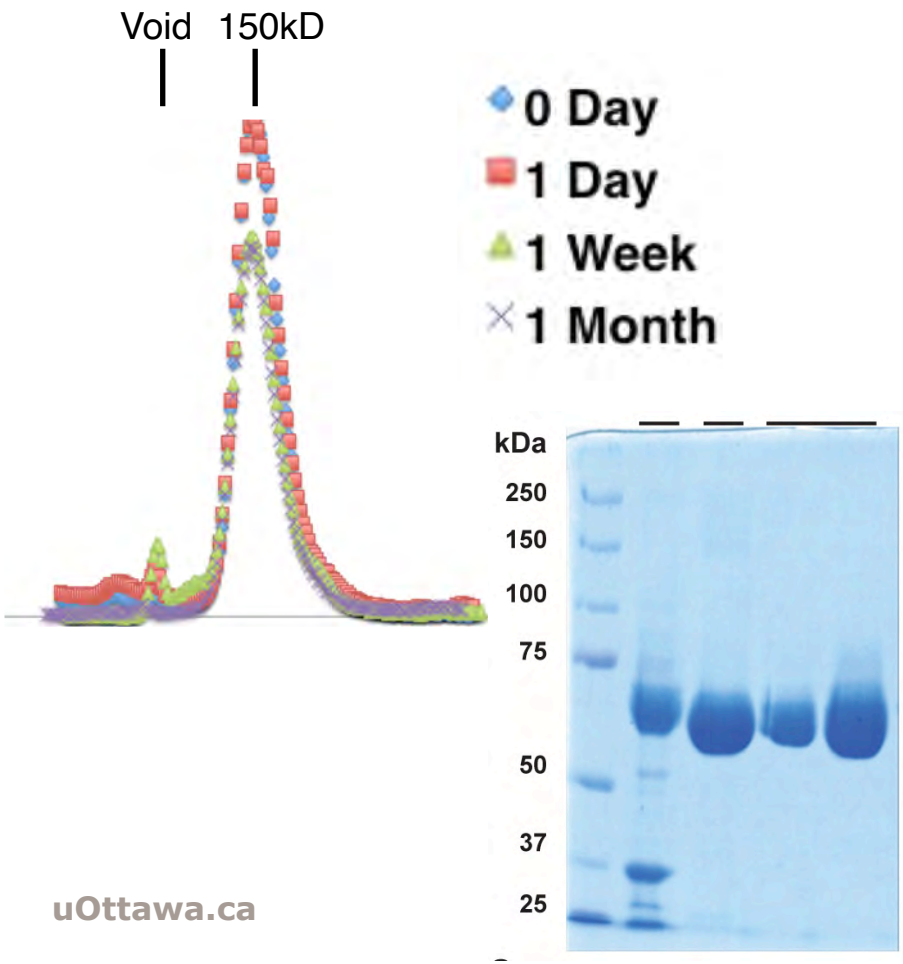
- Core knowledge (Pt I & else)
- **Evaluation of protein biochemistry**
- EM sample preparation
  (Single particles & Micro-/2-D crystals)
- Data collection
- Image processing and analysis

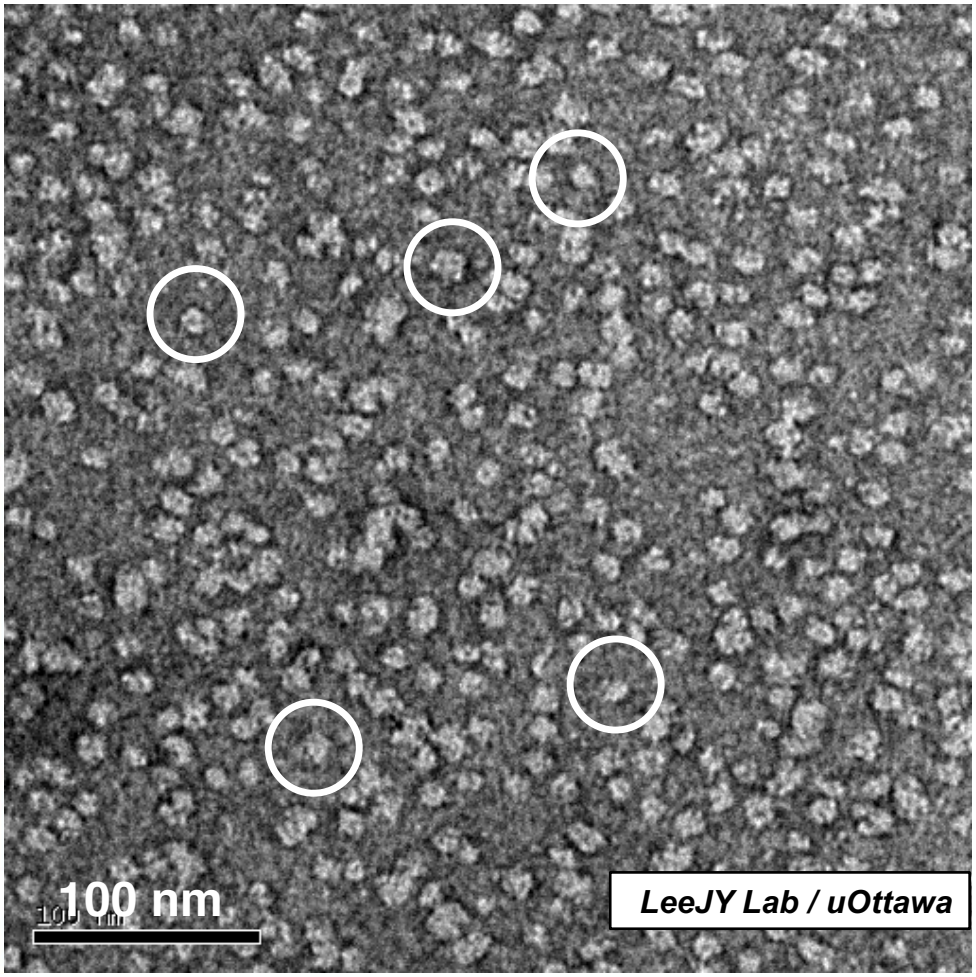
Cryo-EM Workflow



Protein Sample Quality

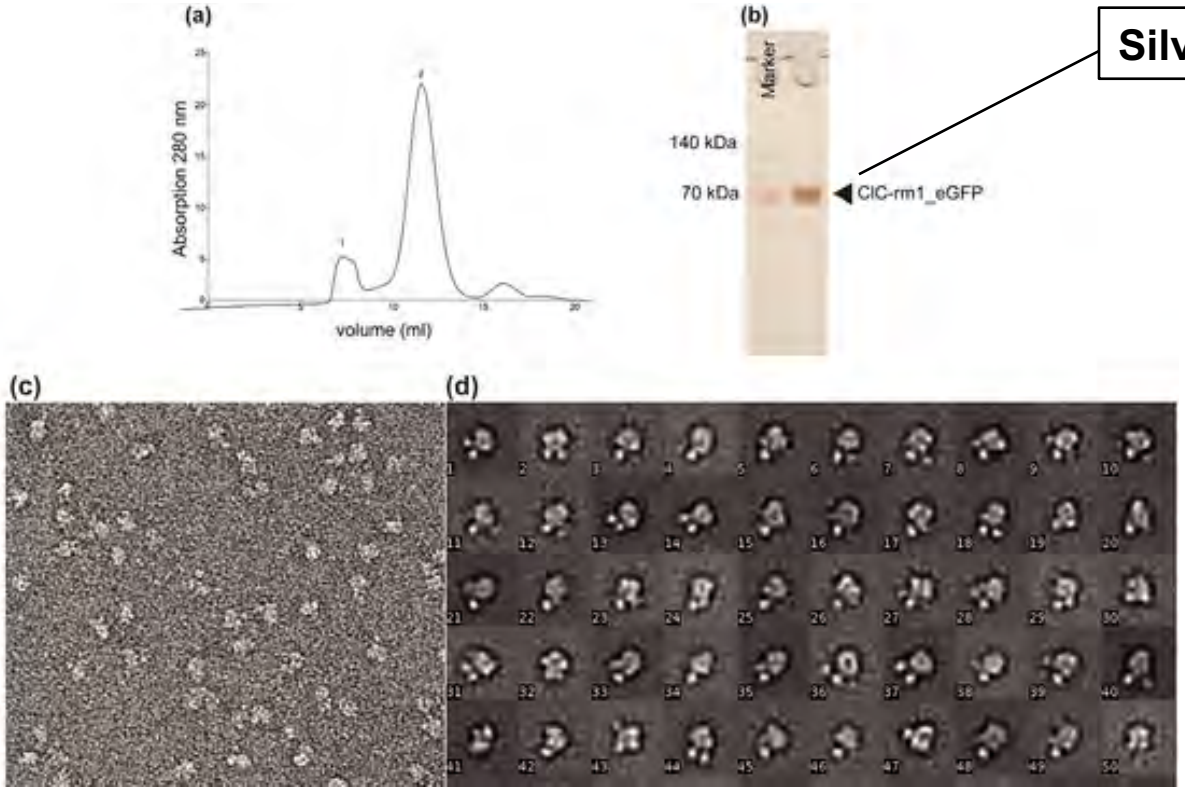
(Same standard as X-ray crystallography)





Protein Sample Quality

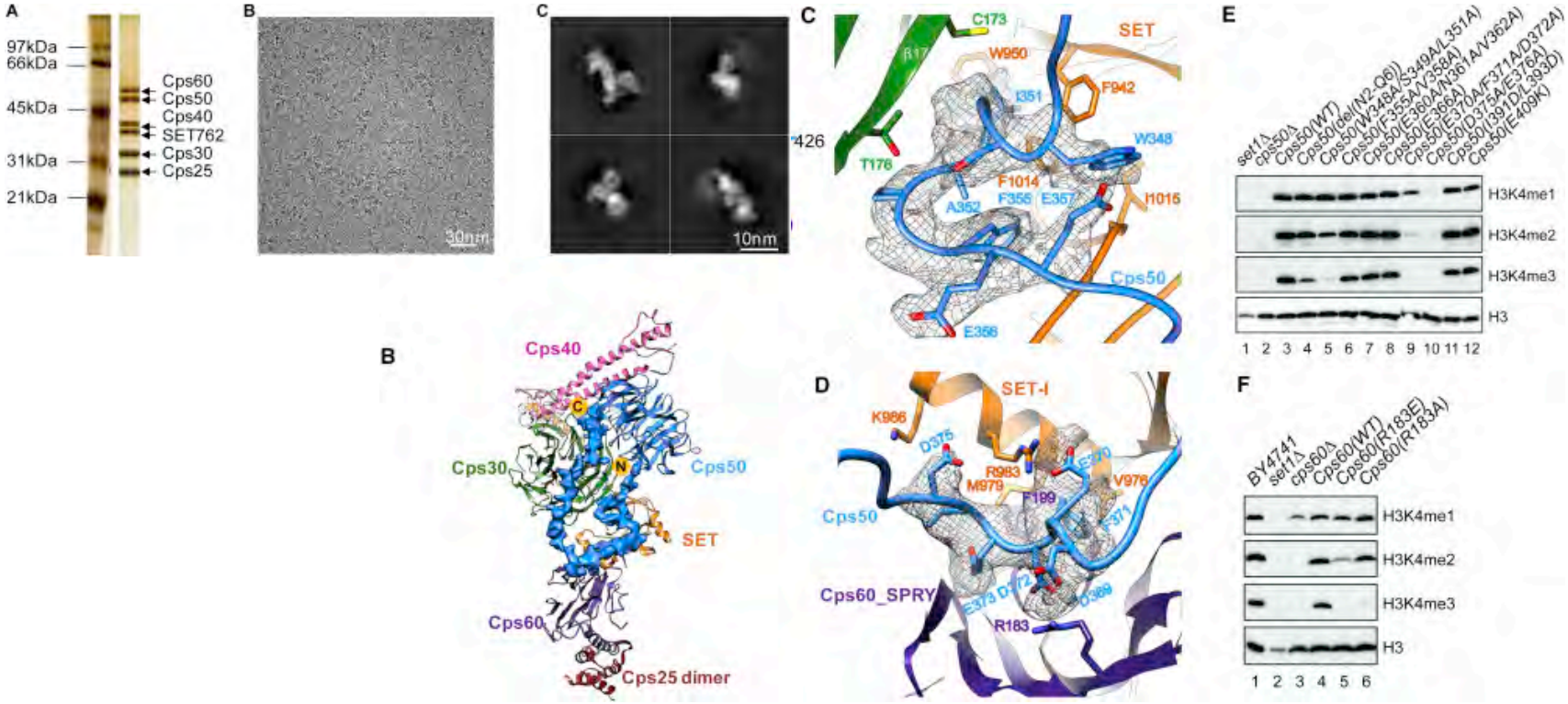
(But no need to purify several mg proteins!)



(Abeyrathne & Grigorieff, PLOS ONE, 2017)

Protein Sample Quality

(Functionally characterized)

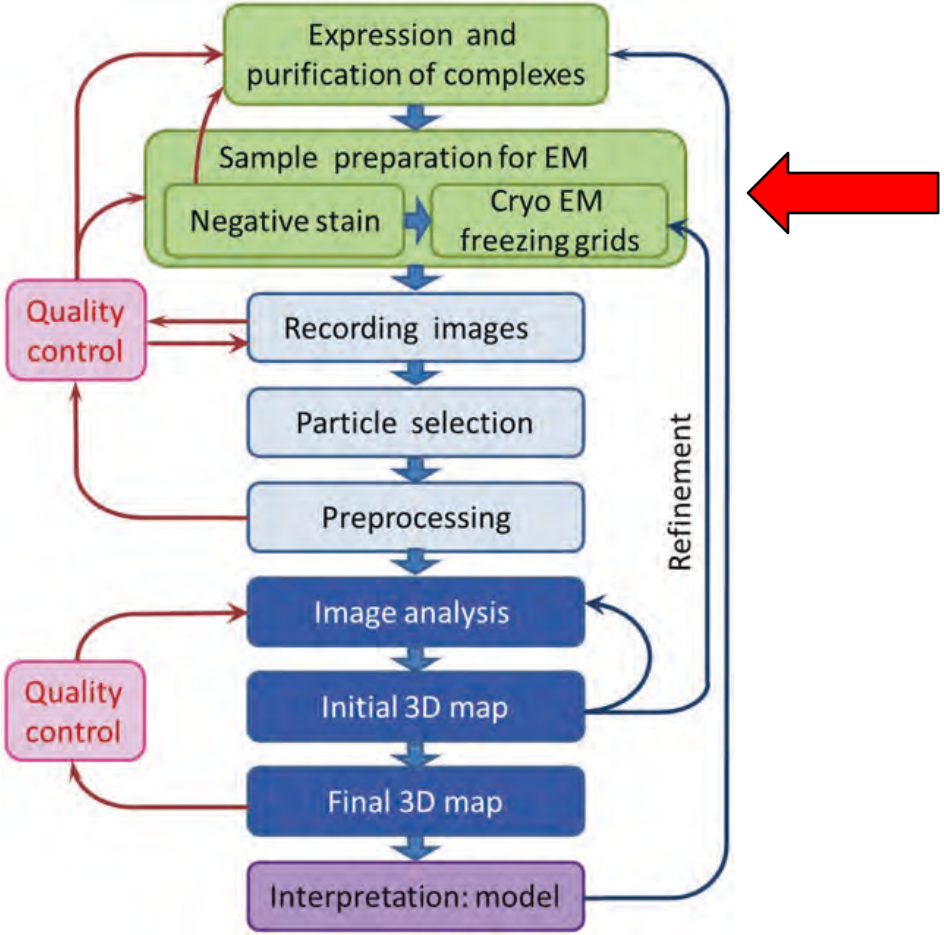


(Qu/Takahashi/Yang et al, Cell, 2018)

Outline

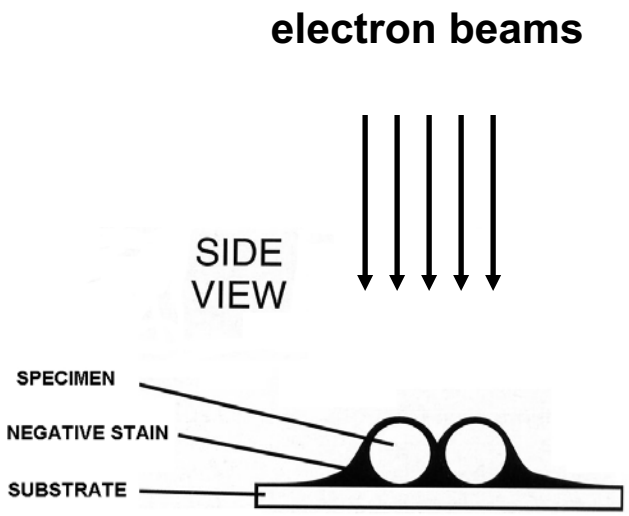
- Core knowledge (Pt I & else)
- Evaluation of protein biochemistry
- **EM sample preparation**
  (Single particles & Micro-/2-D crystals)
- Data collection
- Image processing and analysis

Cryo-EM Workflow



EM Samples: Negative Staining

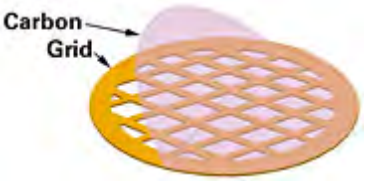
- Principle:
  - Embedding objects in a layer of heavy-metal salts that surround the proteins like a shell.
  - Shape of objects are visible in contrast to the optically opaque stains.
- Benefits:
  - Small amount of proteins (0.01 mg/mL)
  - Easy and quick (preparation and imaging)
  - No need of high-end microscope; diagnostic
- Downsides:
  - Low resolution (*e.g.*, high noise from stains)
  - Artifacts (lack of hydration)

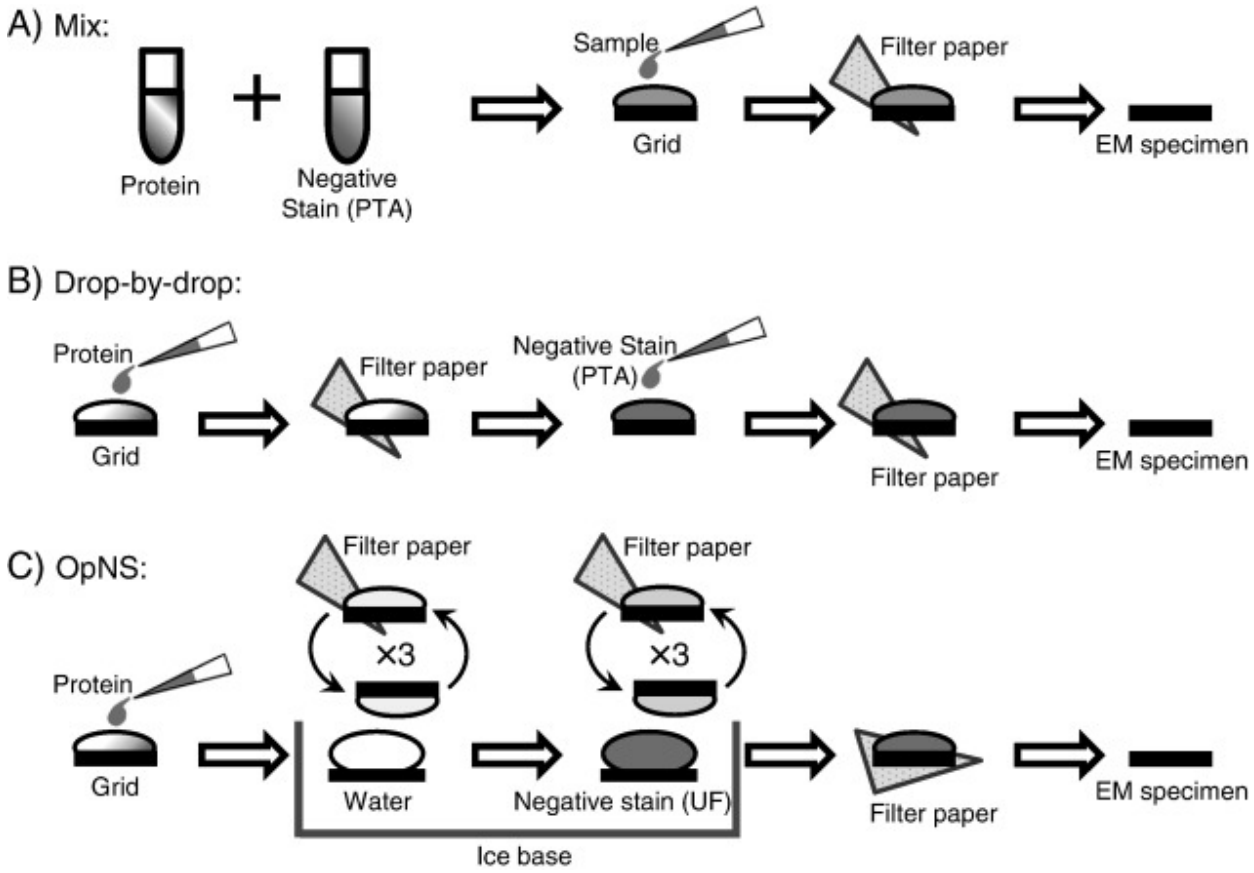


(Brenner & Horne, BBA, 1959)

EM Samples: Negative Staining

<3 nm thick (ultrathin)

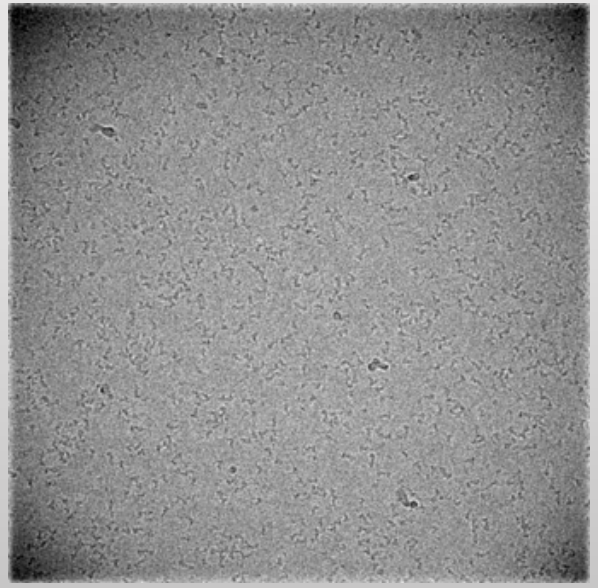




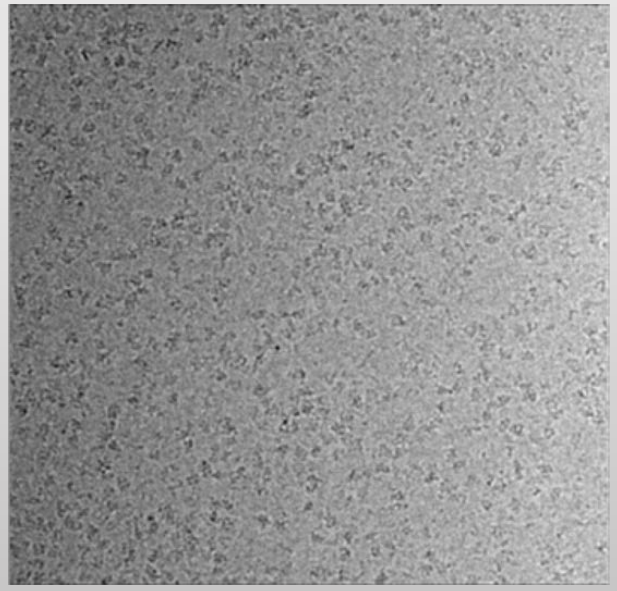
(Zhang et al, BBA, 2013)

EM Samples: Protein Concentrations

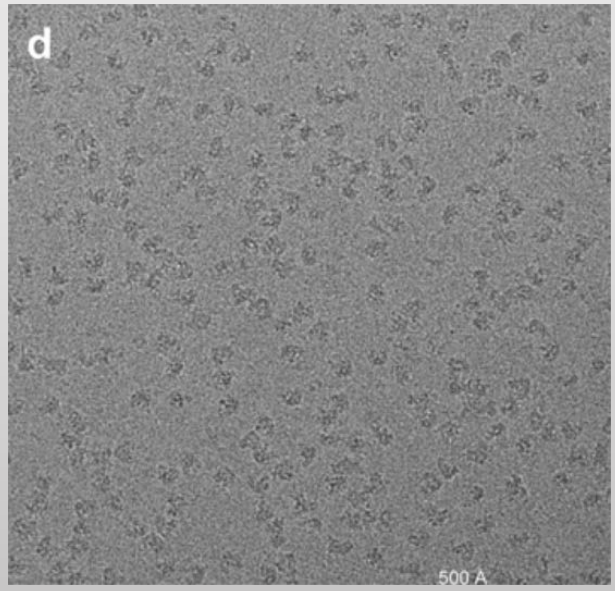
- Negative staining: 0.01-0.05 mg/mL
- Cryo: 0.1-5 mg/mL



CTF3 complex, 130 kDa
0.2 mg/ml



Cas12a-AcrVA4/5 complex, 200 kDa
3 mg/ml



APC/C complex, 1.2 MDa
0.1 mg/ml on continuous carbon film

EM Samples: Buffer Conditions

- Using negative-stain EM, screen buffer conditions to evaluate protein stability on EM grids.

- Avoid the following conditions:
  - High salt
  - High glycerol

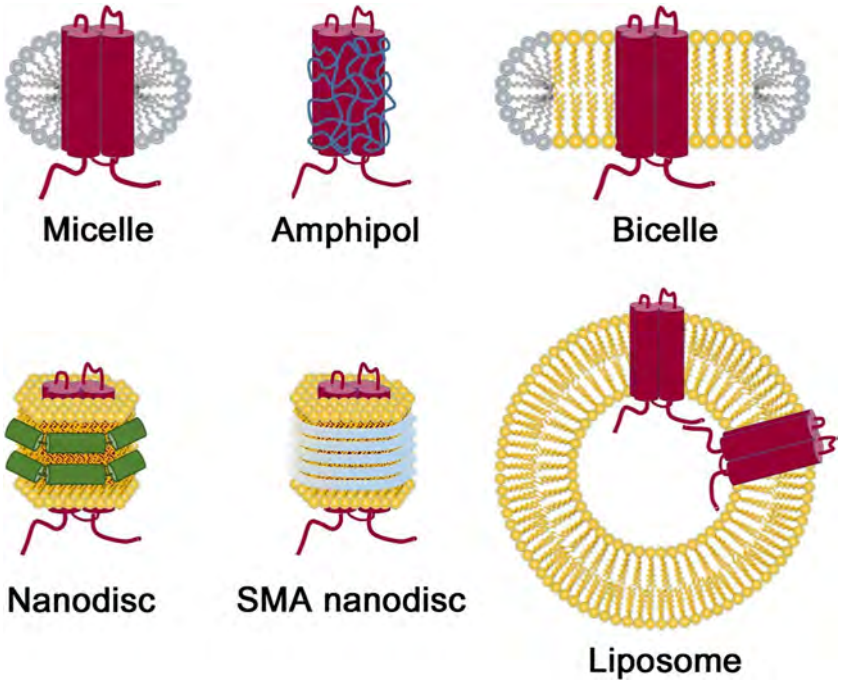
- Cross-linking? Should not be used unless absolutely necessary.

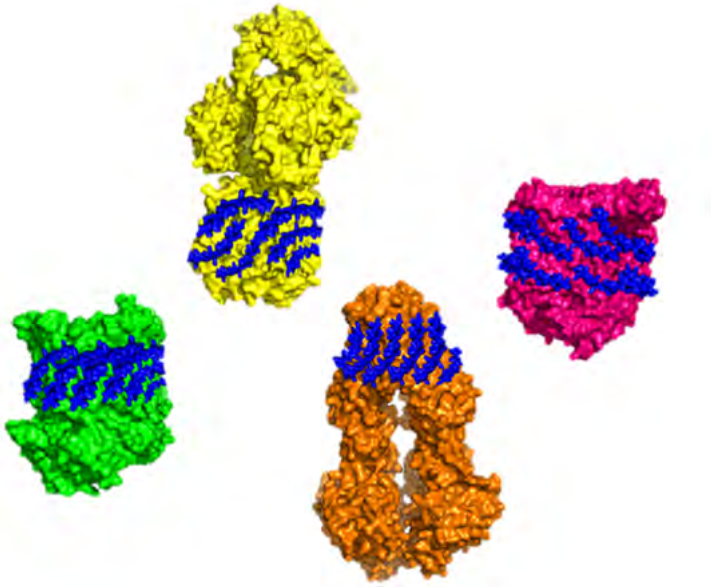
EM Samples: Membrane Proteins

- Use of detergents: starting with the one used for purification, *e.g.*, DDM.
  - Mild non-ionic
  - Amphipols

- Nanoparticles: a membrane-mimetic scaffold that stabilizes proteins in the native lipid-bilayered environment.
  - MSP-nanodiscs (MSP: membrane scaffold protein)
  - SMA nanodiscs (SMA: styrene–maleic acid)
  - Bicelles
  - Peptidiscs

EM Samples: Membrane Proteins





Peptidisc

(Mio & Sato, Biophys Rev, 2018)

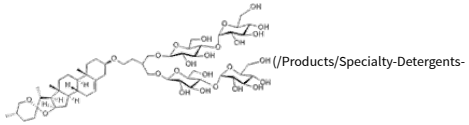
(Carlson et al, eLife, 2018)

EM Samples: Membrane Proteins

DETERGENTS FOR CRYO-EM

The following detergents have been sucessfully used in the Cryo-EM studies of membrane proteins. Want to learn more? Check out our compilations of membrane protein strucutures for 2016(/Landing/2016/Cryo-EM-Update-Sept16), 2017(/Landing/2017/Cryo-EM-Update-Oct17), and 2018(/Landing/2018/Cryo-EM-Update-Oct18).

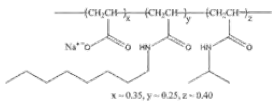
GDN101 - GDN(/PRODUCTS/SPECIALTY-DETERGENTS-PRODUCTS/COMPLEX/GDN101)



(/Products/Specialty-Detergents-Products/COMPLEX/GDN101)

Digitonin is commonly used for Cryo-EM, but there are many drawbacks including batch-to-batch variability and solubility. GDN has been shown to be an effective drop-in substitute for Digitonin which is being used in a number of recent structures.

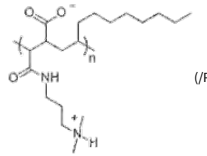
A835 - AMPHIPOL A8-35(/PRODUCTS/SPECIALTY-DETERGENTS-PRODUCTS/AMPHIPOL/A835)



(/Products/Specialty-Detergents-Products/AMPHIPOL/A835)

First described in 1996 by Jean-Luc Popot, amphipols are a class of polymers that can stabilize membrane proteins in a detergent-free, aqueous solution. To date, there have been over 20 Cryo-EM structures of membrane proteins determined using Amphipol A8-35.

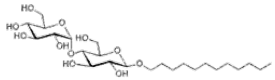
P5008 - AMPHIPOL PMAL-C8(/PRODUCTS/SPECIALTY-DETERGENTS-PRODUCTS/AMPHIPOL/P5008)



(/Products/Specialty-Detergents-Products/AMPHIPOL/P5008)

In recent years, PMAL-C8 has been gaining traction for use in Cryo-EM(/Landing/2018/PMAL-July18) with a number of unique structures published. PMAL amphipols are zwitterionic, and contain repeating units of a carboxyl, ammoniumamidate, and alkyl chain.

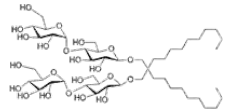
D310 - DDM(/PRODUCTS/DETERGENTS/MALTOSIDES/D310)



(/Products/Detergents/MALTOSIDES/D310)

The most commonly used detergent in membrane protein crystallization, Dodecyl Maltoside (DDM), has also been used in the Cryo-EM structures of a number proteins. DDM is also often used as a mixture with Cholesteryl Hemisuccinate (CHS)(/Products/Detergents/MALTOSIDES/10-1-DDM-CHS-Pre-Made-Solution).

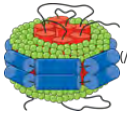
NG310 - LMNG(/PRODUCTS/DETERGENTS/NG-CLASS/NG310)



(/Products/Detergents/NG-CLASS/NG310)

Due to its very low CMC, the concentration of LMNG in the buffer can often be reduced to low concentrations, reducing the amount of free detergent micelles, and reducing background. Like DDM, LMNG is often used as a mixture with Cholesteryl Hemisuccinate (CHS) (/Products/Detergents/NG-CLASS/LMNG-CHS-Pre-Made-Solution).

LIPID NANODISCS(/PRODUCTS/LIPIDS/LIPIDS)



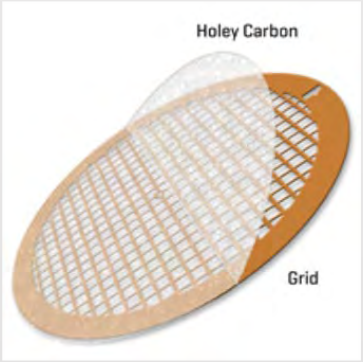
(/Products/Lipids/LIPIDS)

Lipid nanodiscs allow for the reconstitution of a detergent solubilized membrane protein into a lipid environment, and are being increasingly used in Cryo-EM. Anatrace offers a full selection of the lipids commonly used in nanodisc reconstitution.

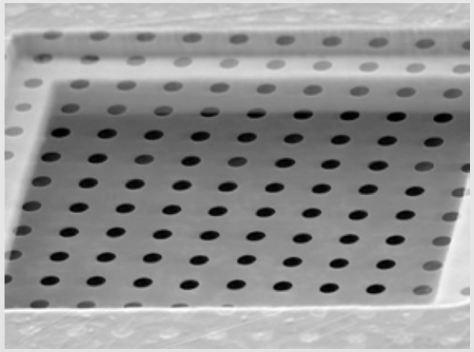
(Anatrace, Inc.)

Preparation of cryo-EM Grids

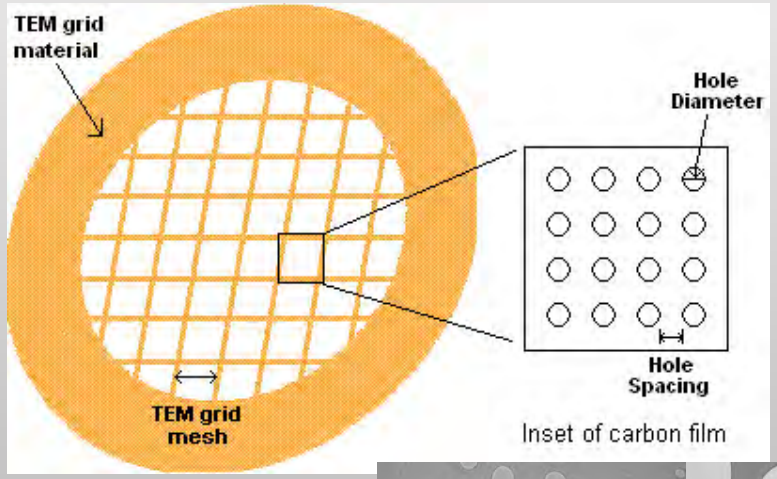
- Supporting grids for cryo-EM
  - Holey carbon grids
    - Quantifoil
    - C-Flat
    - Lacey
  - Gold grids (Quantifoil UltrAuFoil® Holey Gold Films)

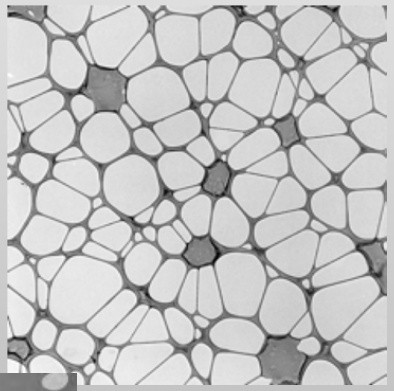


Holey carbon

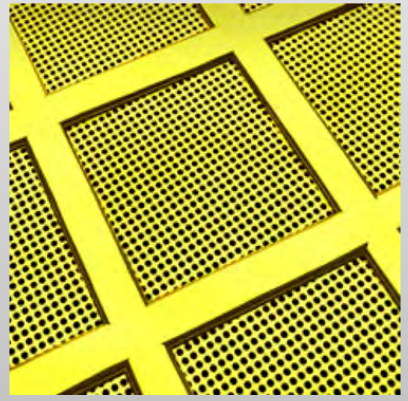


Quantifoil

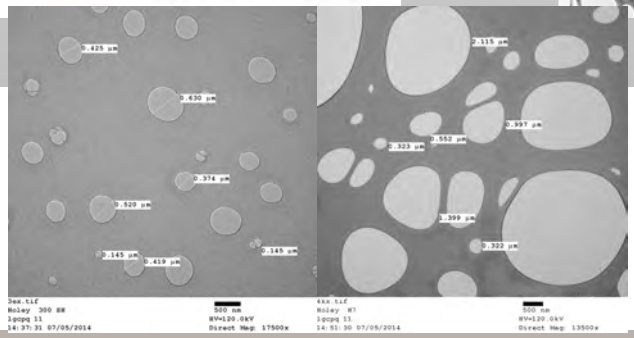




Lacey

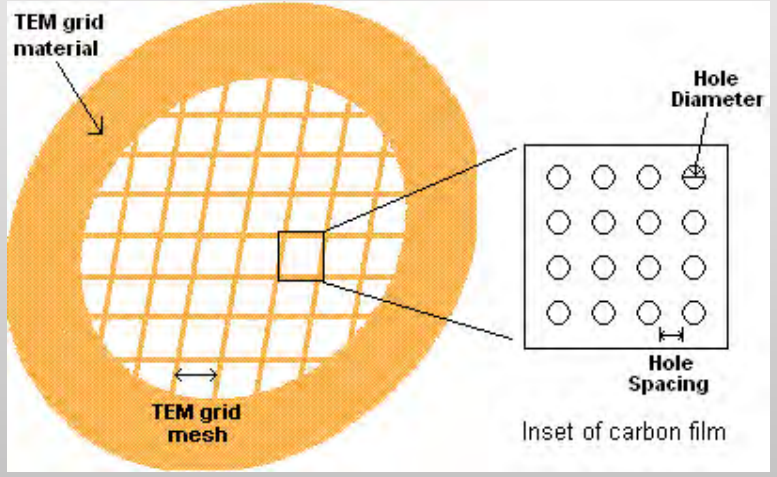


Gold grids (Quantifoil UltrAuFoil)



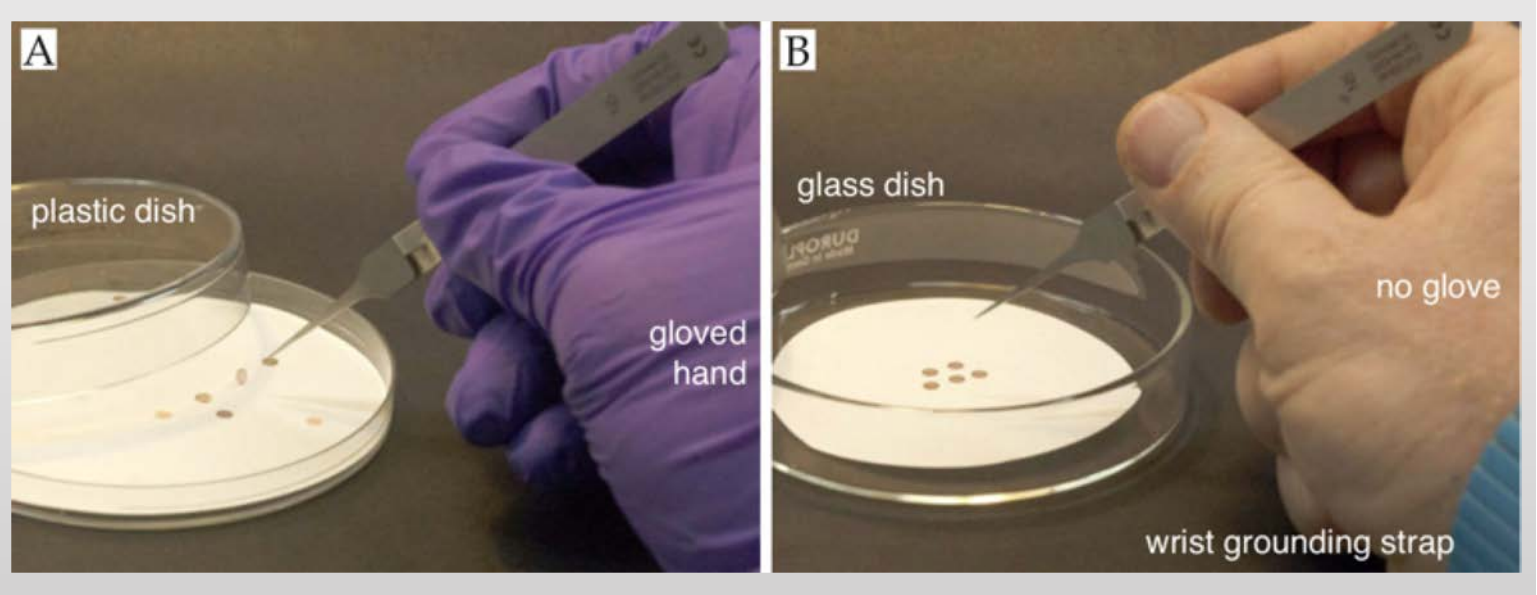
Preparation of cryo-EM Grids

- Supporting grids for cryo-EM
  - Holey carbon grids
    - Quantifoil
    - C-Flat
    - Lacey
  - Gold grids (Quantifoil UltrAuFoil® Holey Gold Films)



Preparation of cryo-EM Grids

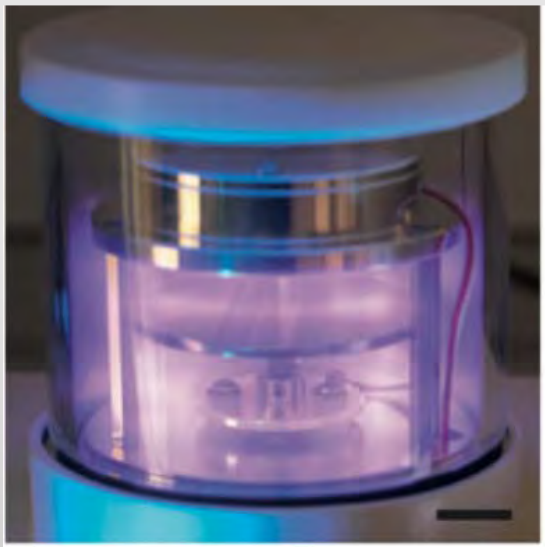
Handling the EM Grids



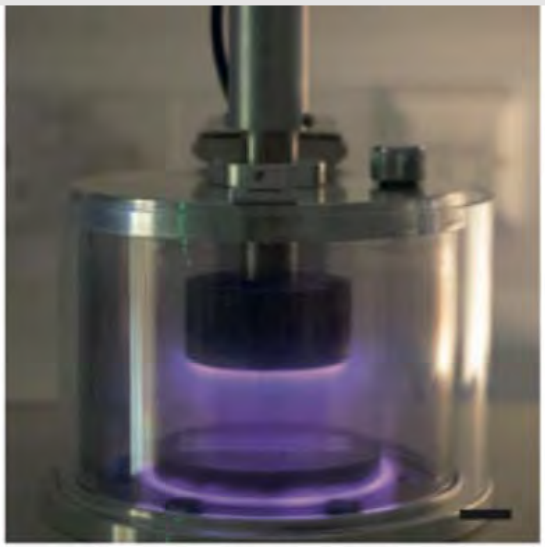
- **Glass dish, instead of plastic petri dish.**
- **Dry environment for long-term storage**

Preparation of cryo-EM Grids

Glow-Discharging



Ted Pella easyGlow (c. 2015)



Edwards S150B (c. 1995)

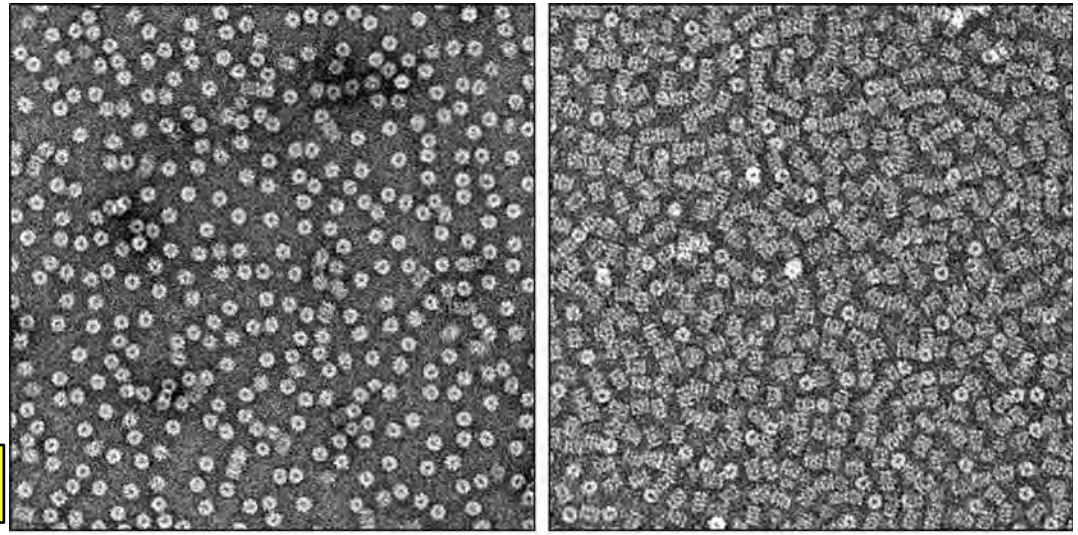
Edwards 12E6 (c. 1962)

- **Ionization-based plasma.**
- **Remove the organic contamination.**
- **Make surface hydrophilic.**

Preparation of cryo-EM Grids

Glow-Discharging

air only



air + pentylamine

(Proteasome, Morris & Fonseca, Acta Cryst D, 2017)

- **Different machines.**
- **Different discharging duration.**
- **Different air conditions.**

Preparation of cryo-EM Grids

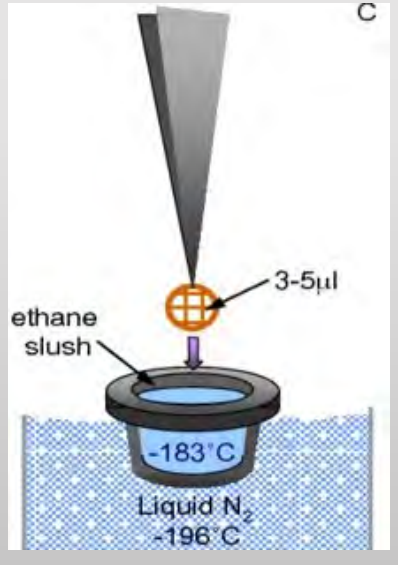
Sample Freezing with a Plunger

EMBL-Heidelberg, 1980s

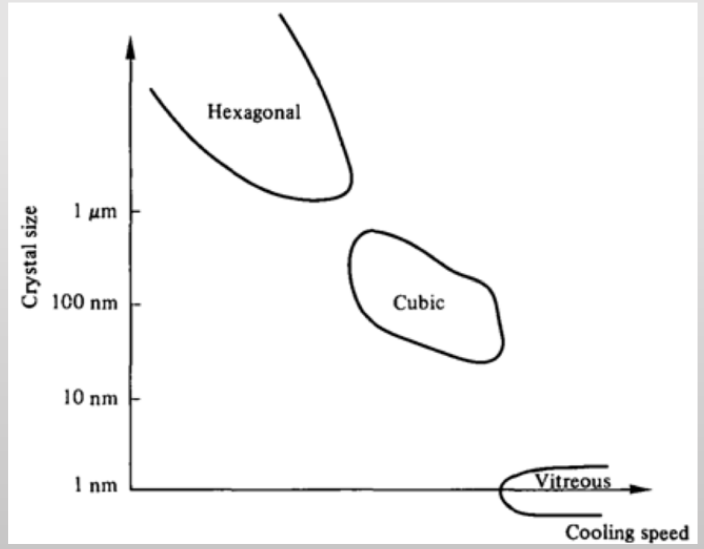
Jacques Dubochet

Preparation of cryo-EM Grids

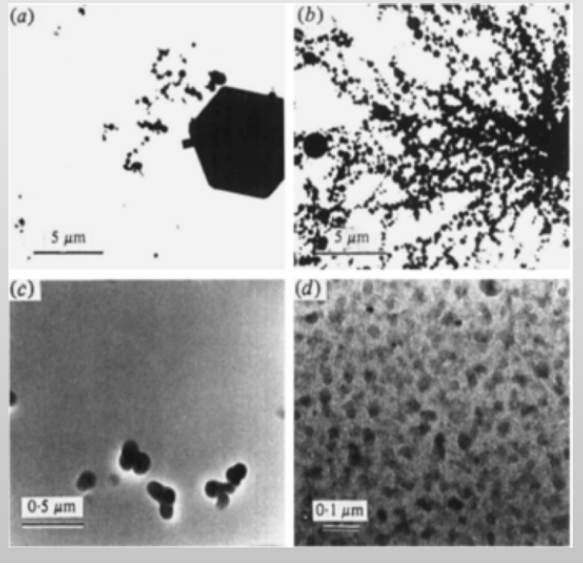
Sample Freezing with a Plunger



Liquid ethane
(Liquid N_2 boils
on contact)



Cooling speed
v.s.
Ice forms

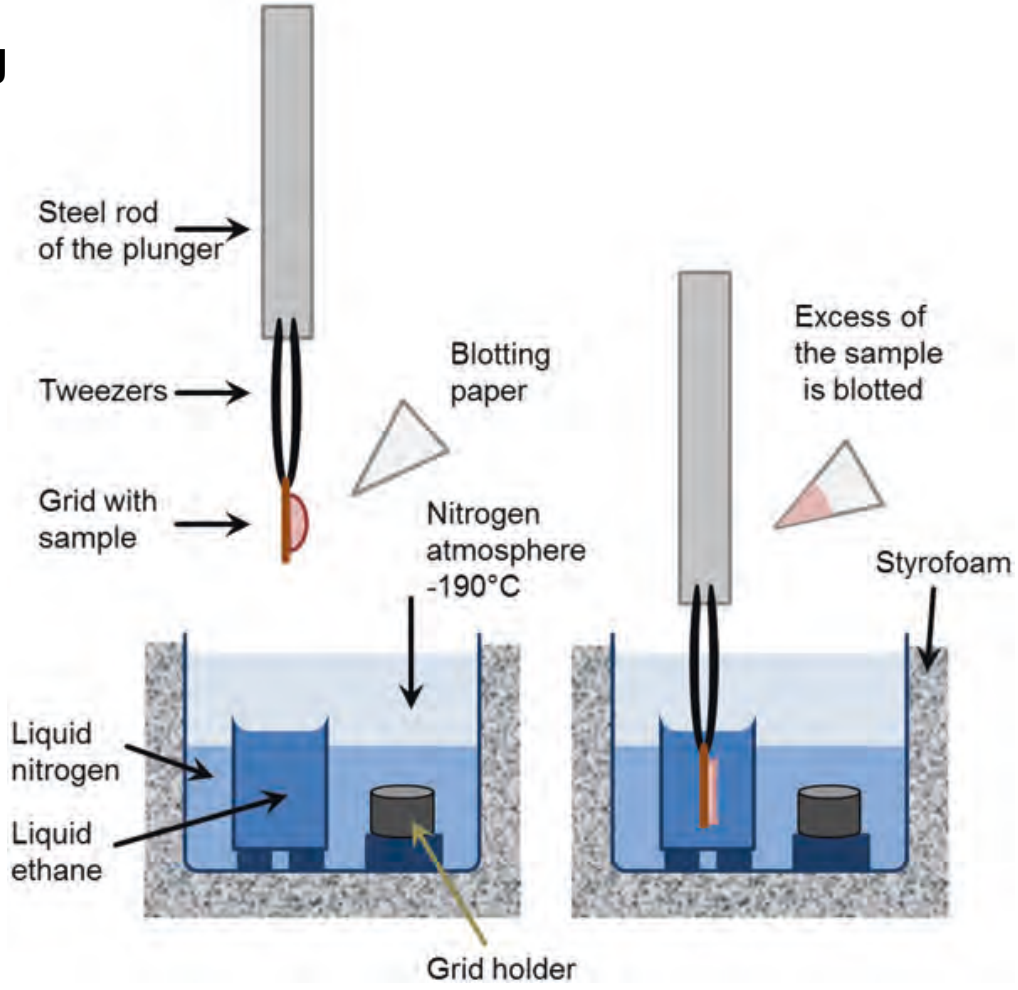


Ice contamination

(Dubochet et al, Q Rev Biophys, 1988)

Preparation of cryo-EM Grids

Sample blotting



Preparation of cryo-EM Grids

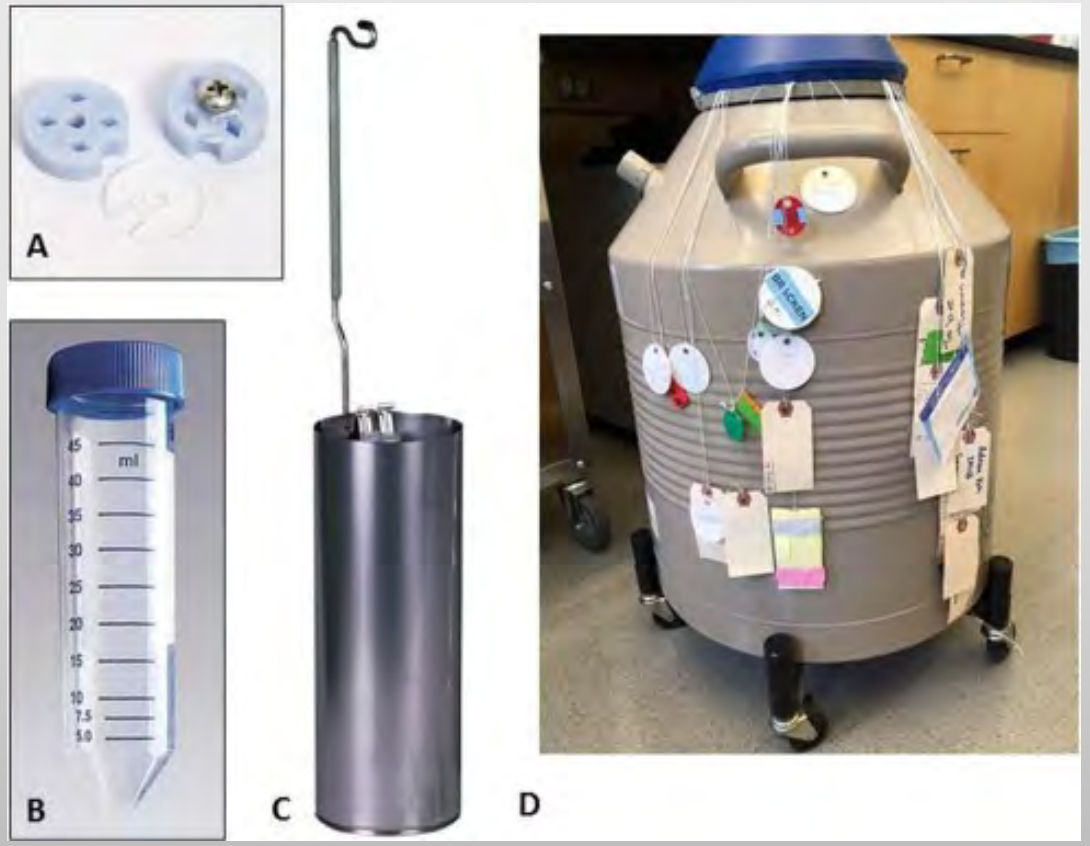
Home-made Plungers

FEI Vitrobot

Gatan CP3

Leica EM GP2

Storage and Transfer



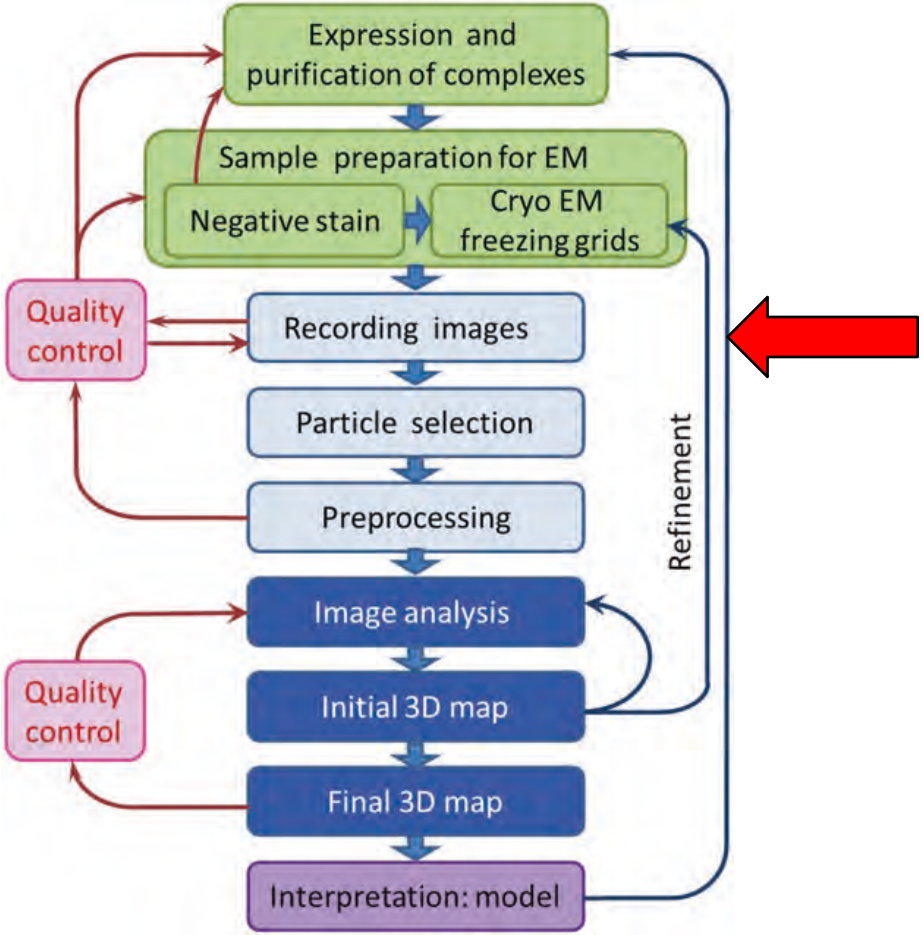
Gatan 626 holder

Toolset with the autoloader

Outline

- Core knowledge (Pt I & else)
- Evaluation of protein biochemistry
- EM sample preparation
  (Single particles & Micro-/2-D crystals)
- **Data collection**
- Image processing and analysis

Cryo-EM Workflow



Screening and Optimization

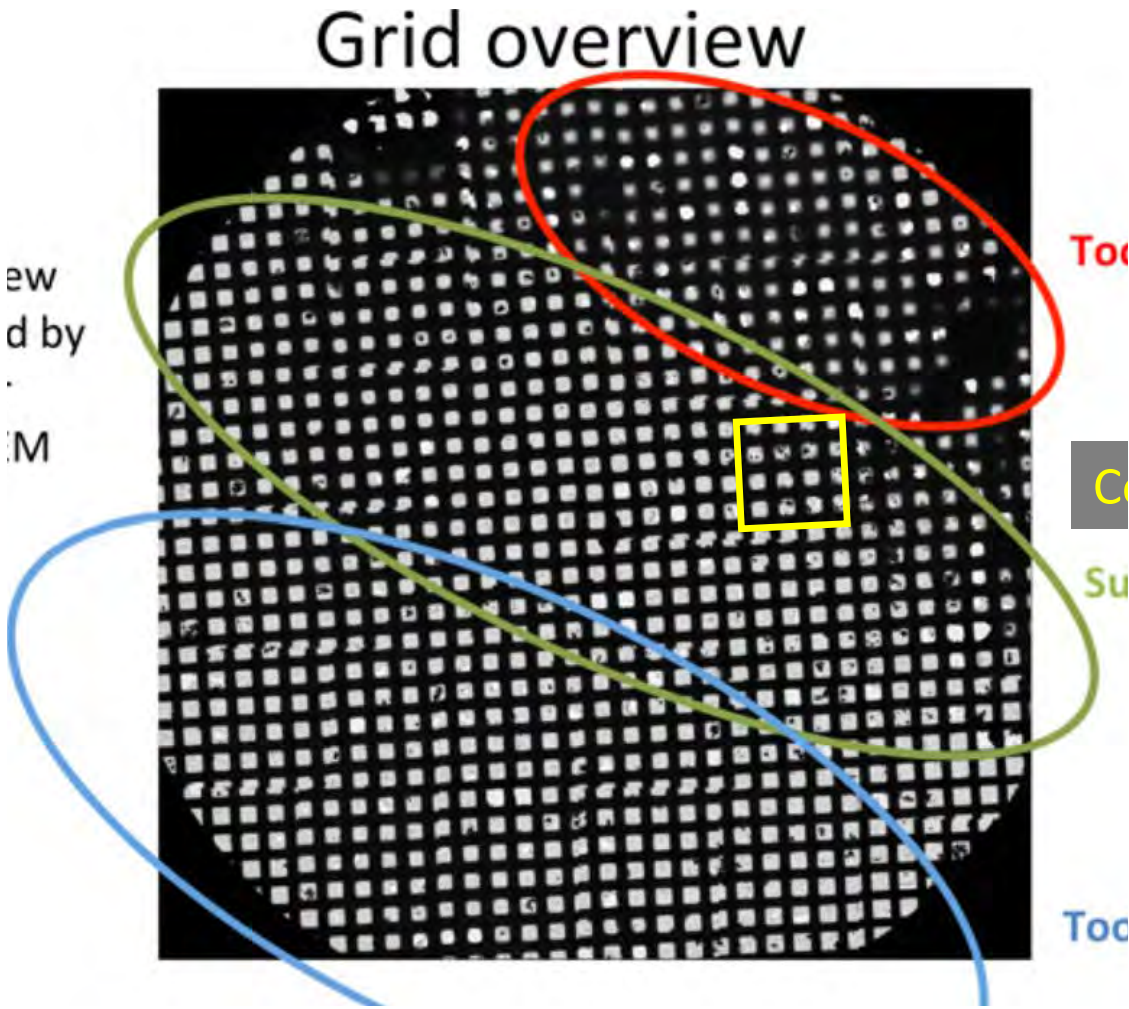
- Very important for data collection
  - Learn to compare different areas (squares and holes) on the grids
  - Good grids, good and fast data collection

- Literature search for similar cases

- Types of EM grids
- Protein concentrations ± additives
- Blotting conditions: time, force, humidity & ethane temperature

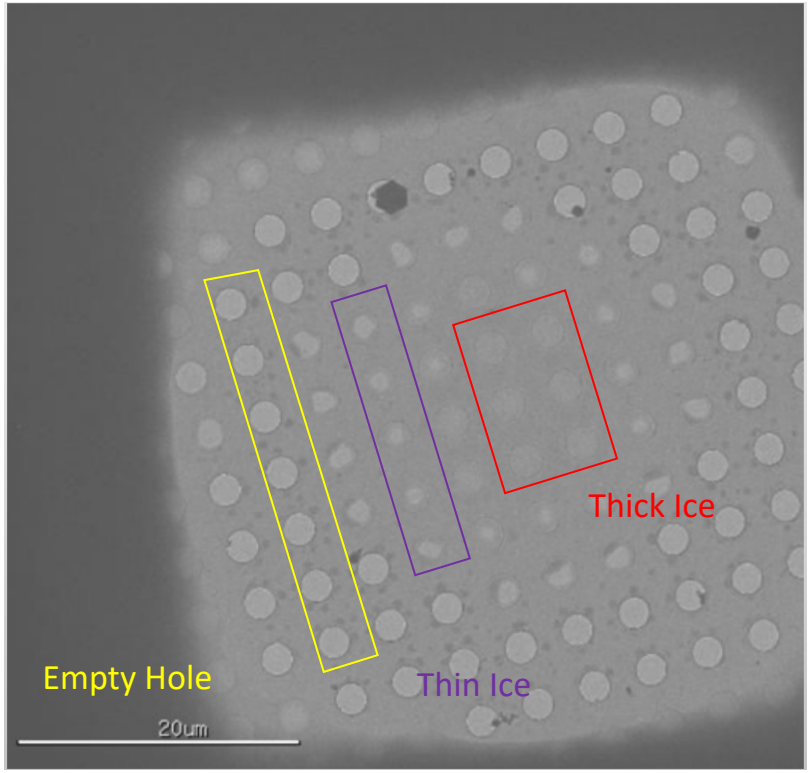
Screening and Optimization

- Most area of a whole grid is likely not ideal.
  - Red: too thick
  - Blue: too thin
  - Green: suitable



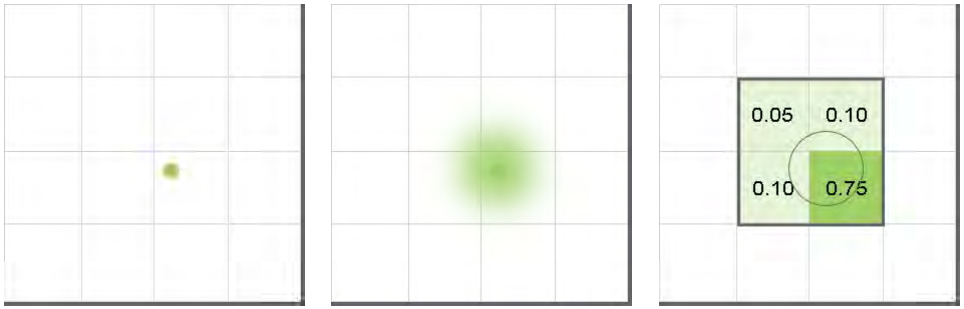
Screening and Optimization

- Too thick
  - Low contrast
  - Low S/N ratios
- Too thin
  - Not thick enough to accommodate particles
  - Protein denaturation
  - More preferred orientation
  - Poor support and large motion during imaging
  - Vulnerable to radiation damage



Imaging: Parameters to Consider

- Magnification / pixel size
  - Theoretical maximal resolution = 2x pixel size
  - Smaller pixel size --> better DQE at high frequency (DQE: detective quantum efficiency)
  - High mag: few images
  - Low mag: 2x pixel size for maximal resolution
  - In general, use "2.5-3x pixel size"



Electron enters detector.

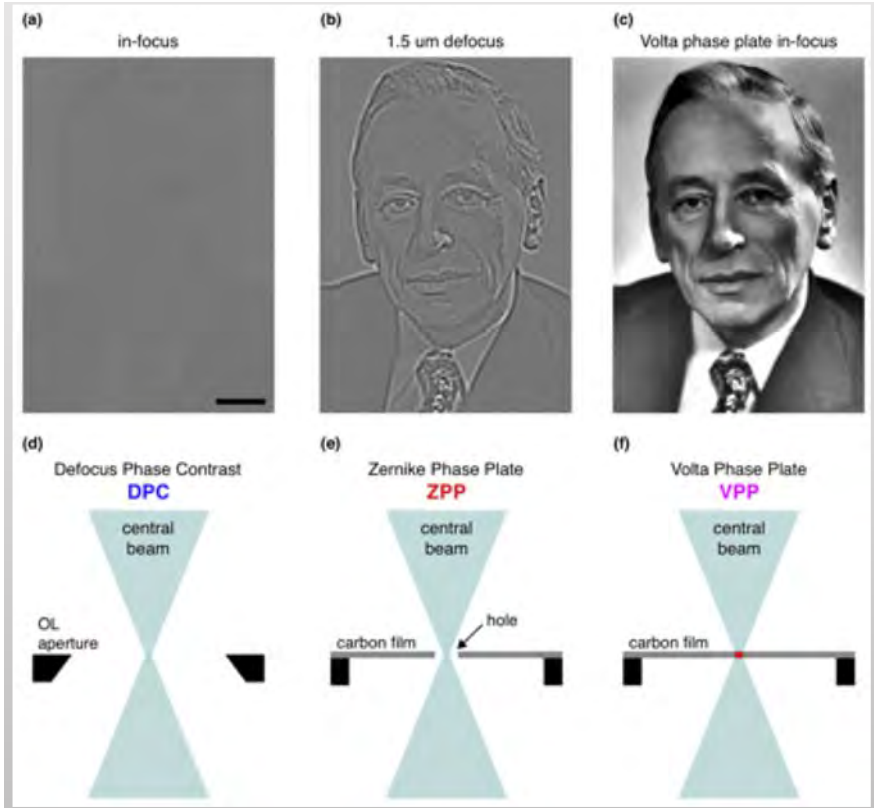
Electron signal is scattered.

Charge collects in each pixel.

(Gatan, Inc.)

Imaging: Parameters to Consider

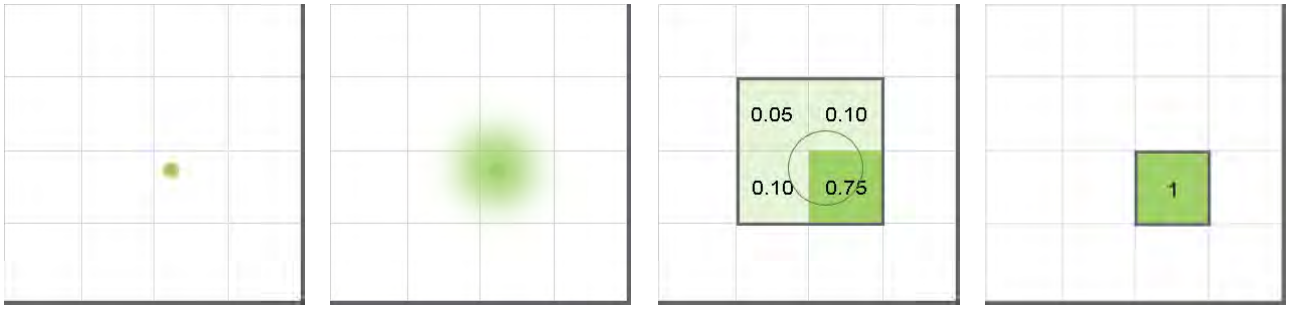
- Phase plate: increasing contrast while "in-focus"



(Danevet & Baumeister, Curr Opin Struct Biol, 2017)

Imaging: Parameters to Consider

- Super-resolution or counting mode



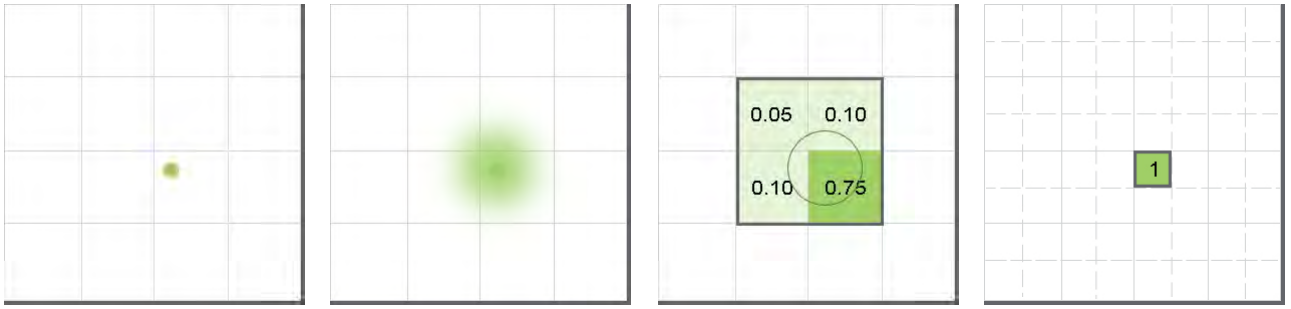
Counting

Electron enters detector.

Electron signal is scattered.

Charge collects in each pixel.

Events reduced to highest charge pixels.



Super-resolution

Electron enters detector.

Electron signal is scattered.

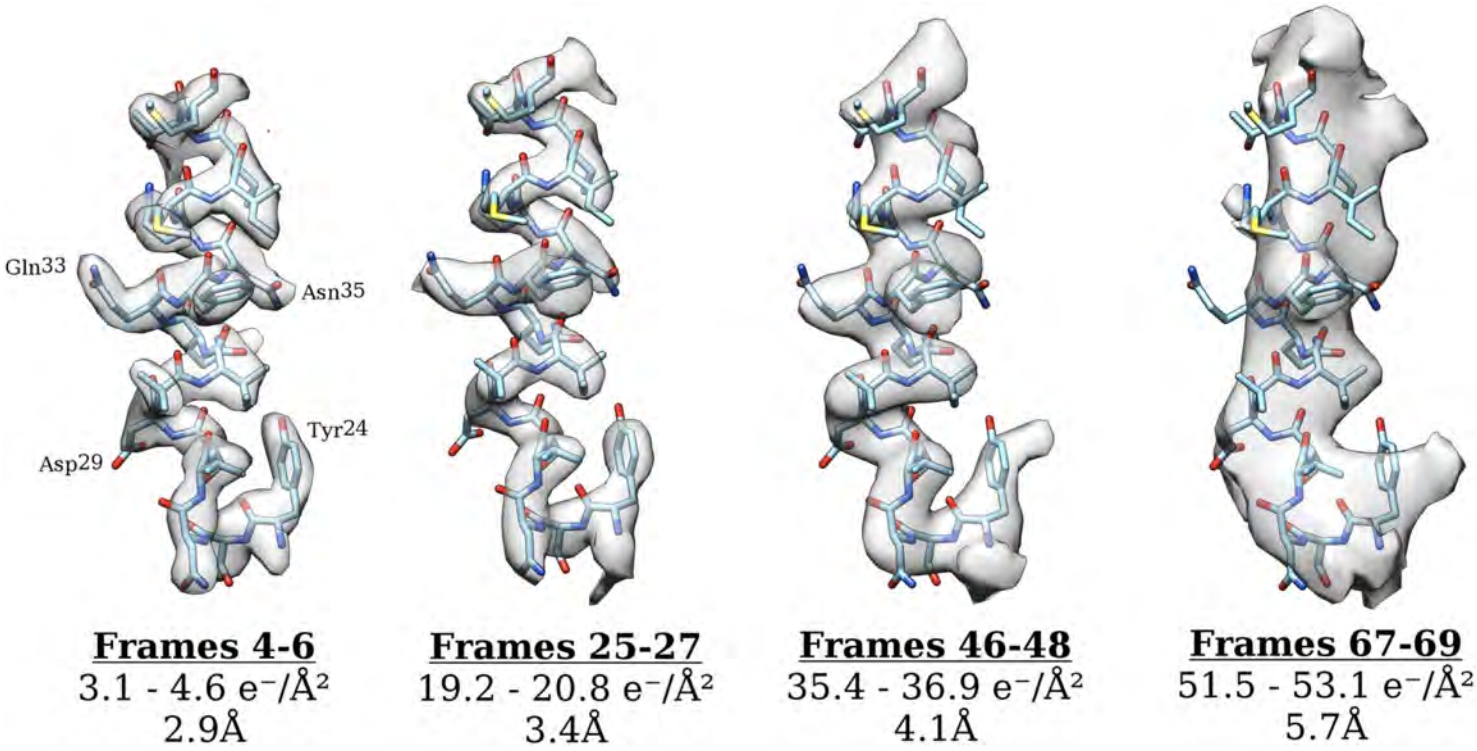
Charge collects in each pixel.

Events localized to sub-pixel accuracy.

(Gatan, Inc.)

Imaging: Parameters to Consider

- Dose rate: "low-dose" mode

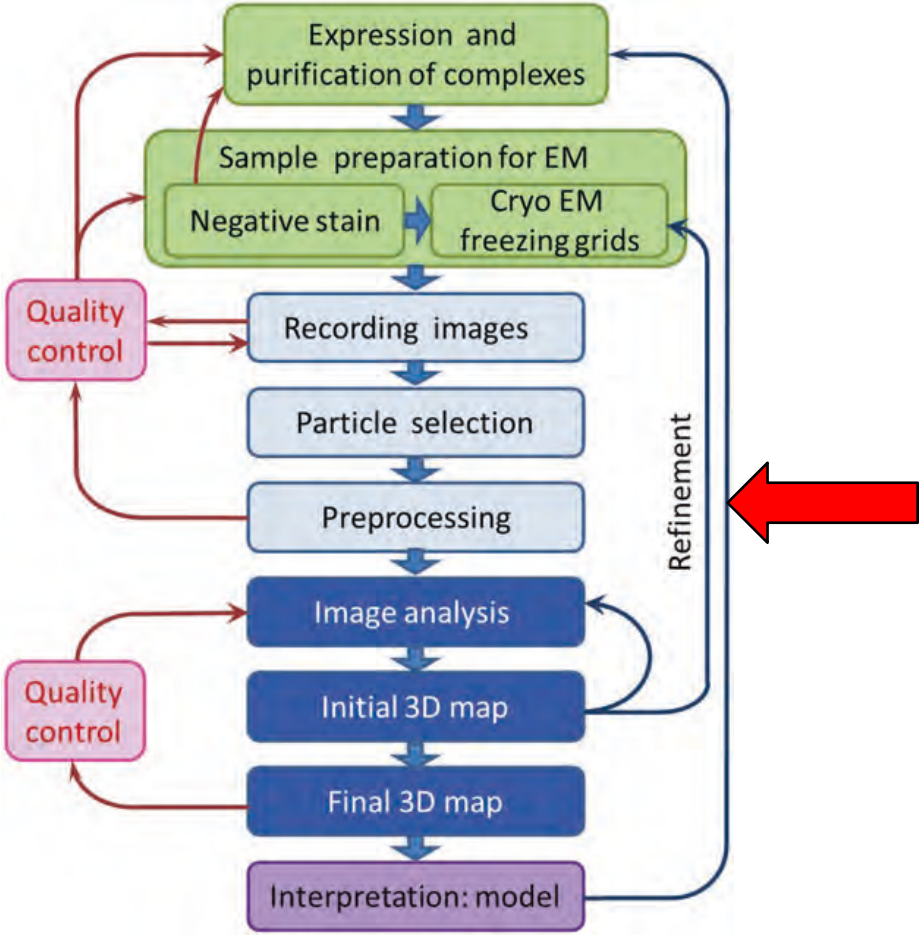


(Grant & Grigorieff, eLife, 2015)

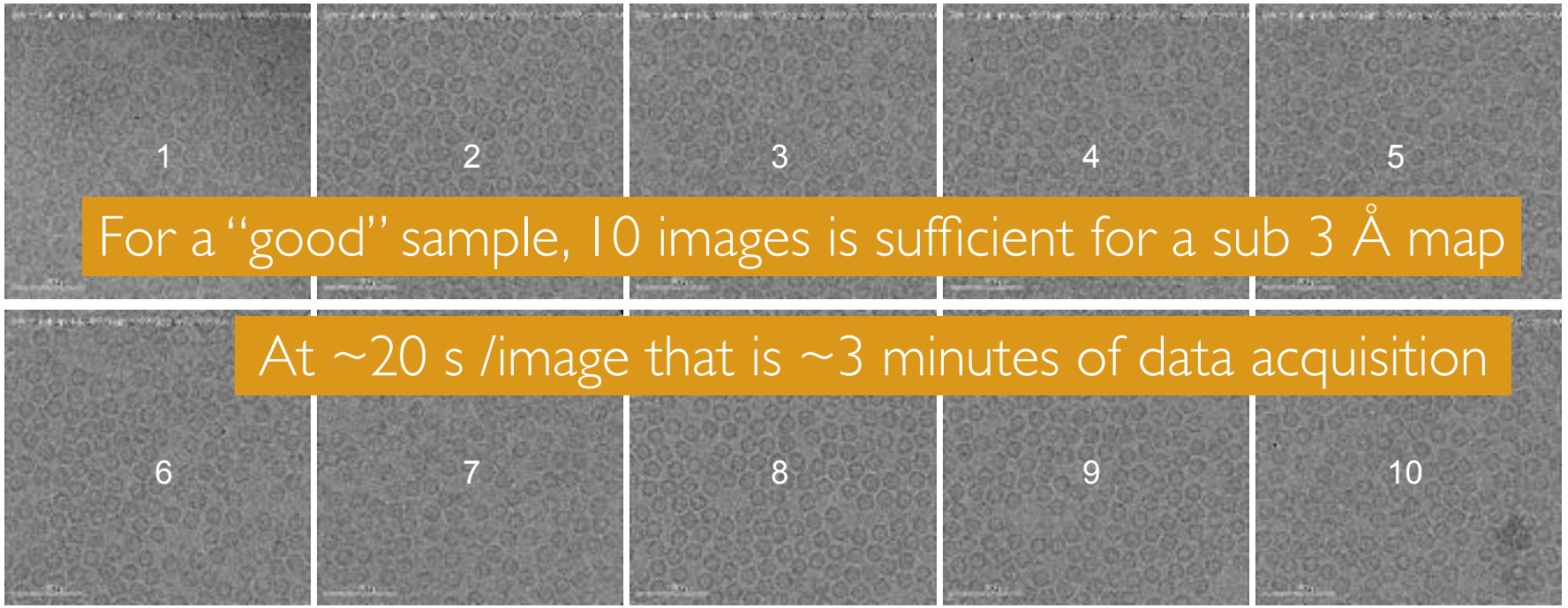
Outline

- Core knowledge (Pt I & else)
- Evaluation of protein biochemistry
- EM sample preparation
  (Single particles & Micro-/2-D crystals)
- Data collection
- **Image processing and analysis**

Cryo-EM Workflow

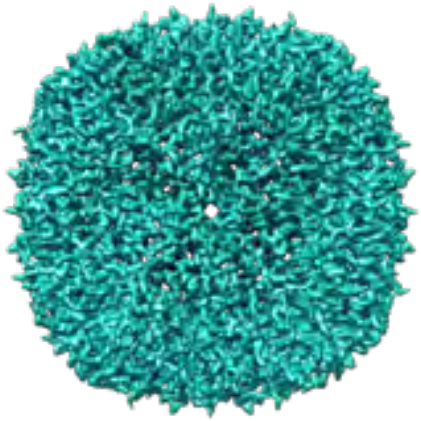


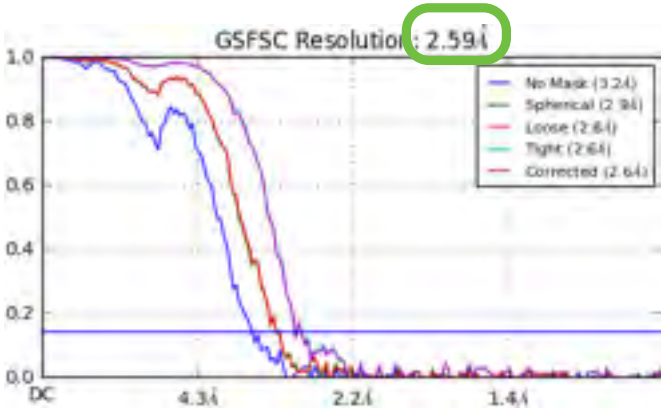
How many images do we need?



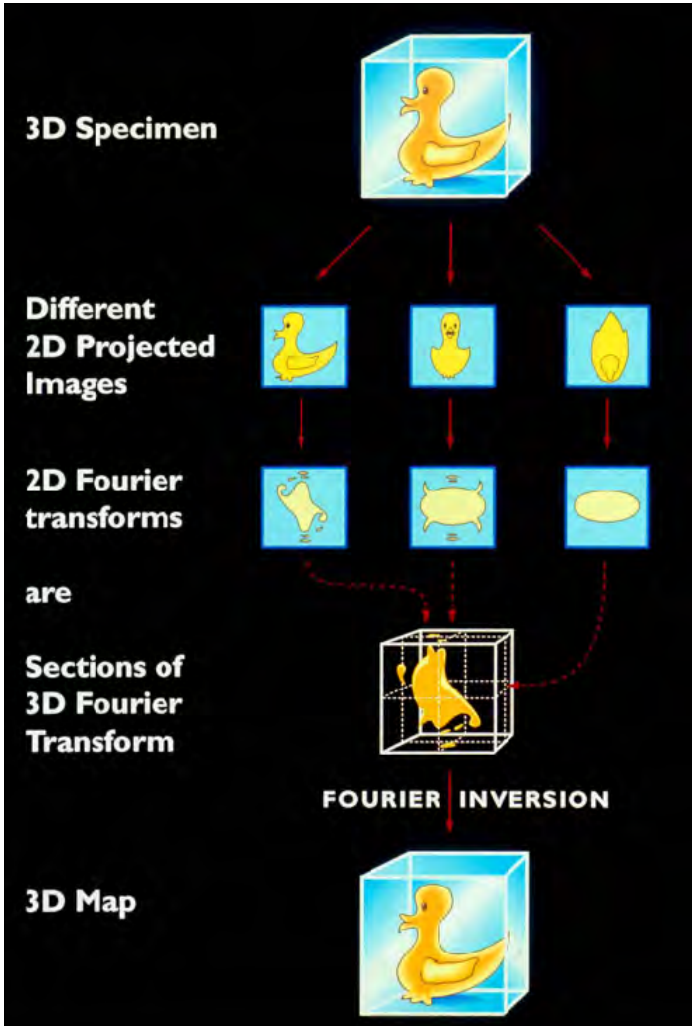
For a "good" sample, 10 images is sufficient for a sub 3 Å map

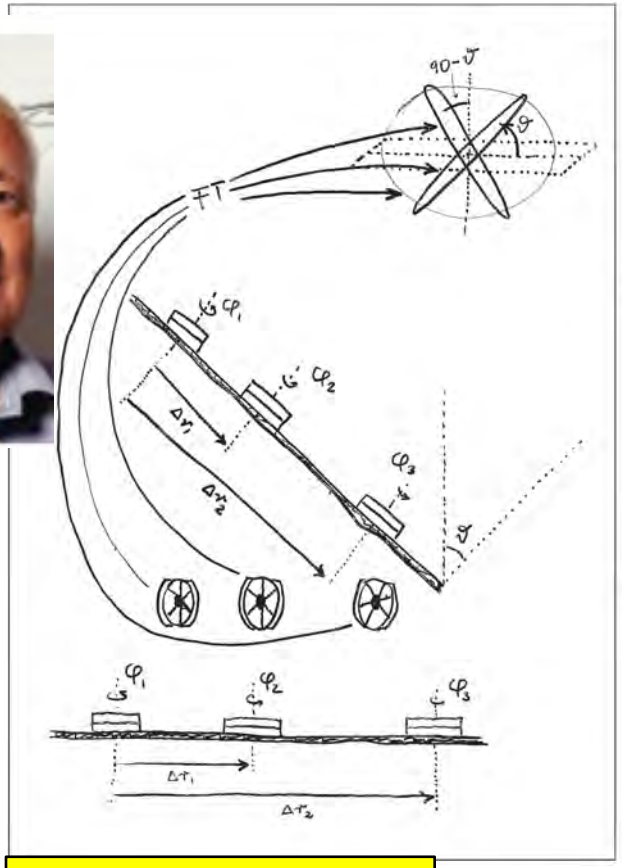
At ~20 s /image that is ~3 minutes of data acquisition

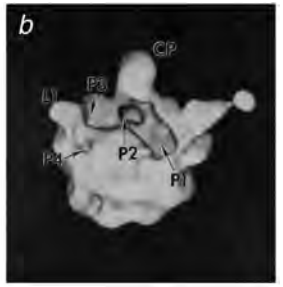


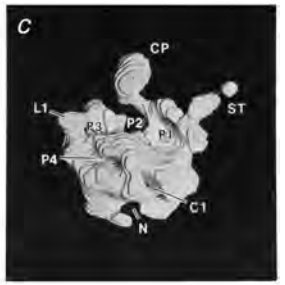


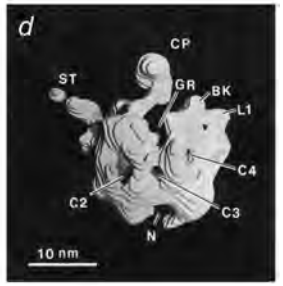
2-D images to 3-D Reconstruction









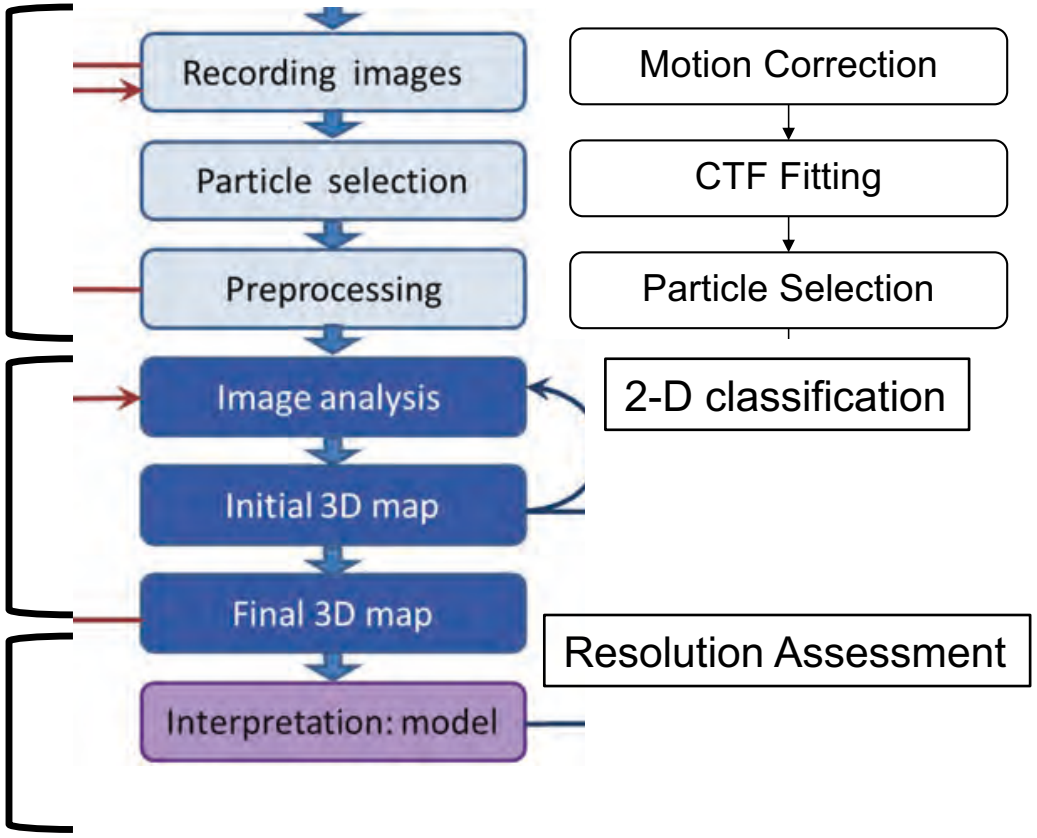


1979 hand-drawing

(Frank, Q Rev Biophysc, 2009)

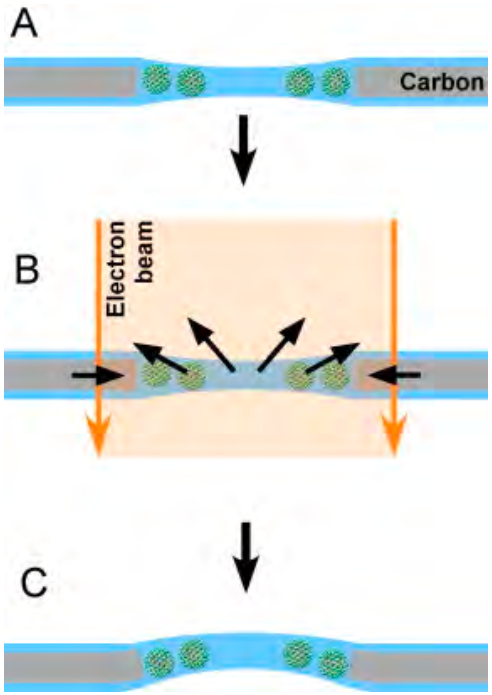
Software & Workflow

- cryoSPARC (GPU)
- Relion (GPU)
- EMAN/EMAN2
- Frealign/cisTEM
- Xmipps/Scipion
- Spider
- IMAGIC
- MRC/2dx (2-D crystals/MicroED)
- …

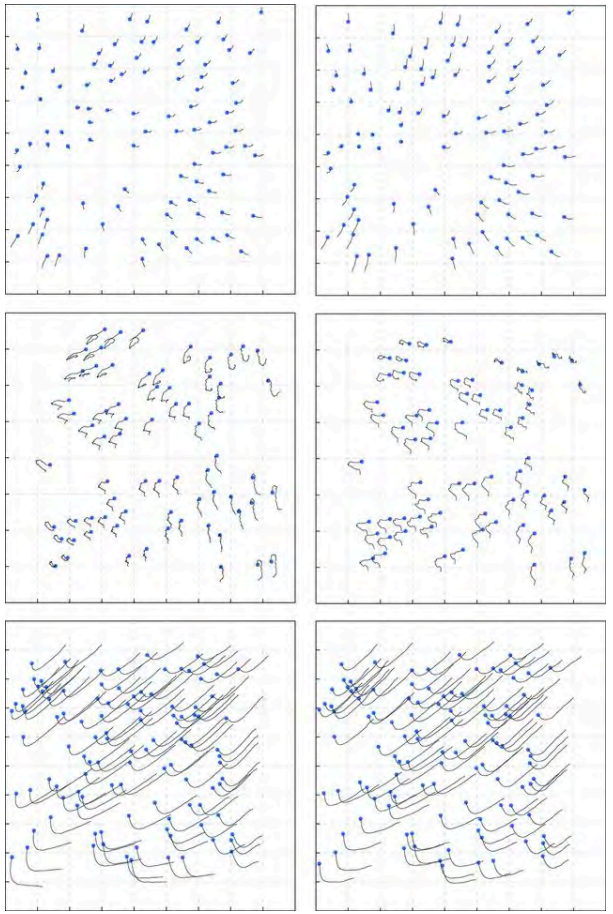


Motion Correction

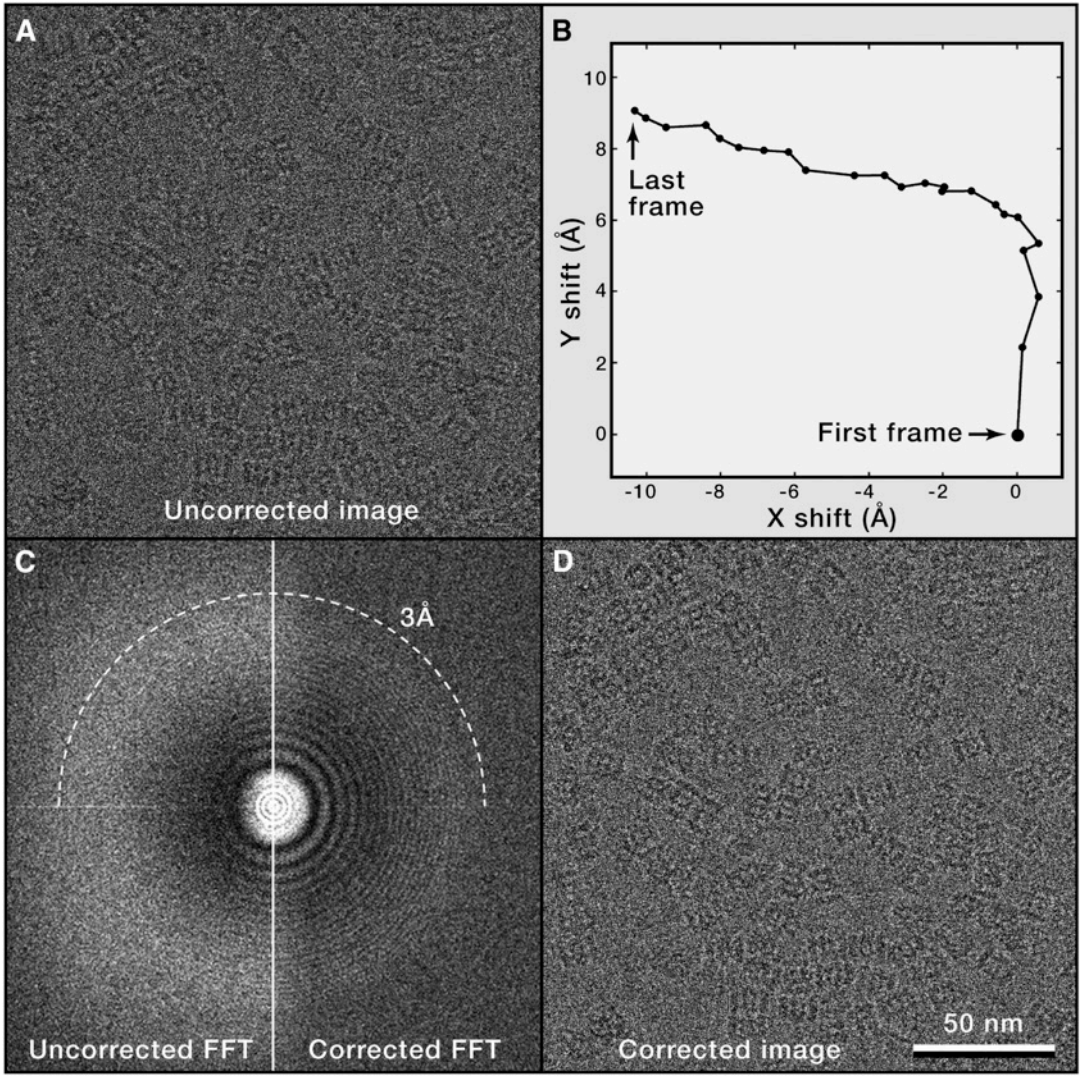
- Stage drift
- Beam-induced sample motion



(Brilot et al, J Struct Biol, 2012)



(Zivanov et al, IUCrJ, 2019)



(Cheng et al, Cell, 2015)

CTF Assessment (Power Spectra)

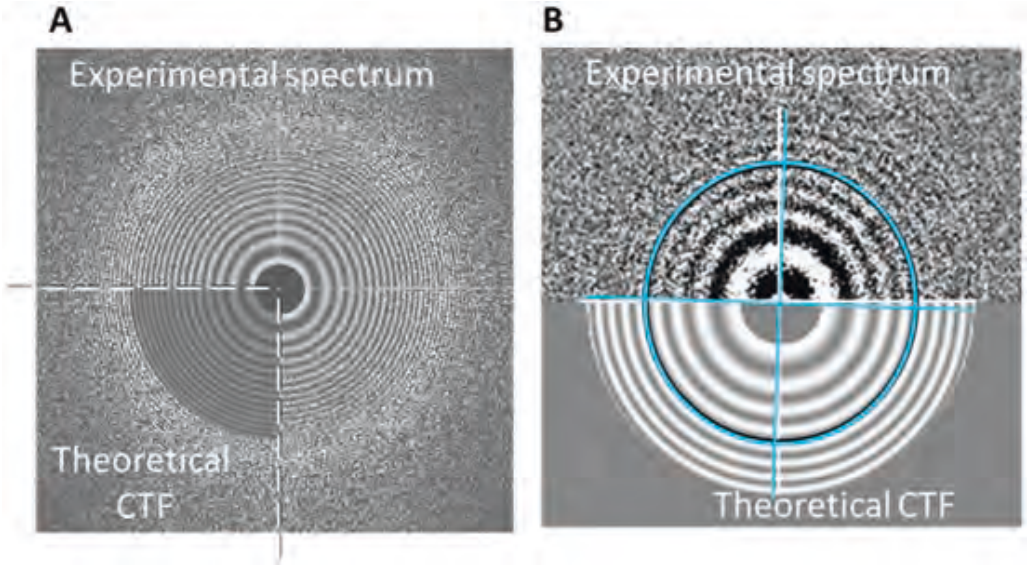
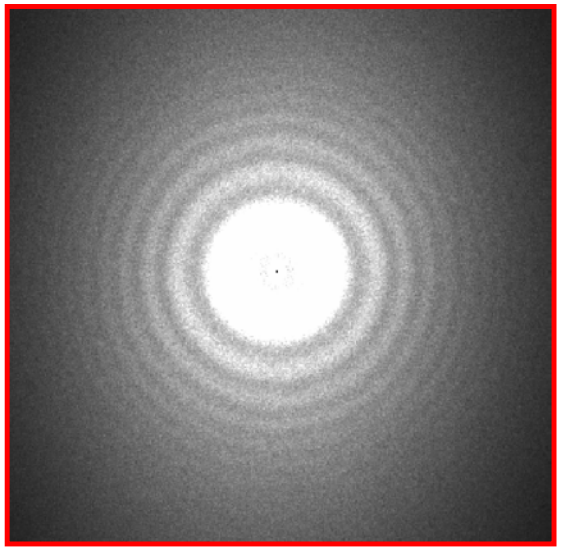


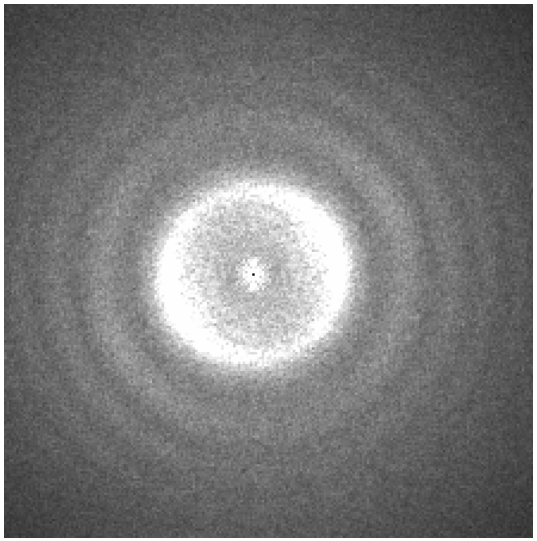
Fig. 8 Assessment of CTF parameters. (**a**) Comparison of theoretically calculated CTF (*left bottom quadrant*) with CTF seen in experimental spectrum. For an accurate CTF determination the Thon rings from both image parts should match accurately. (**b**) Identification of axes of astigmatism which are superimposed over Thon rings of an actual observed power spectrum and compared with the theoretical spectrum. The spectrum of a micrograph shown here indicates that there is a small astigmatism, ~2%, and the axes of ellipse are slightly tilted, shown in *light blue*

(Costa et al, Meth Mol Biol, 2017)

CTF Assessment (Power Spectra)

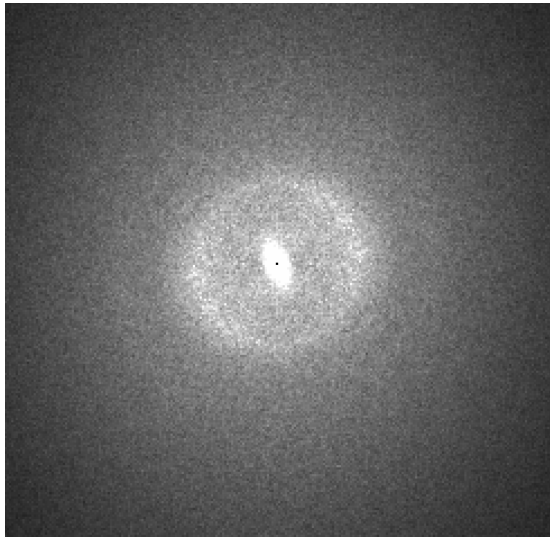
Good:
- Isotropic
- Thon rings at high resolution

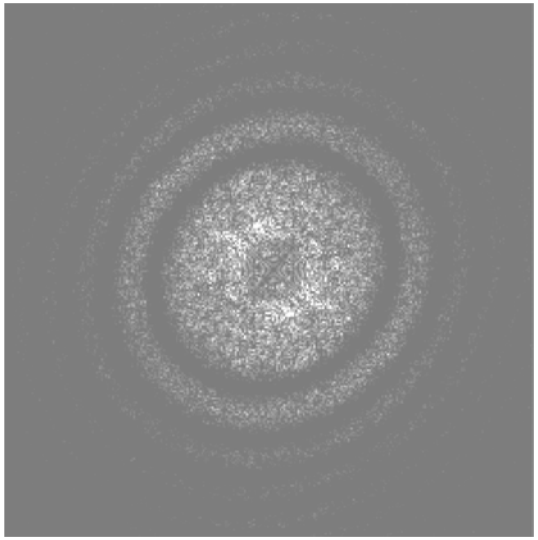




Bad:
Missing Thon rings at certain direction due to drift (*can be corrected if movies are recorded*)

Bad:
Thon rings only at low resolution

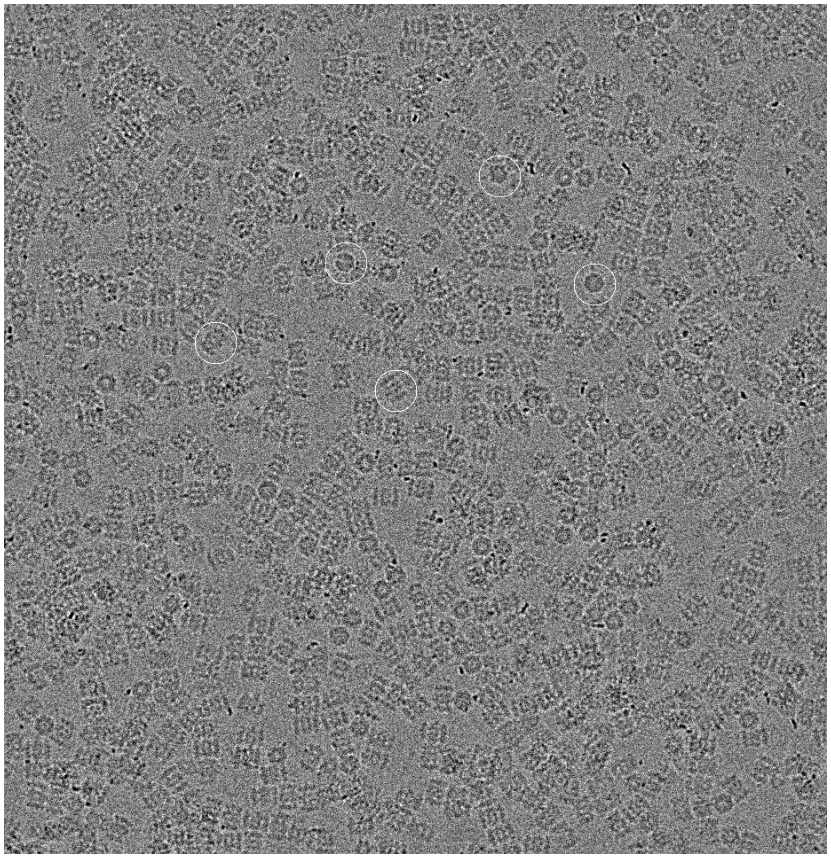




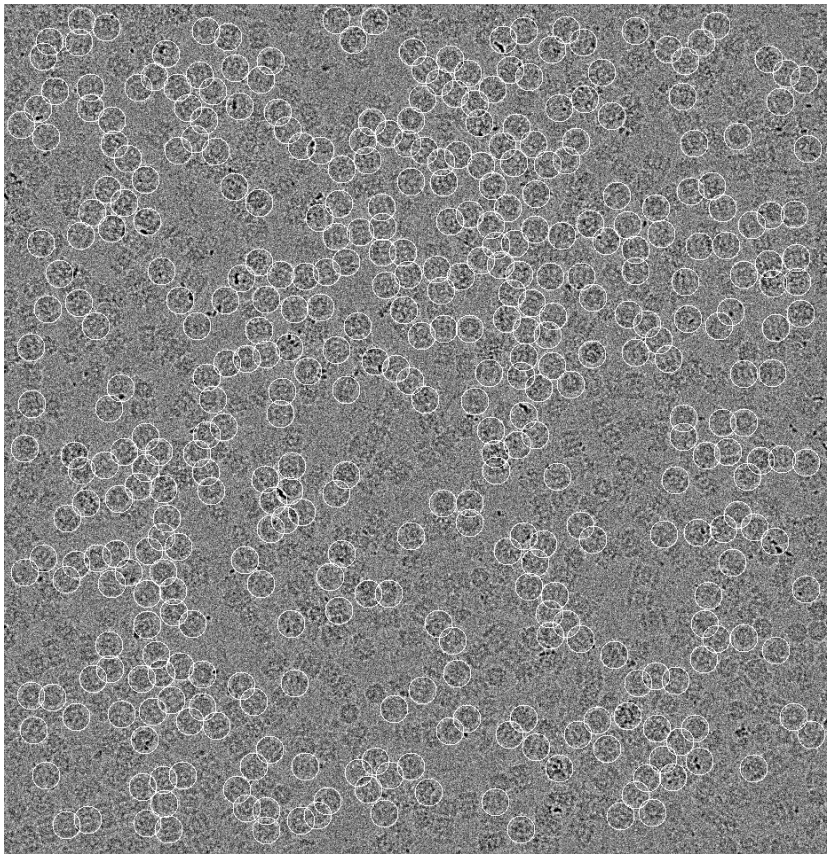
Bad:
Elliptic Thon rings due to astigmatism (*can be useful if properly processed*)

Particle Selection & 2-D Classes

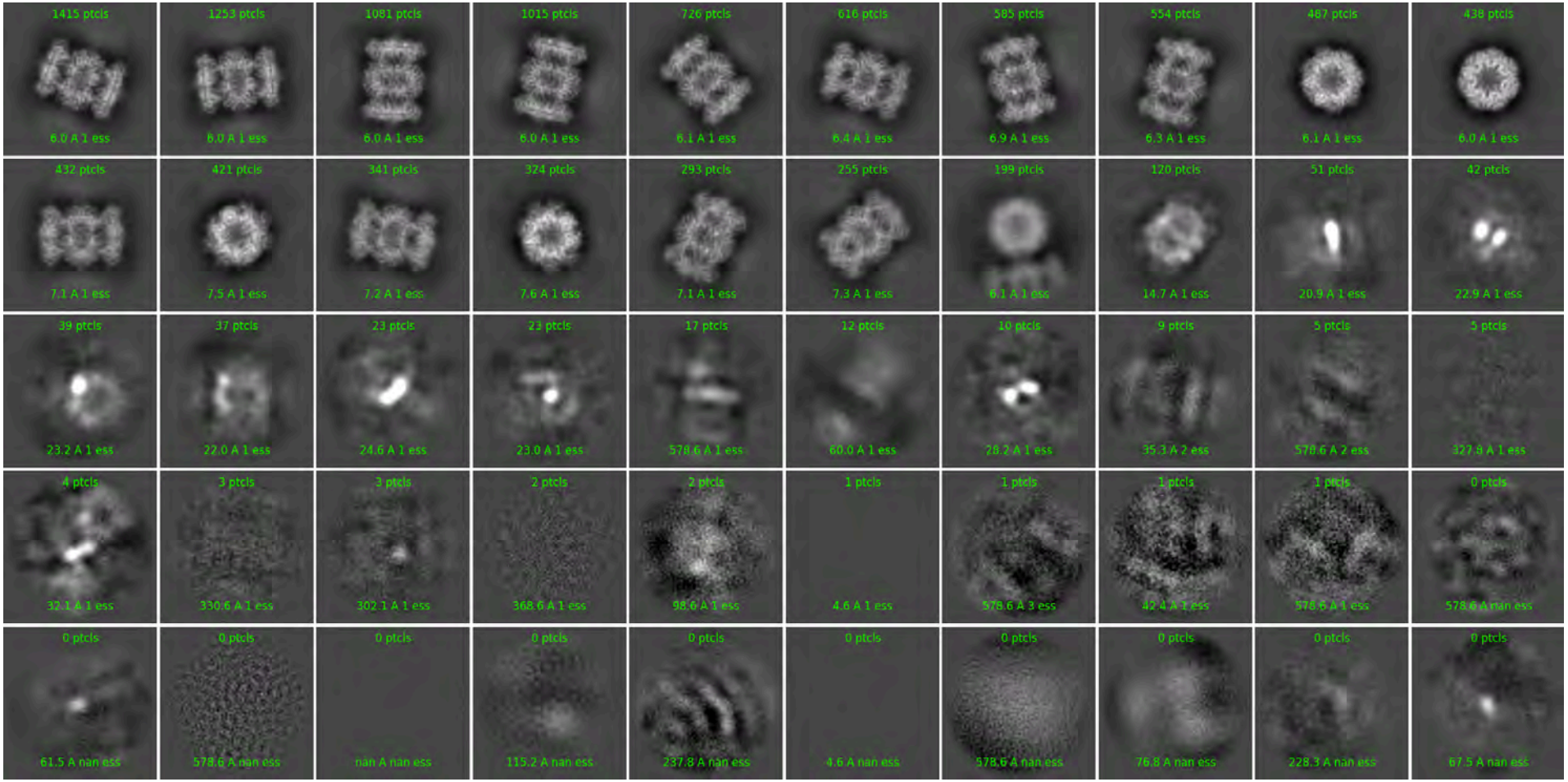
Manual



Automated
(template/deep-learning)

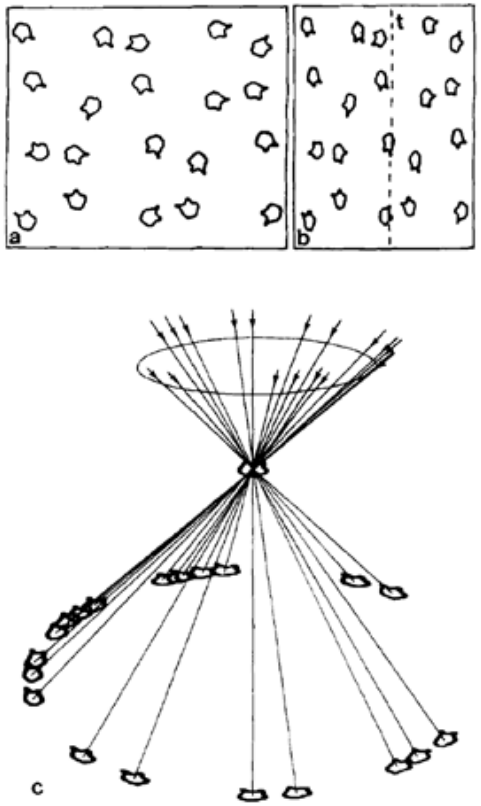


Particle Selection & 2-D Classes



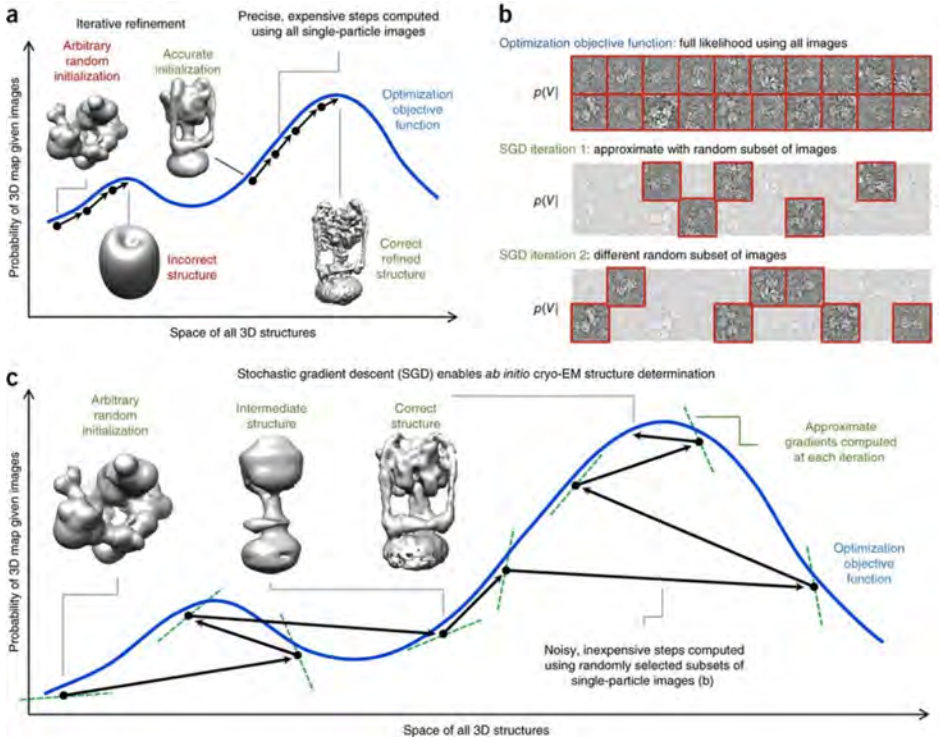
Initial Model & 3-D Classes

Random Conical Tilt



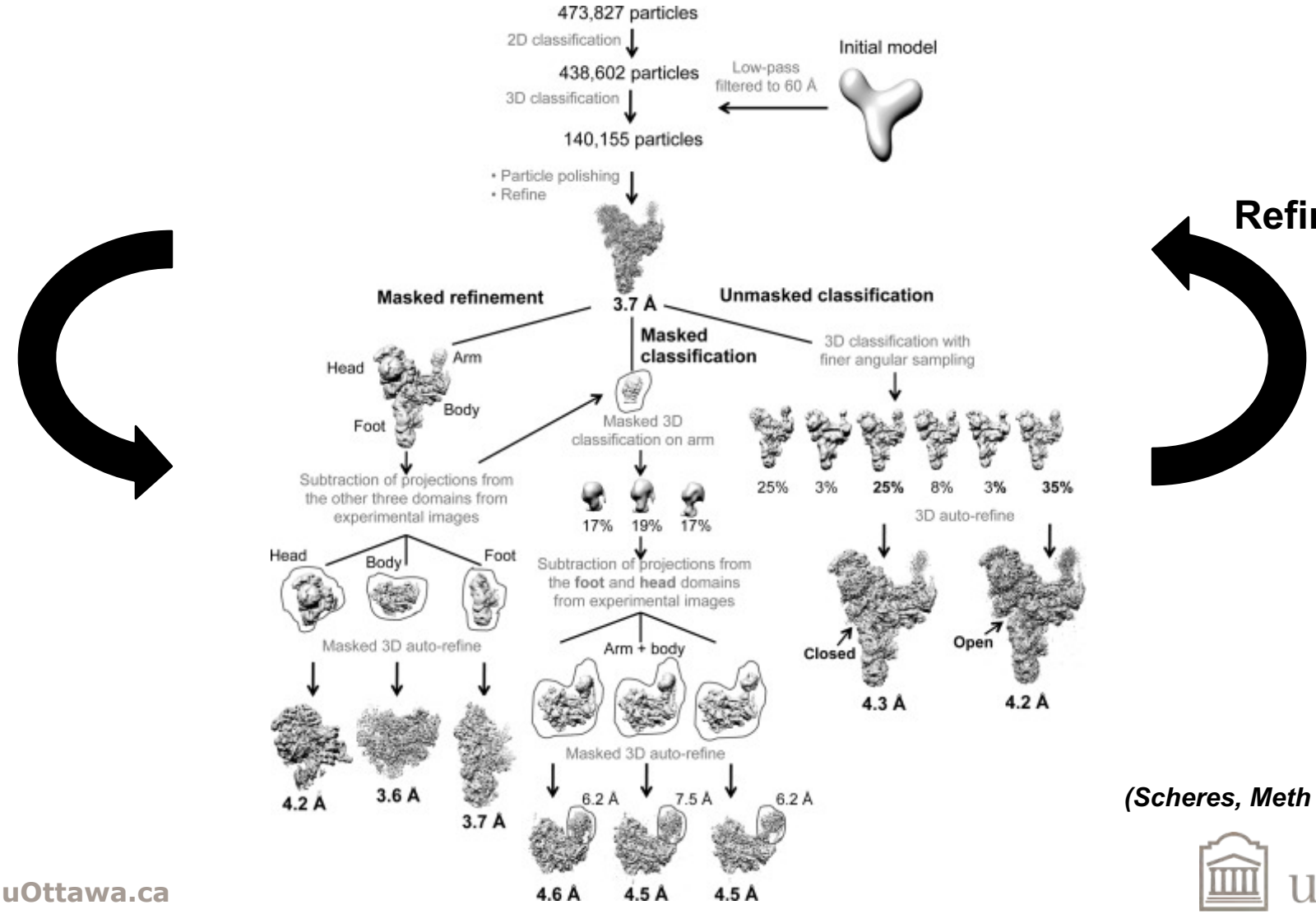
(Radermacher et al, J Microsc, 1986)

Stochastic Gradient Descent



(Punjani et al, Nat Methods, 2017)

Initial Model & 3-D Classes



Refinement

(Scheres, Meth Enzymol, 2016)

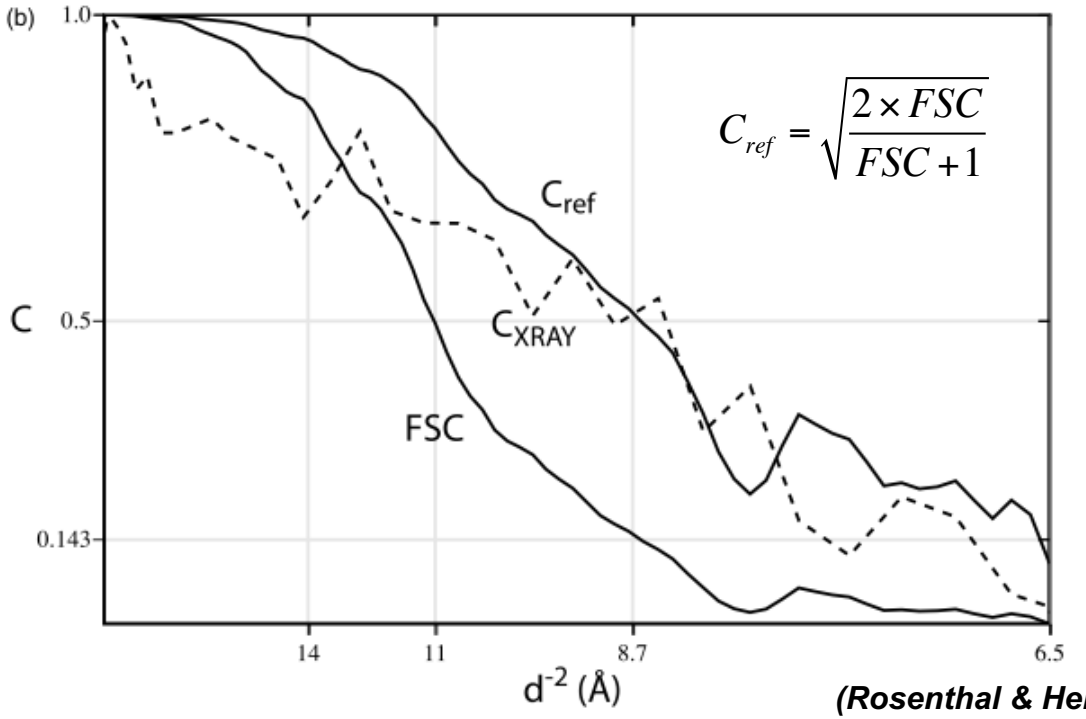
Resolution Assessment

- Split particles into even and odd halves, reconstruct and compare models
- Early literature (<2005): differential phase residue
- Now: Fourier Shell Correlation (FSC)
  - Easy to compute. Invariant to filtering/sharpening level
  - Threshold? (0.5, 0.33, 0.14, 3σ)
  - Masking?
  - Split data? When?

$$FSC = \frac{\sum (F_1 \cdot F_2^*)}{\sqrt{(\sum |F_1|^2)(\sum |F_2|^2)}}$$

$$FSC = \frac{SNR}{SNR + 1}$$

Resolution Assessment



(Rosenthal & Henderson, JMB, 2003)

Full dataset map	FSC of half-data set maps from truly independent reconstructions	0.143
Full dataset map	FSC of full-data set map vs atomic model	0.5
Full dataset map	FSC of half-data set maps from semi-independent reconstructions	0.5

Part III

CHALLENGES AND OPPORTUNITIES

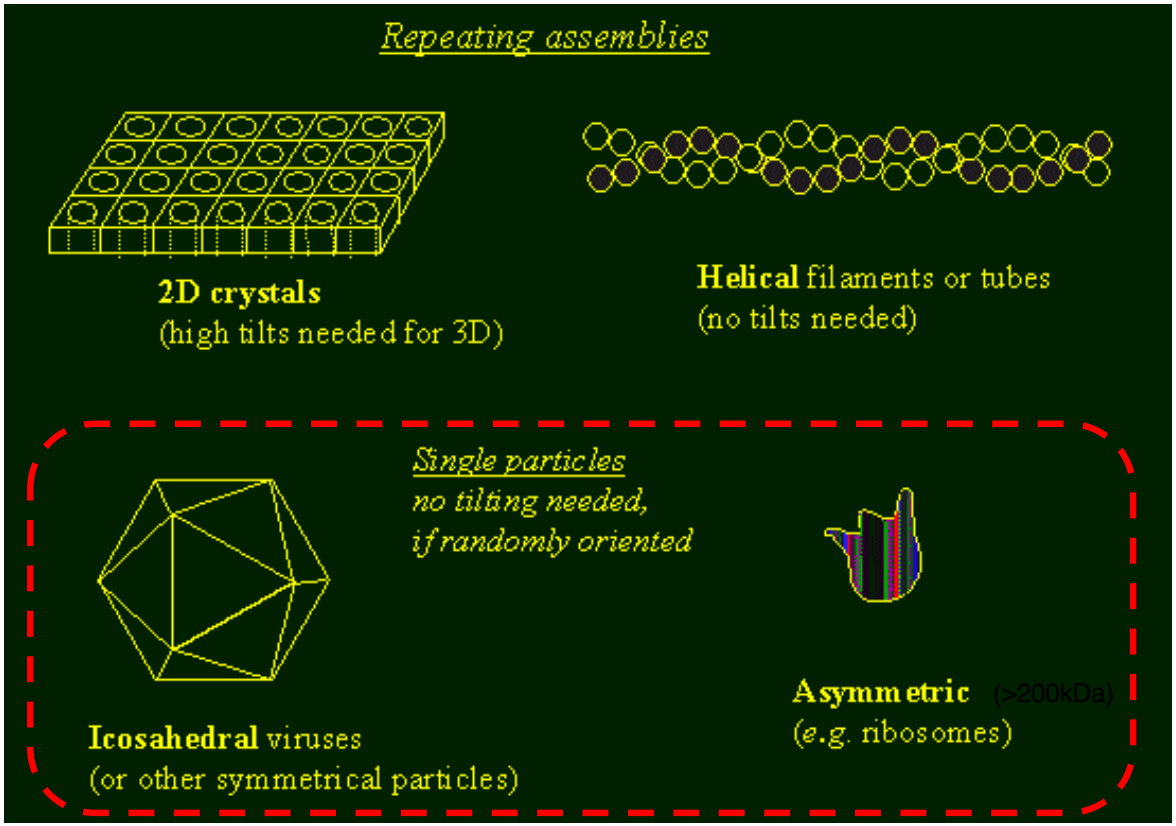
Wet Lab Workshop
"Molecular cryo-EM: Structural Biology without Crystals

Part III: Challenges and opportunities

Jyh-Yeuan (Eric) Lee, Assistant Professor, BMI

October 7th, 2019

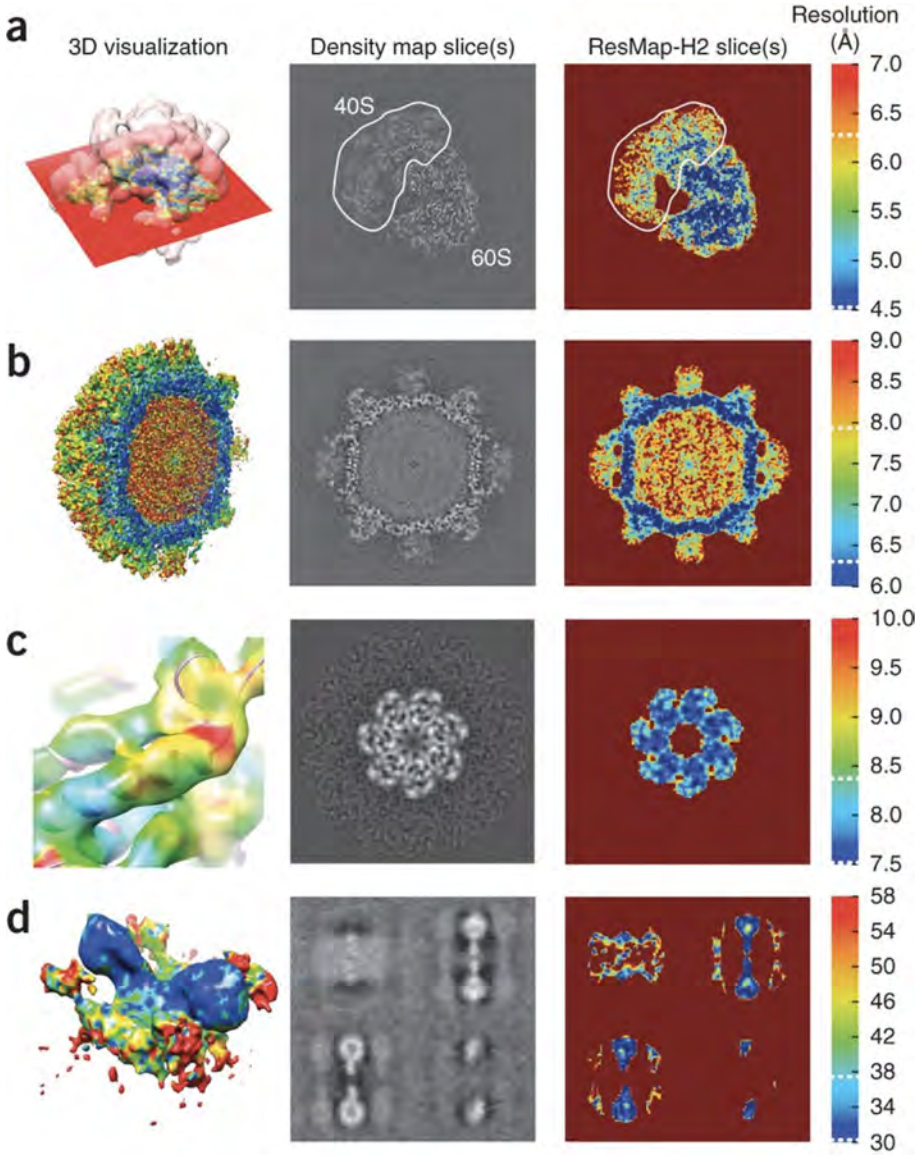
Cryo-EM Friendly Samples



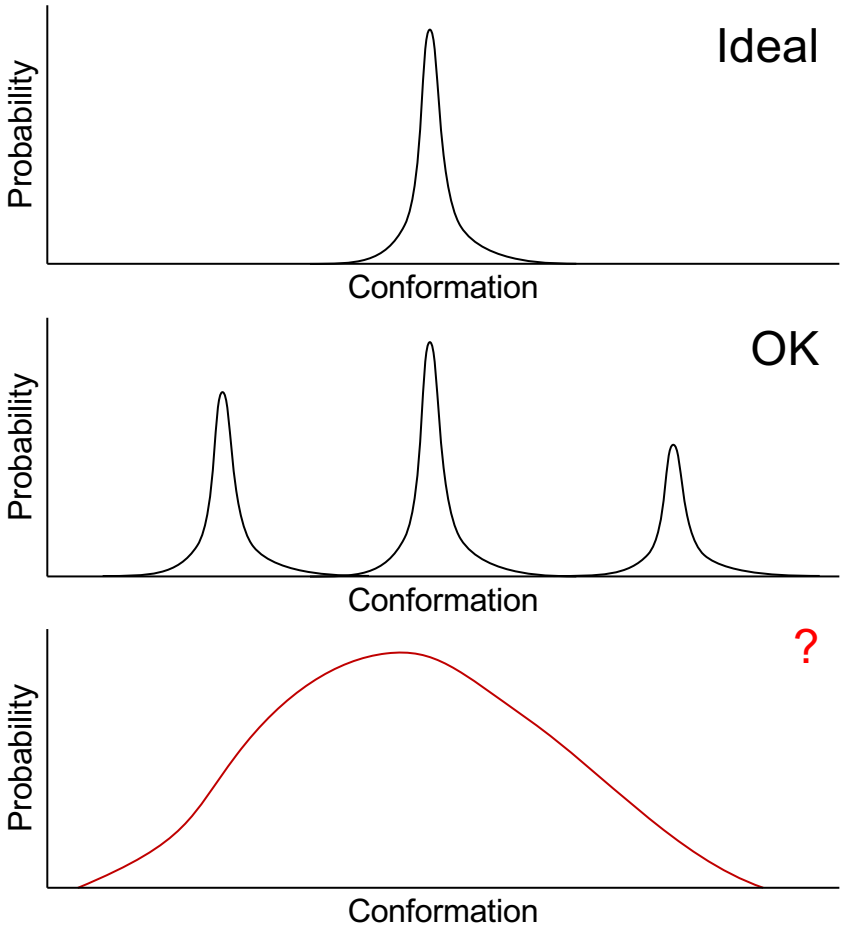
Saibil, Acta Cryst. 2000, D56:1215

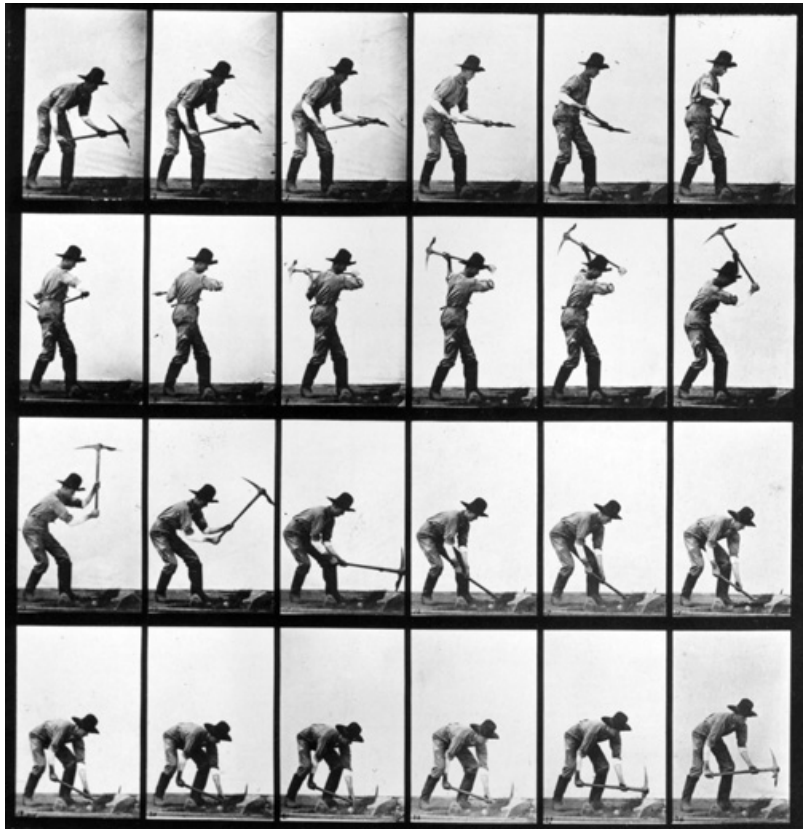
Challenges

- Molecular size:
  - >200 kD
  - 100-200 kD, pushing the limit
  - Otherwise growing crystals
- Resolution:
  - Mostly 3-5 Å
  - Overfitting
- Conformational variability
  - Only a small number of functional states are solved.
  - Preferred orientation
- Validation

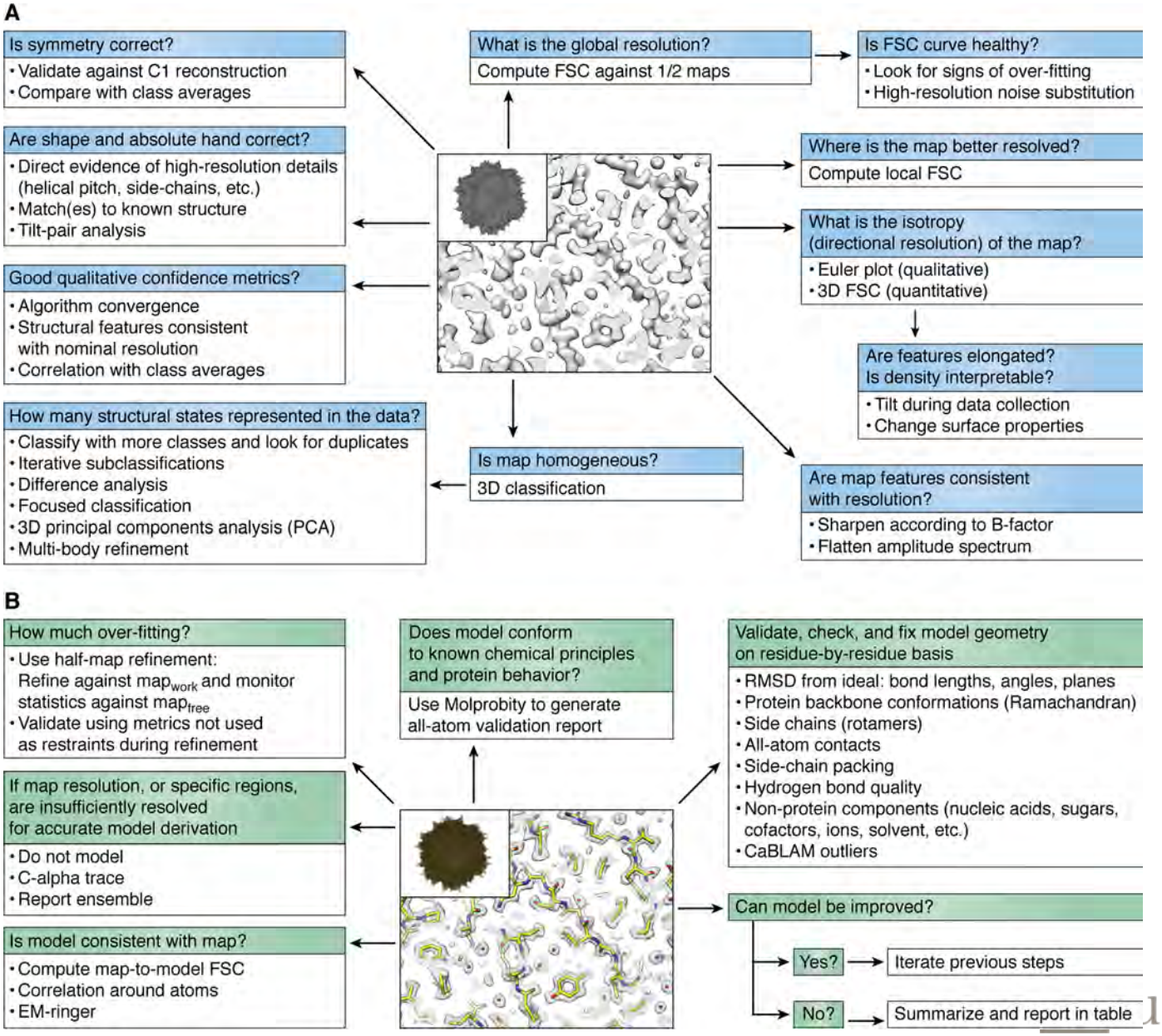


(Kucukelbir et al, Nat Methods, 2014)





Only small number of discrete states are solvable now

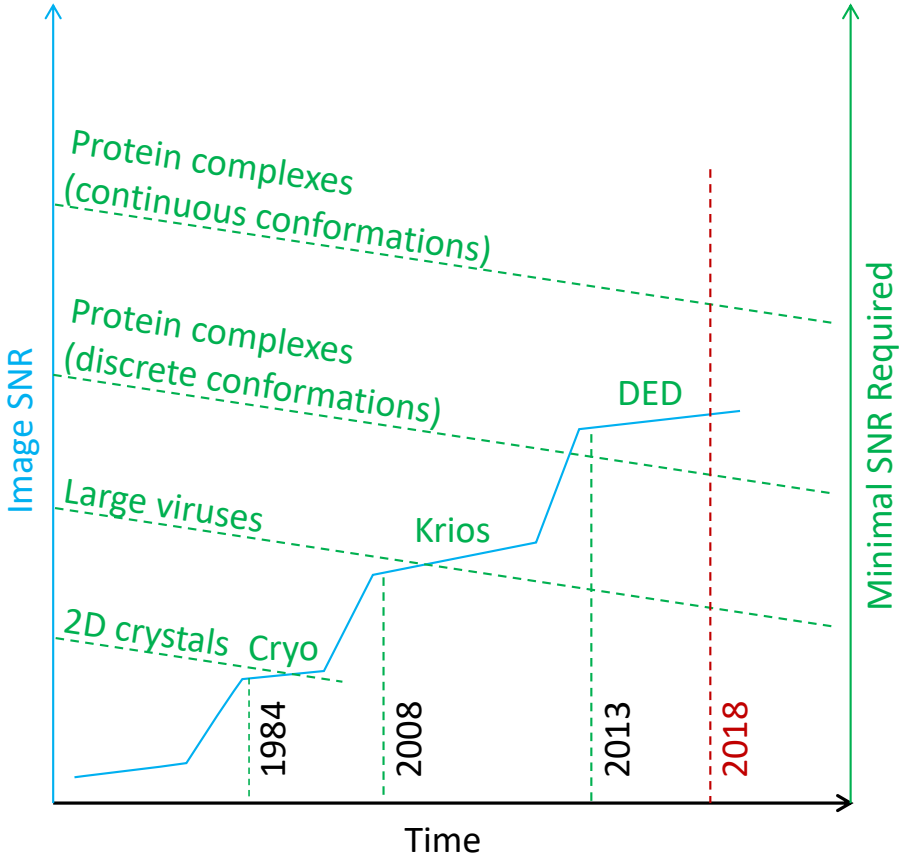


(Dmitry Lyumkis, JBC, 2019)

Opportunities

- Imaging <200 kD samples
- Instrumental improvement:
  - Routinely <3 Å
- Capture the macromolecular complexed directly from small-volume cell cultures
- Determine 1) structures of all states in the mixture and 2) orders of states in terms of their functional processes
- Better validation criteria

It's all about S/N ratio (SNR).



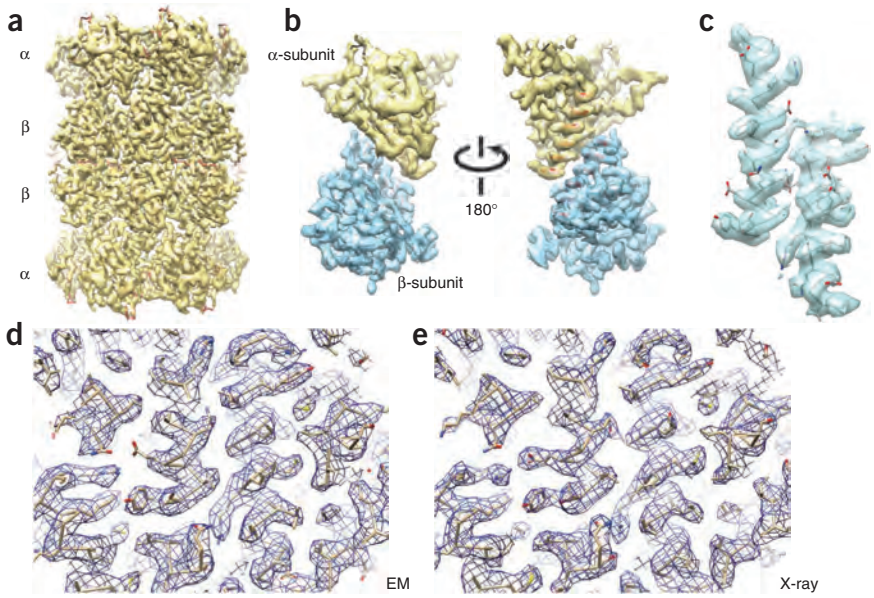
Resolution Revolution?

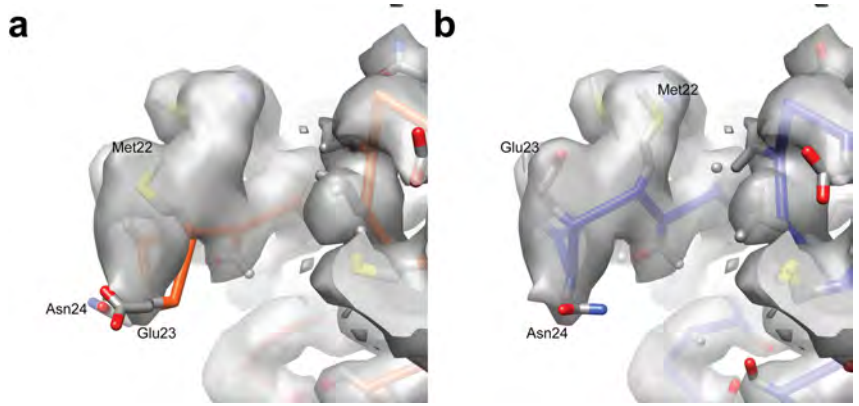
Started 40 years ago.

More to come.

Cryo-EM *v.s.* X-ray Crystallography

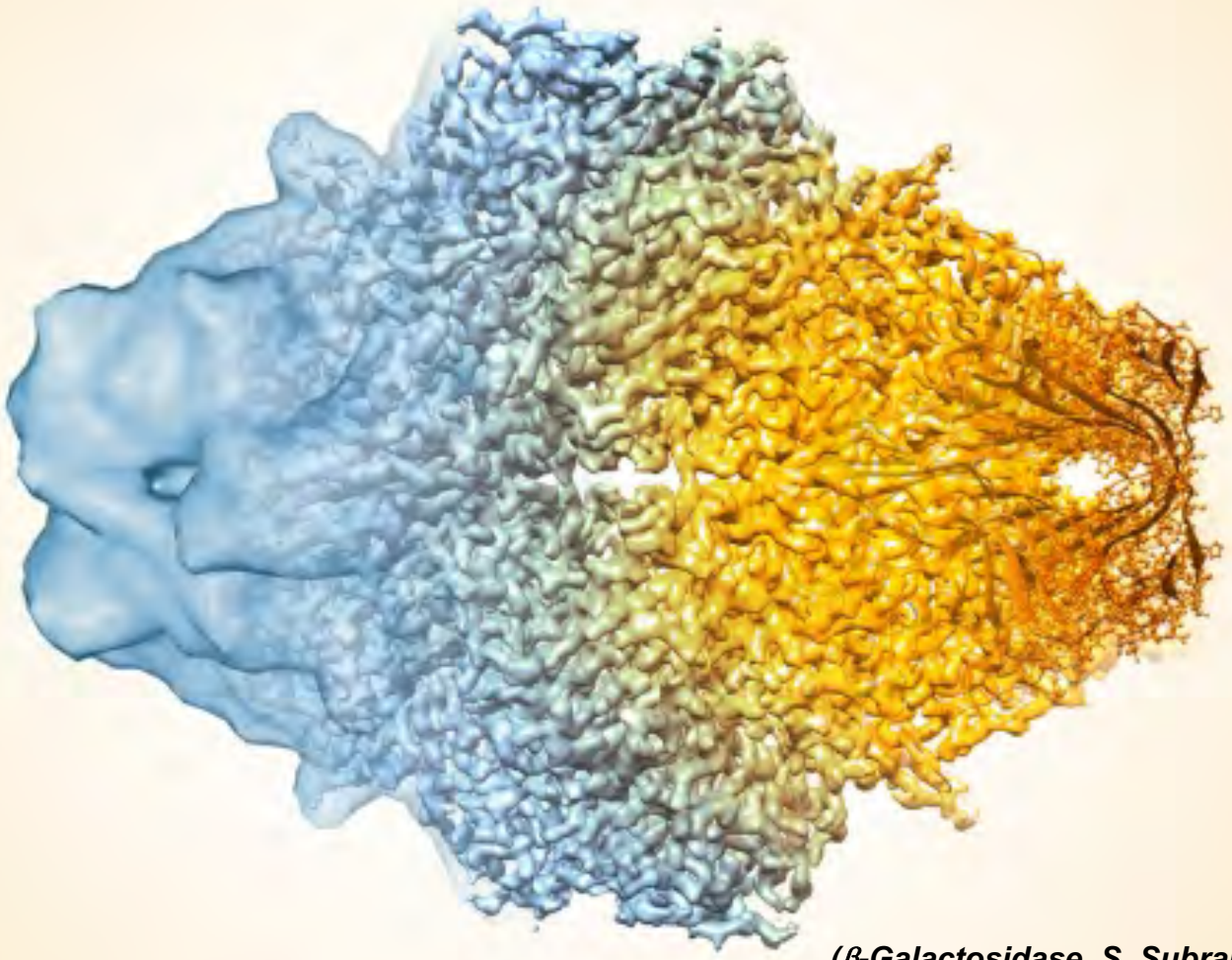
Friends and Rivals





a. A short loop (Met22 ~ Asn24) in the β subunit from the crystal structure does not fit well into the final 3D density map, but is easily corrected. While not in a crystal contact, these residues do have much higher temperature factors in the 3.4 Å crystal structure than the average. **b**. We remodeled this loop to fit better into our 3.3 Å density map.

(Li et al, Nat Methods, 2012)



(β-Galactosidase, S. Subramaniam, UBC)