BIO-315
Structural Biology

Introduction to Electron Microscopy
- Lecture 2 -

Aleksandar Antanasijević

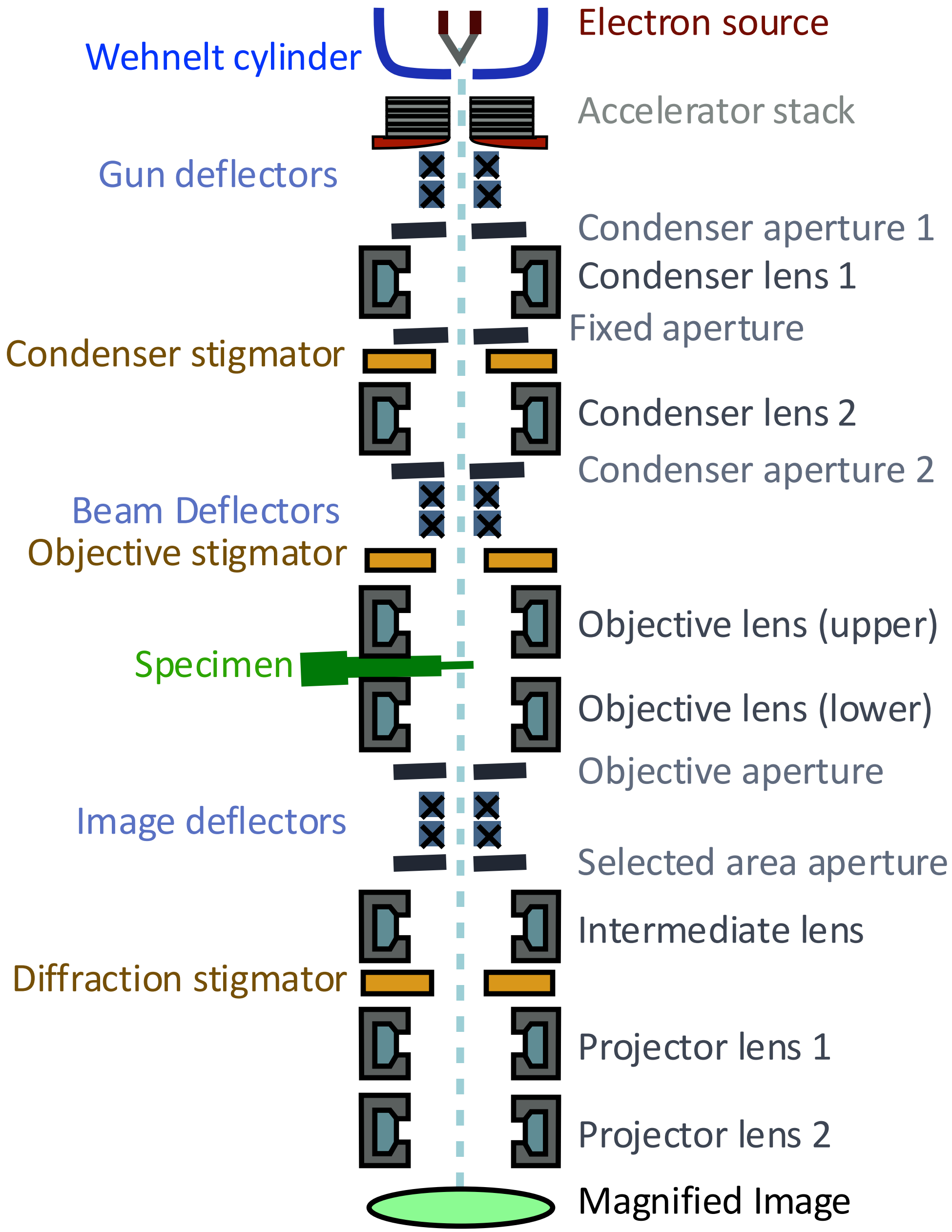
Global Health Institute

EPFL

19-03-25

Recap of the last lecture

Microscope components



“Resolution Revolution” and “Democratization” of EM

Improved resolution of EM data

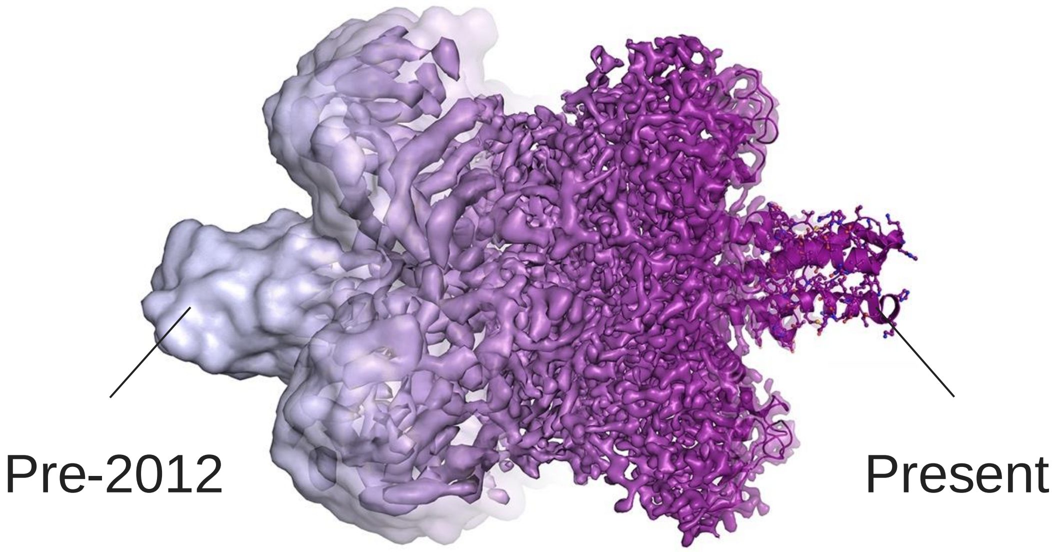


Illustration by Martin Högbom; The Royal Swedish Academy of Science

Nobel Prize in Chemistry, 2017

“for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution”

- High-resolution maps became readily attainable starting ~2012 as a result of technological breakthroughs in the field

Electron Microscopy for Structure Determination

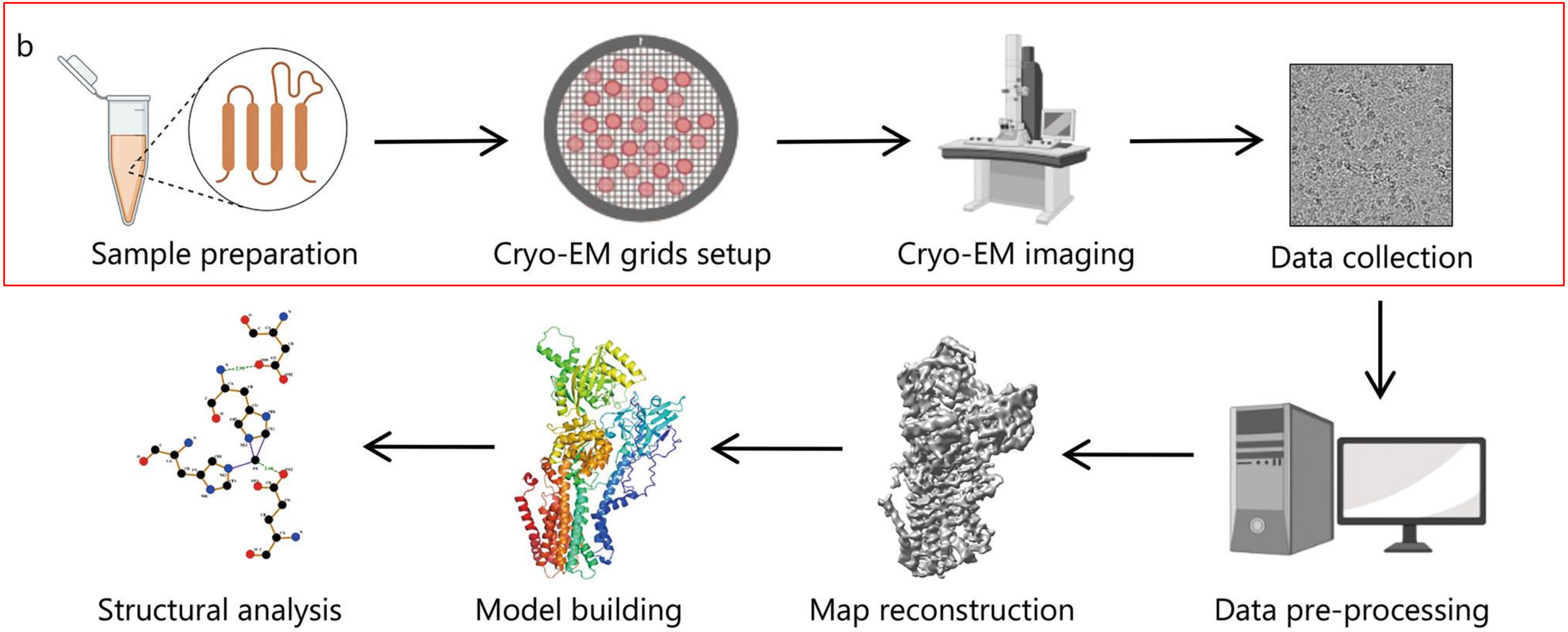
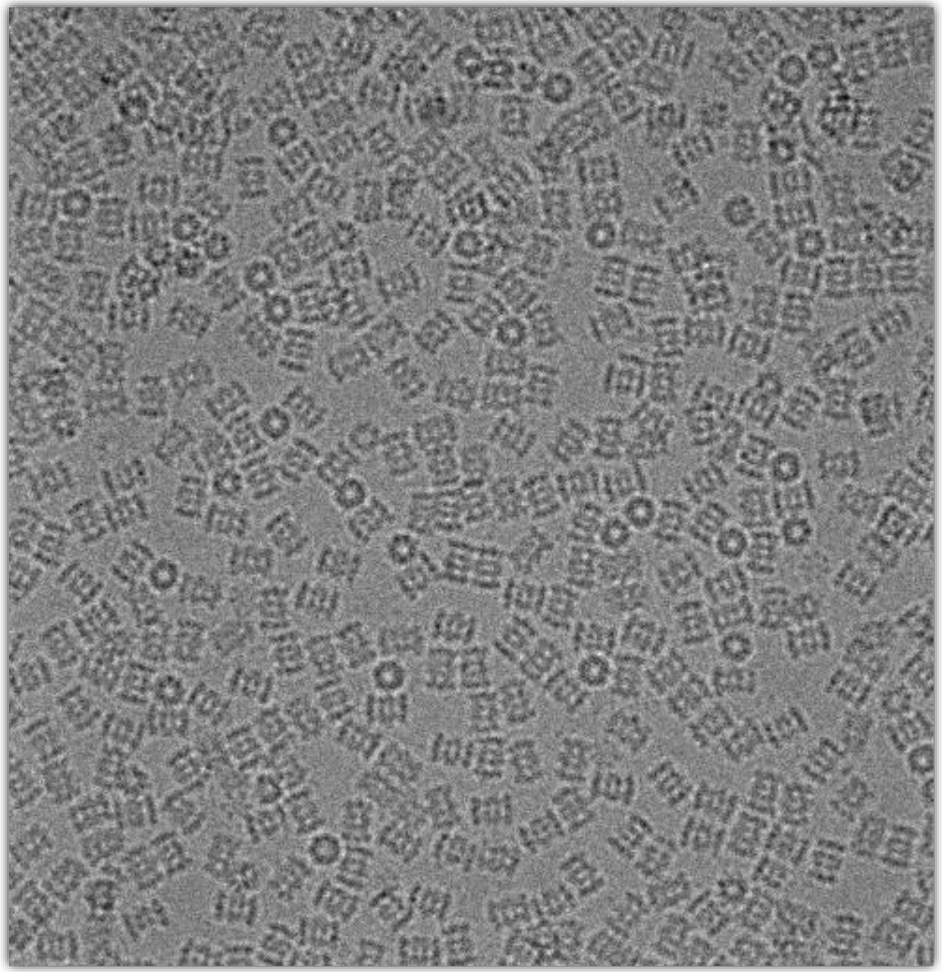


Image generated by cryo-electron microscopy

- Images in cryoEM are generated by electrons scattering from biomolecules
- Due to the transmission mode of data collection, 2D projections of biomolecules appear dark



Real-space image

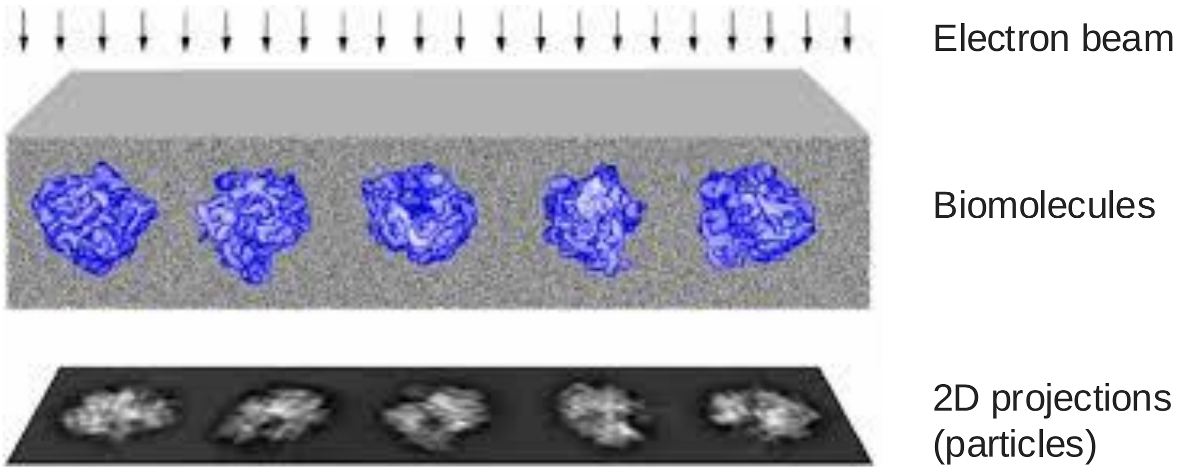
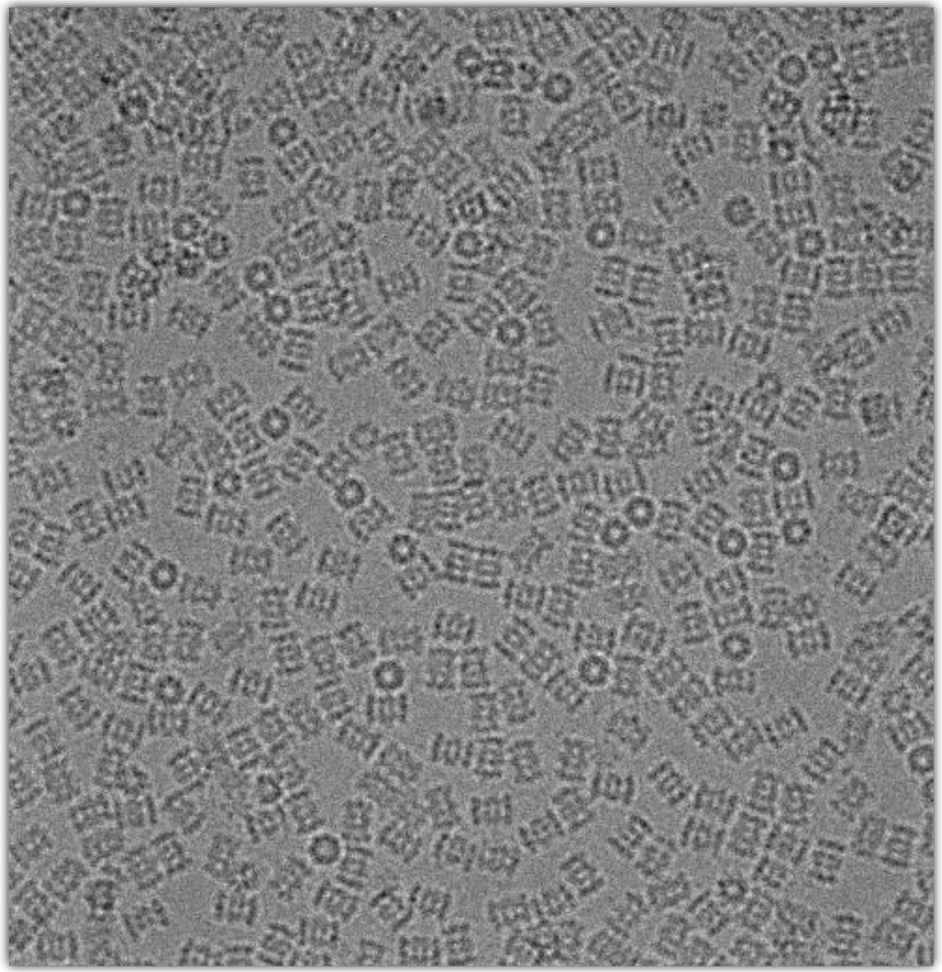


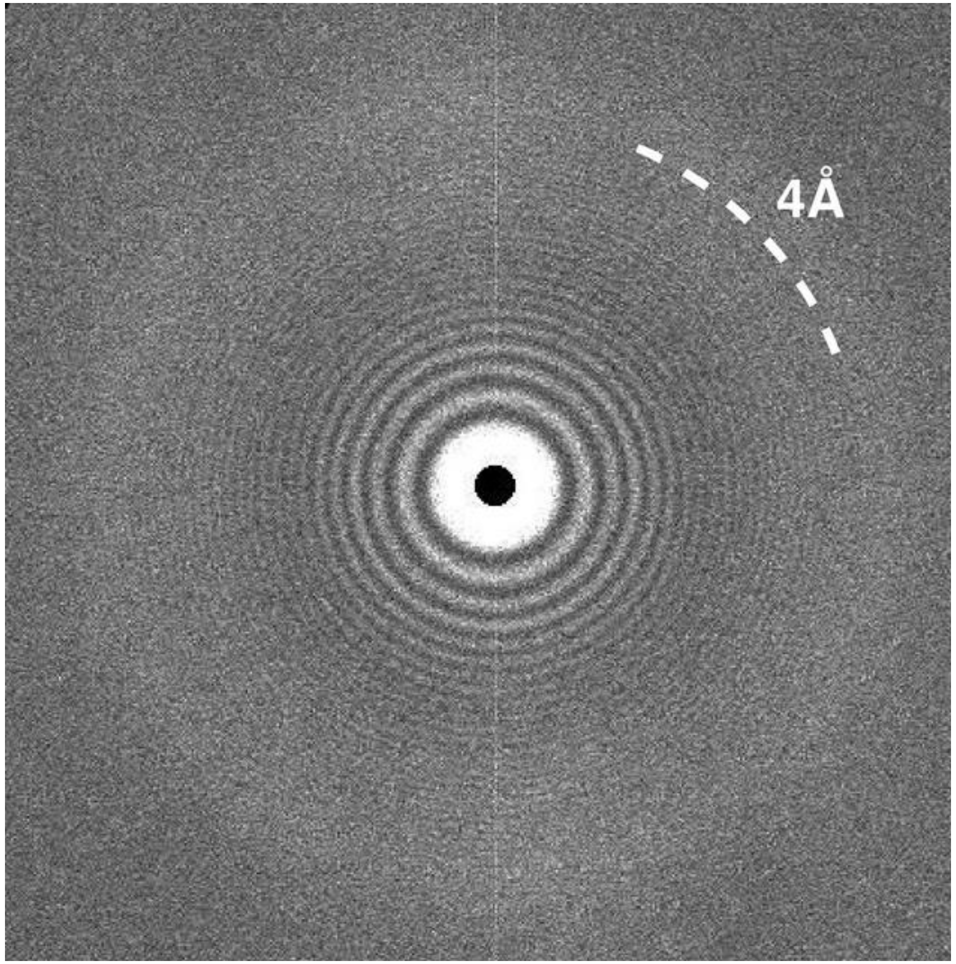
Image generated by cryo-electron microscopy

- Fourier transform of the real-space image is called the **power spectrum**
- It allows to assess basic properties of the image and identify potential problems



Real-space image

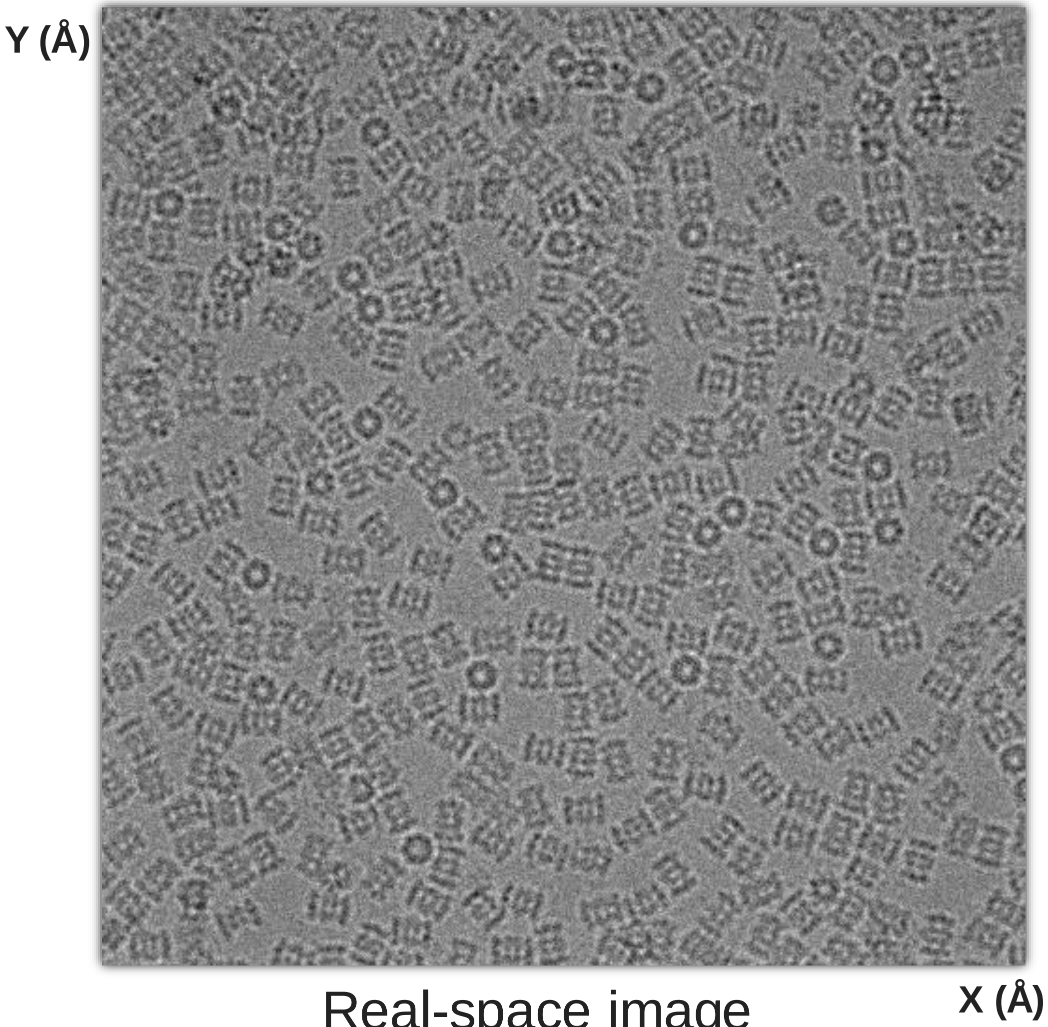
FFT →



Power spectrum

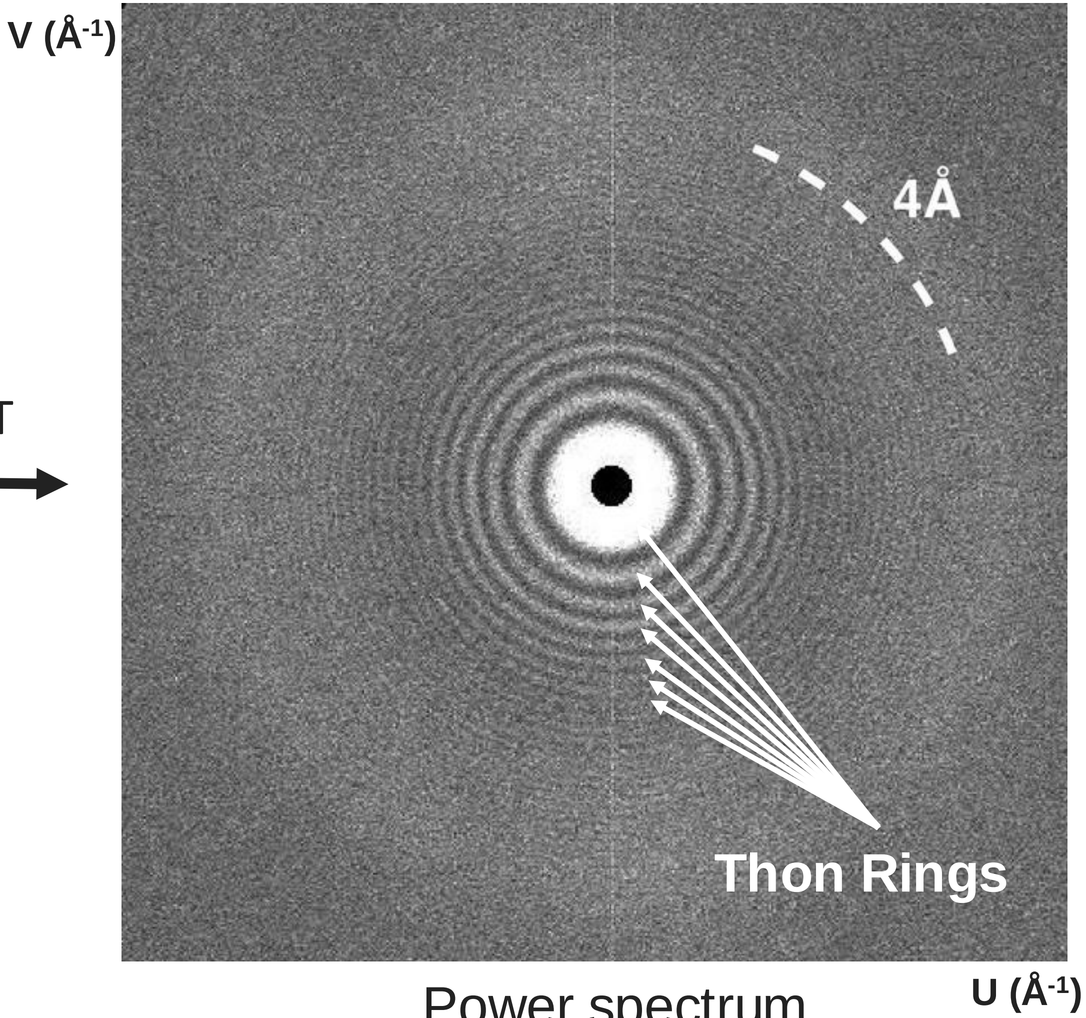
What is the origin of Thon rings?

- FT images have the unit of spatial frequency (**1/distance**)
- Resolution increases as you move away from center (**lower distance values = higher resolution**)



Real-space image

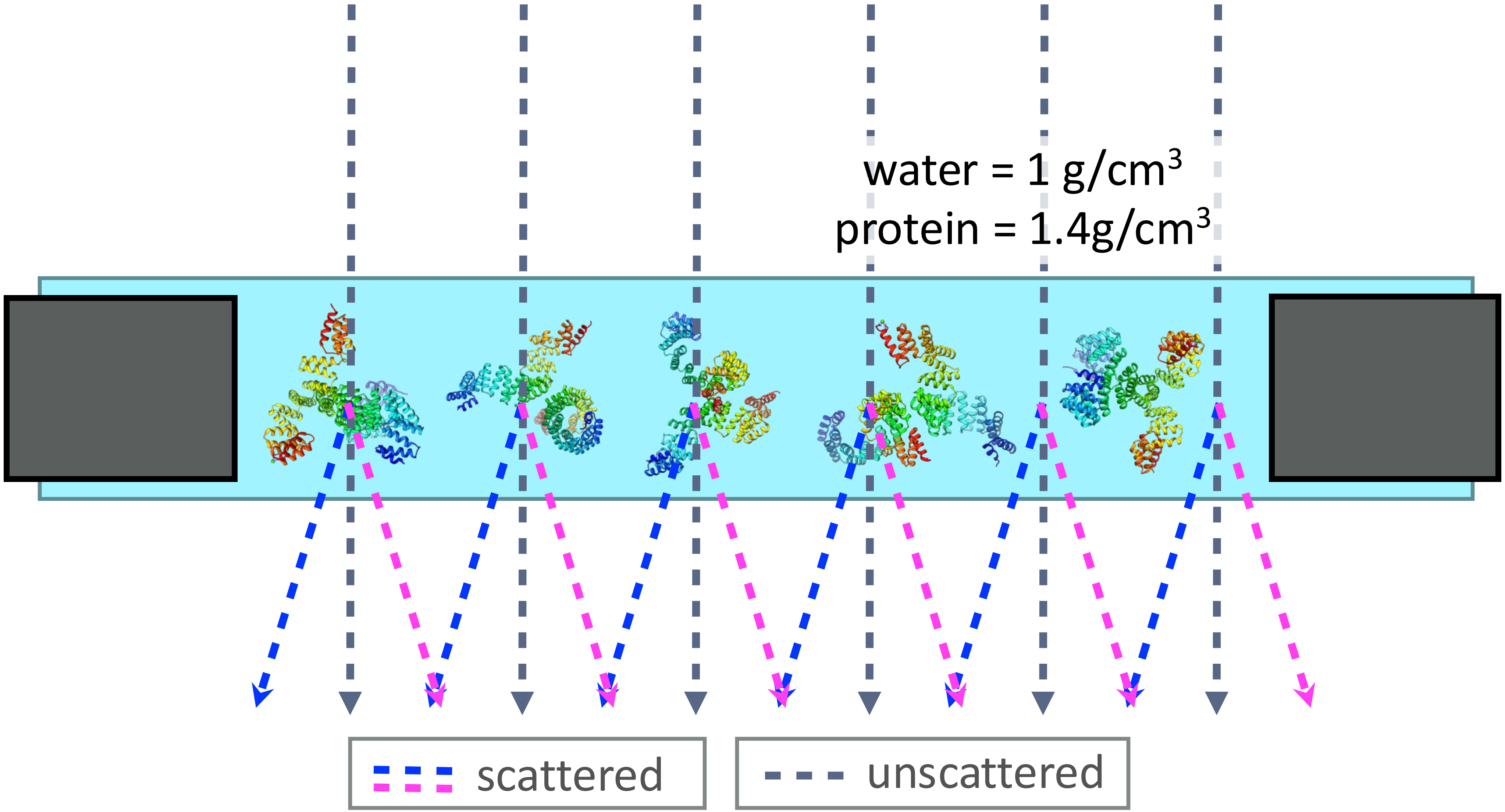
FFT



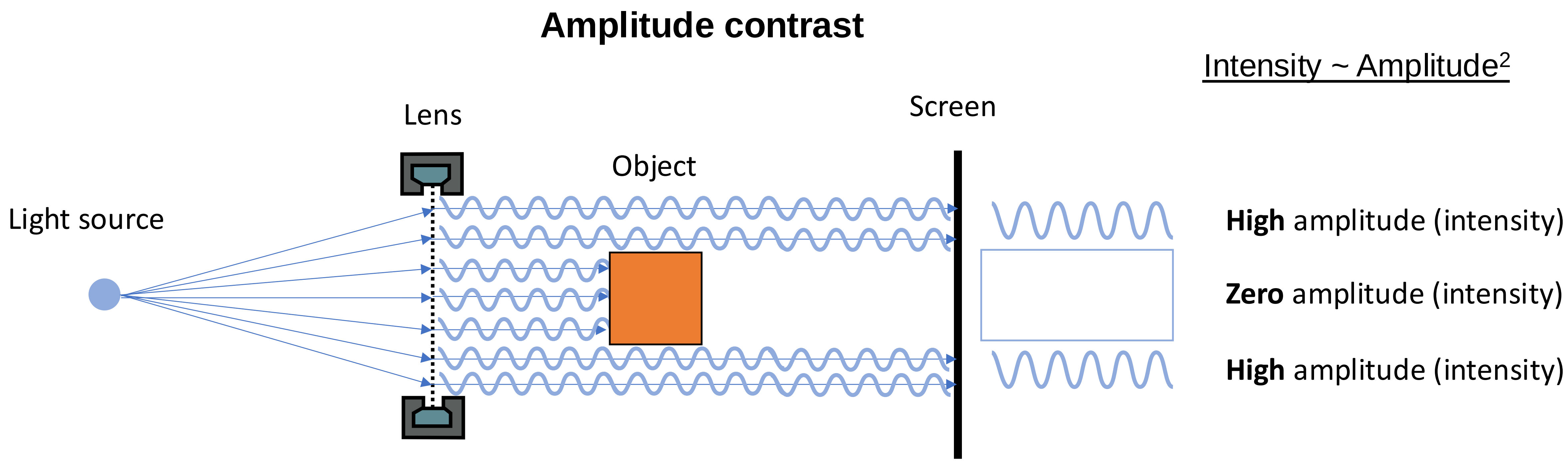
Power spectrum

Contrast Transfer Function

How is contrast generated in EM images



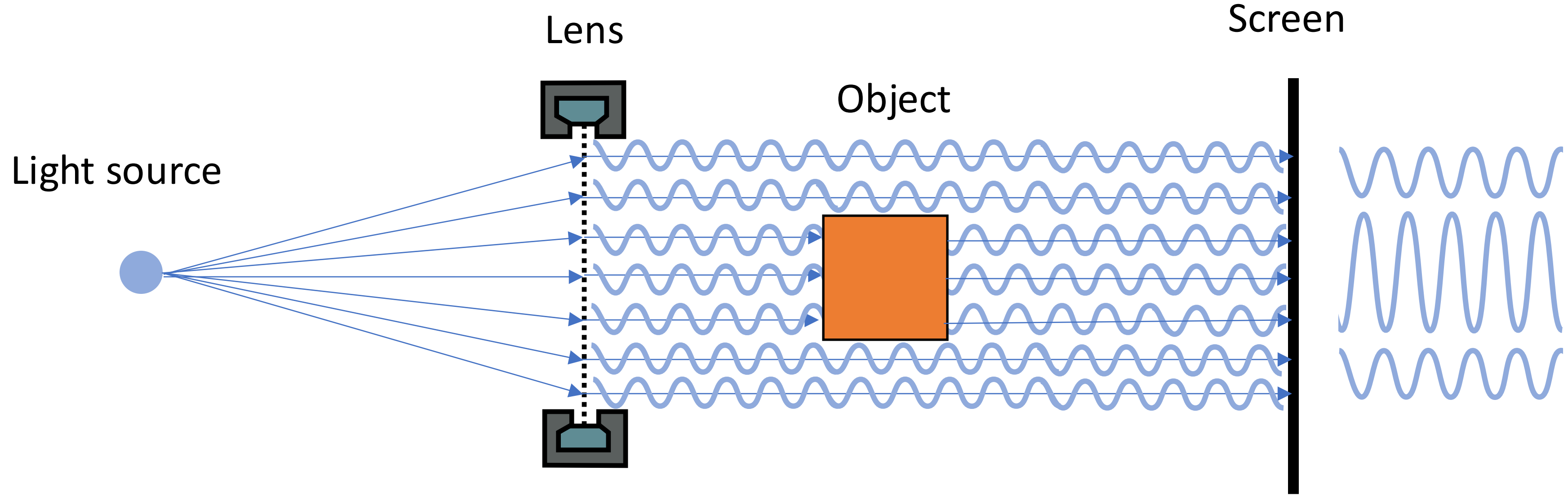
Amplitude Contrast vs Phase Contrast



- Amplitude contrast based imaging produces differential intesities on the screen based on the interaction with the object
- Only ~10% of contrast in cryoEM is generated from amplitude contrast

Amplitude Contrast vs Phase Contrast

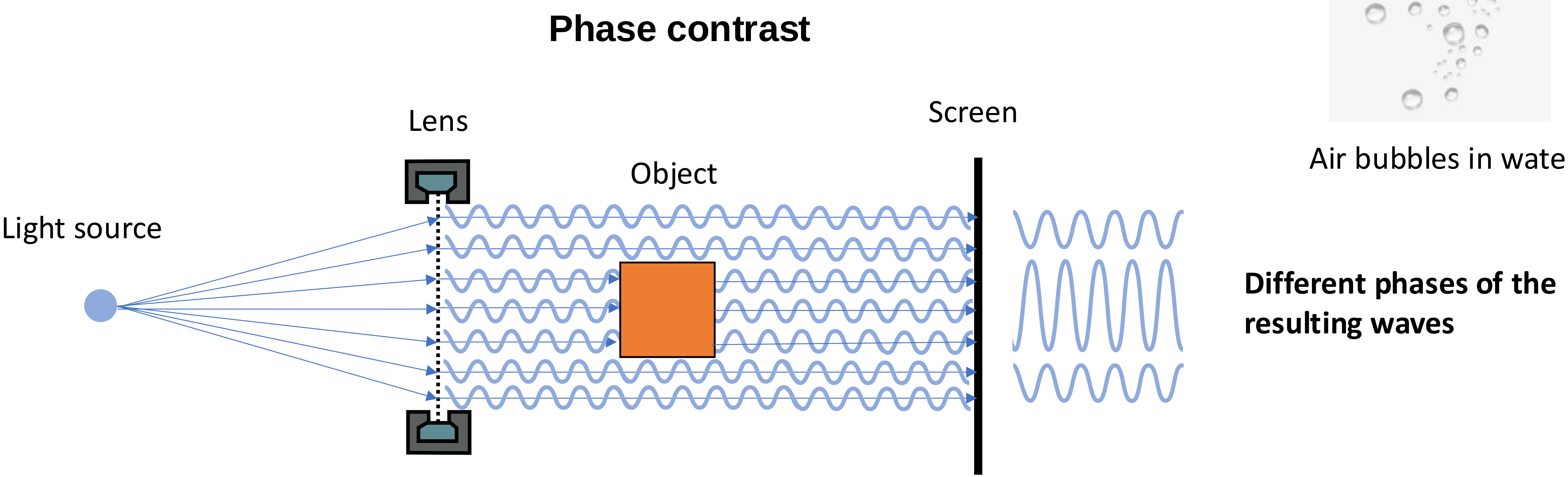
Phase contrast



Different phases of the resulting waves

- Phase contrast is based on bending, refracting or delaying the passage of light through the sample by different amounts
- It is responsible for most contrast seen in cryoEM images

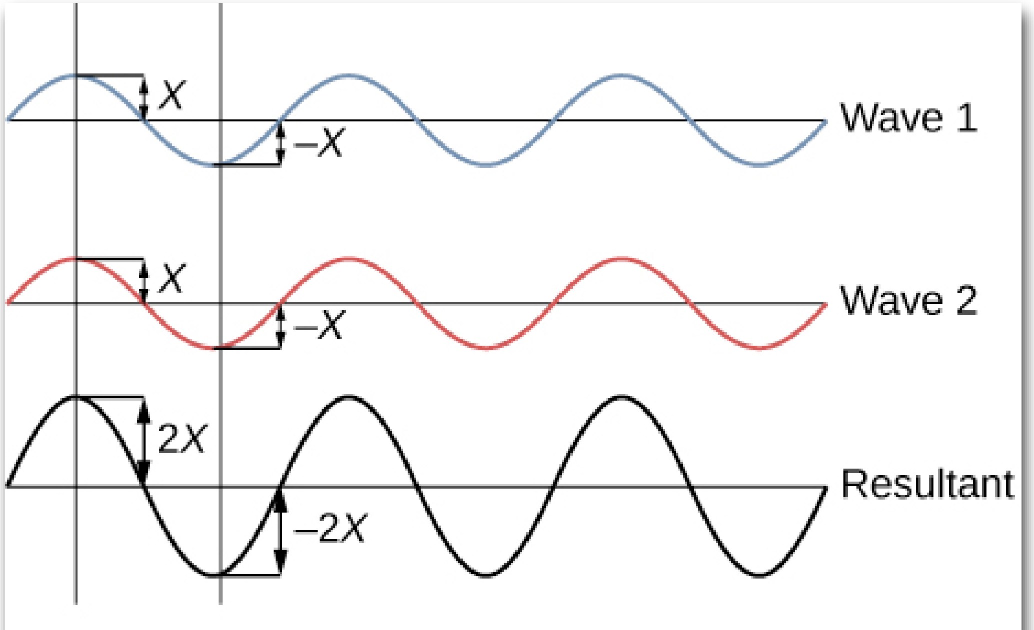
Amplitude Contrast vs Phase Contrast



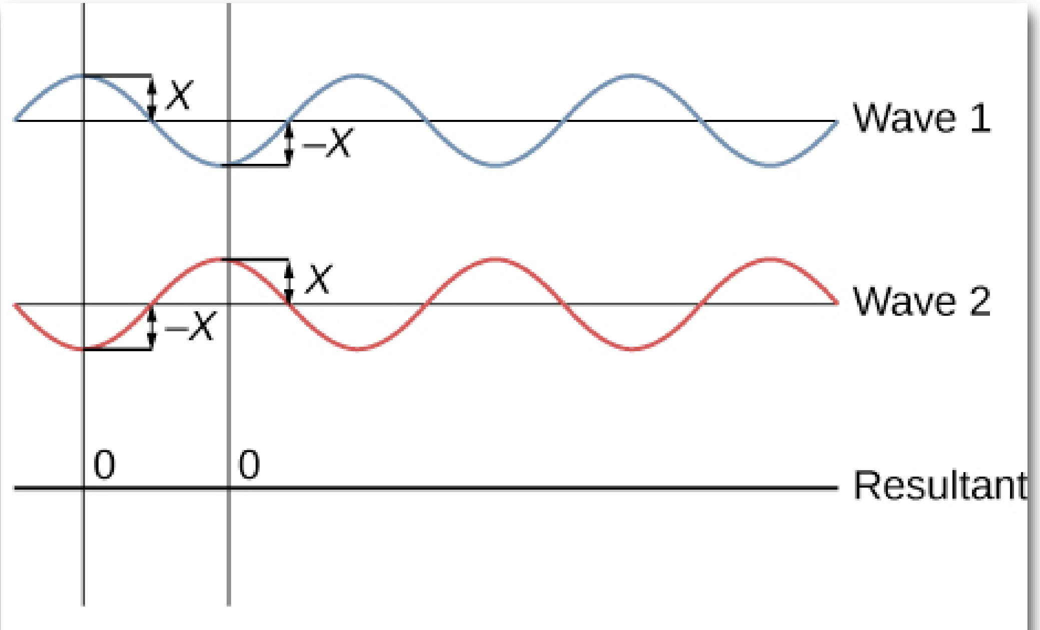
Air bubbles in water

- Phase contrast is based on bending, refracting or delaying the passage of light through the sample by different amounts
- It is responsible for most contrast seen in cryoEM images

Phase contrast works by wave interference



(a) Constructive interference

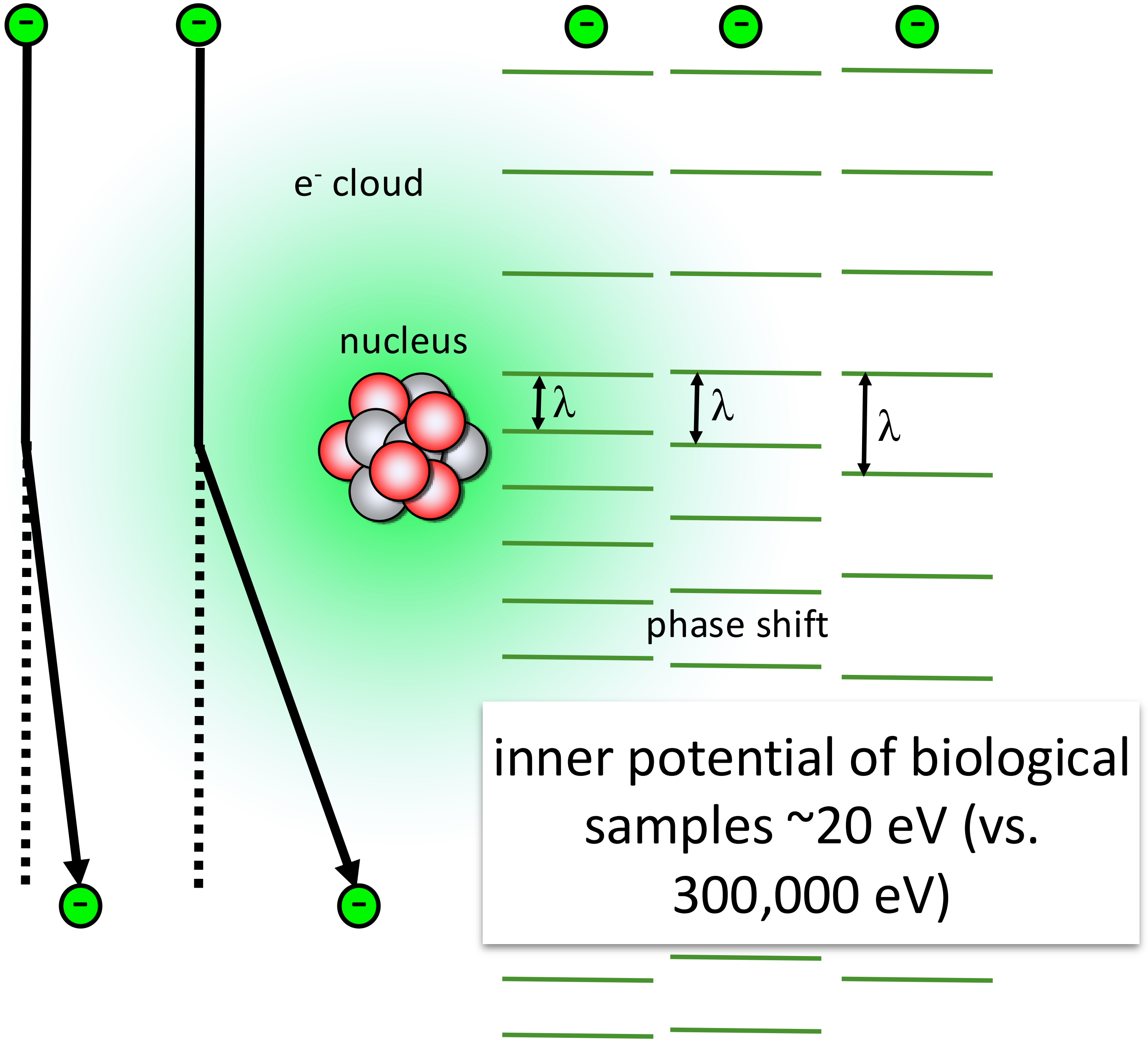


(b) Destructive interference

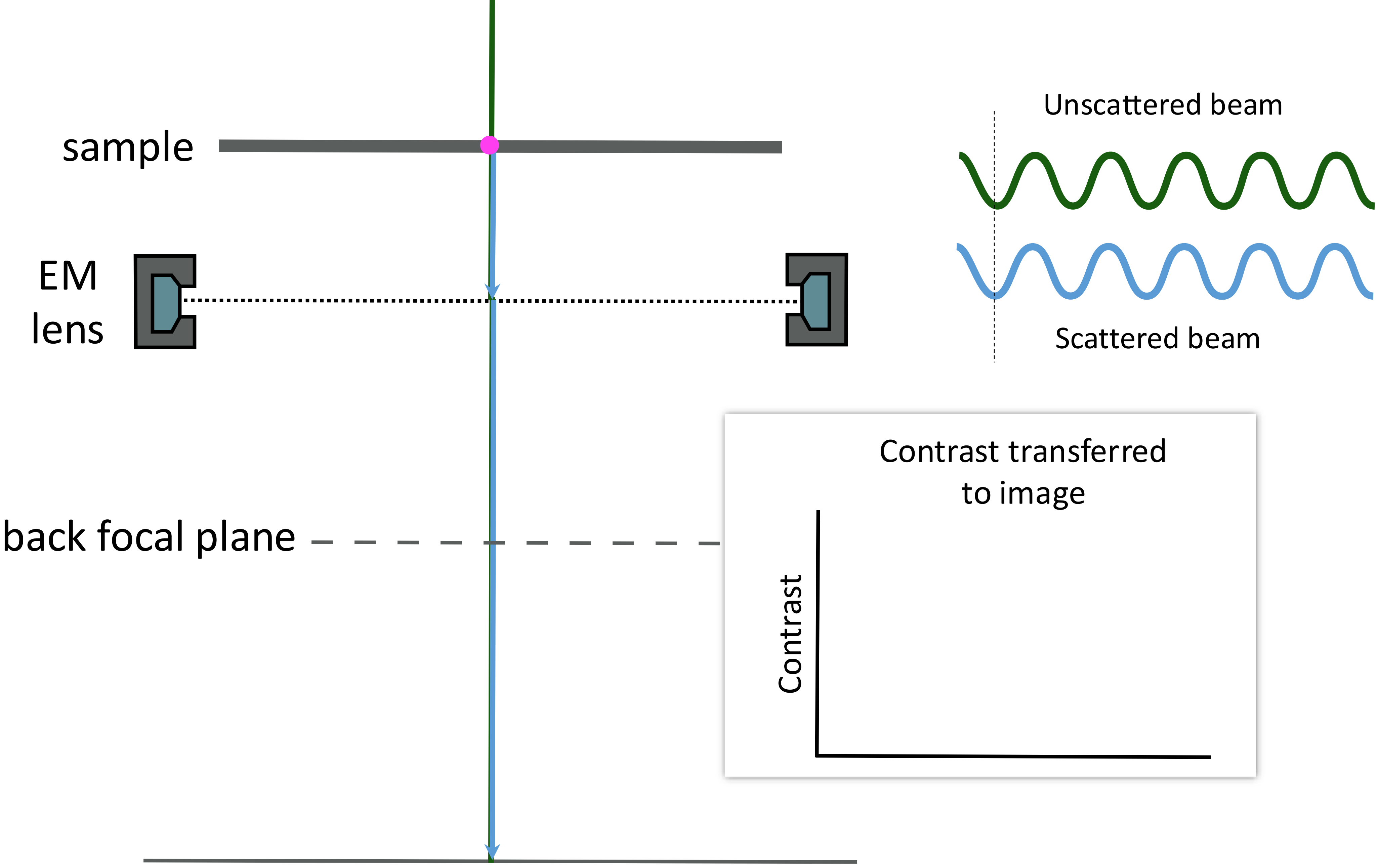
Biological samples are weak phase objects

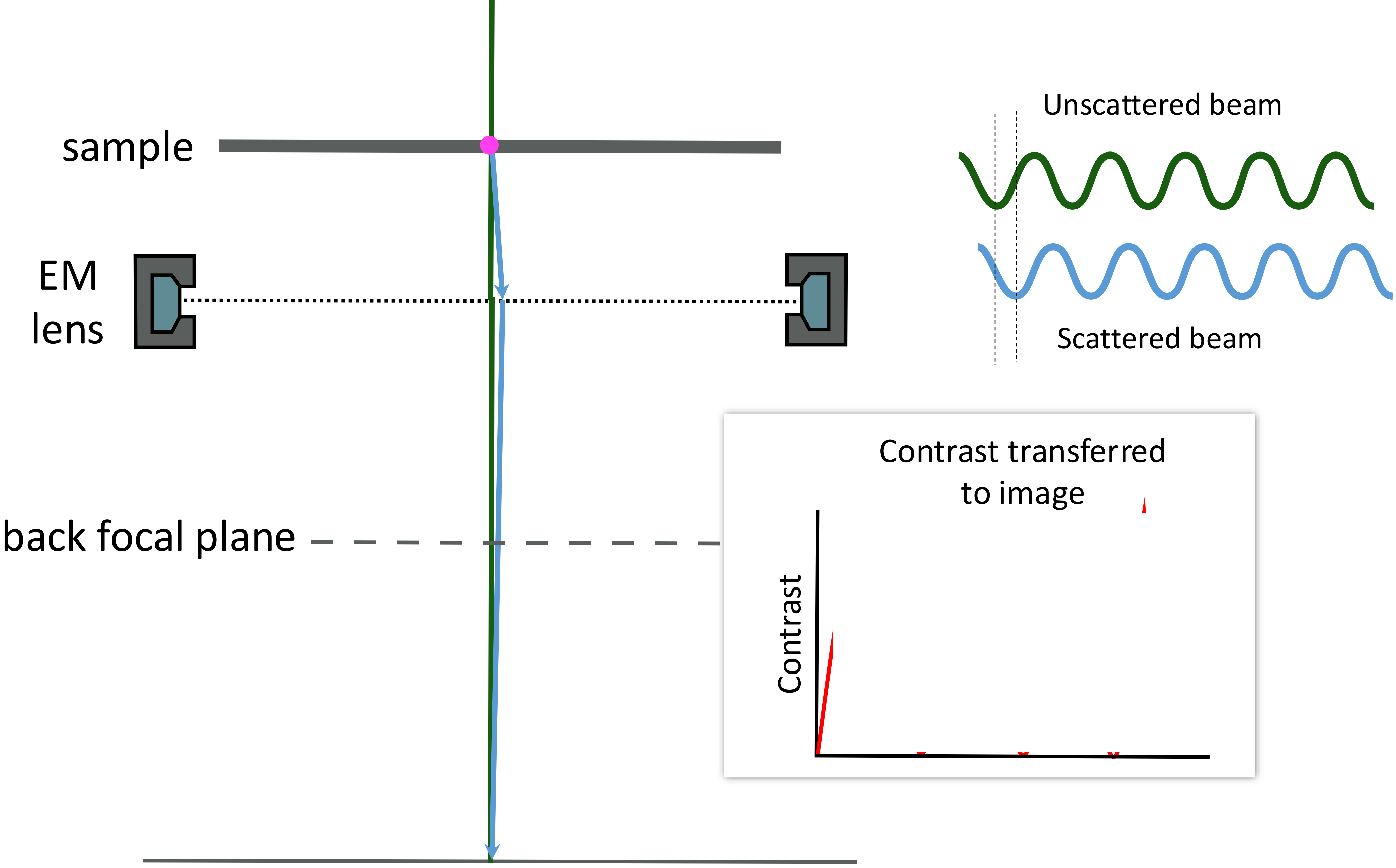
Problems:

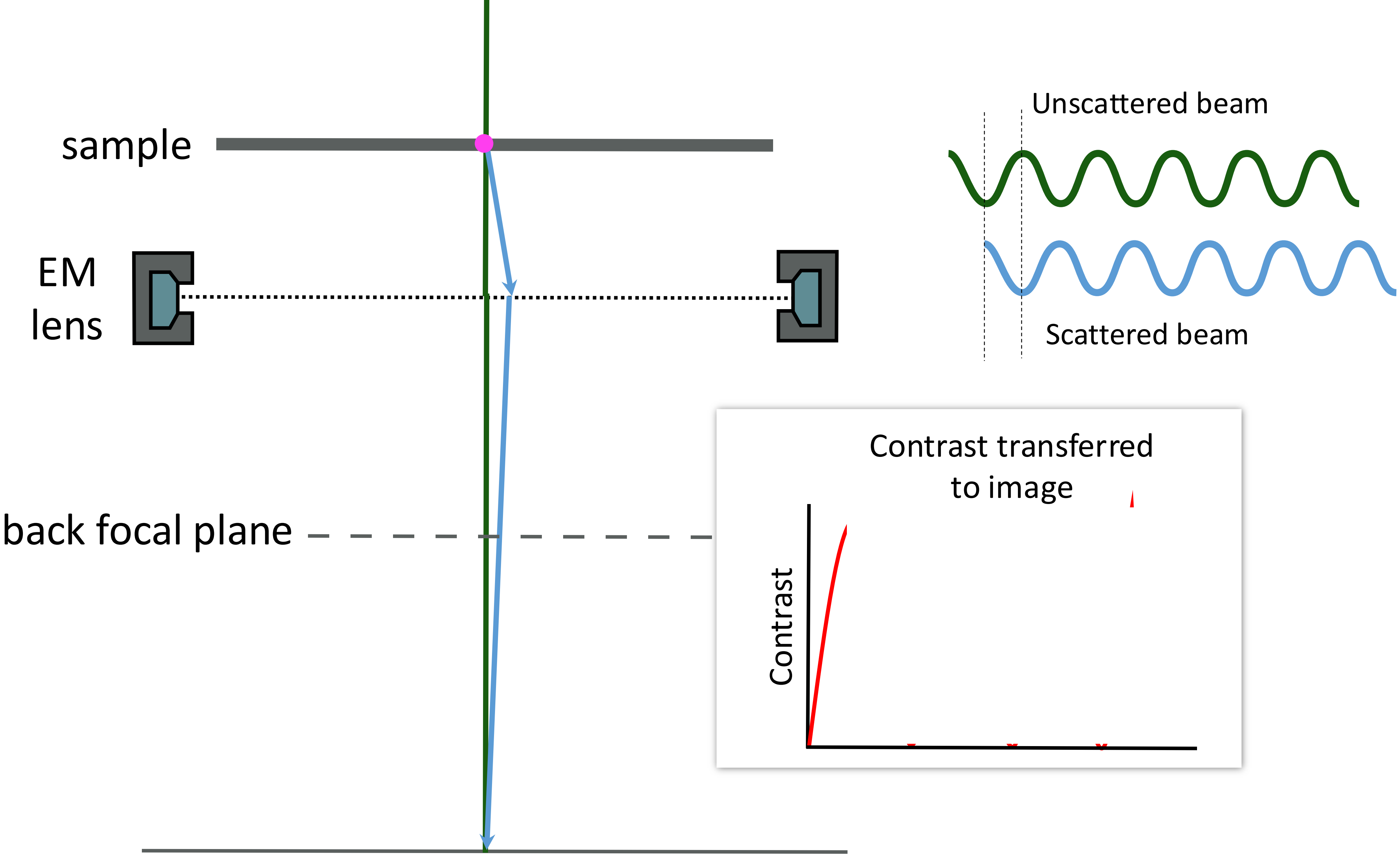
1) Biological specimens are weak phase objects, so the difference between unscattered and scattered is ~300,000 eV vs. 300,020 eV
2) Only ~10% of the contrast in EM images originates from amplitude contrast
3) Cameras cannot detect phase shifts, they can only measure amplitudes

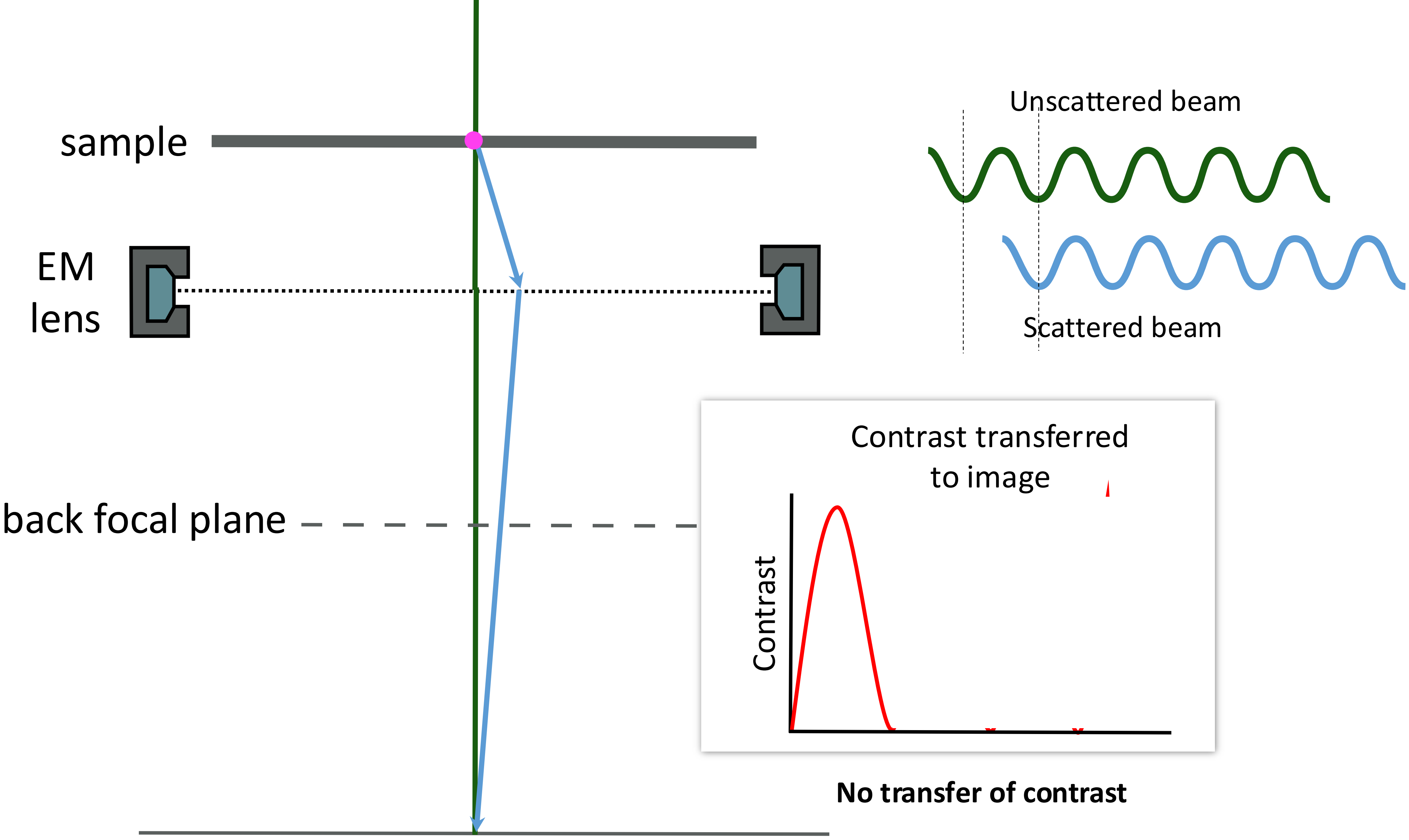


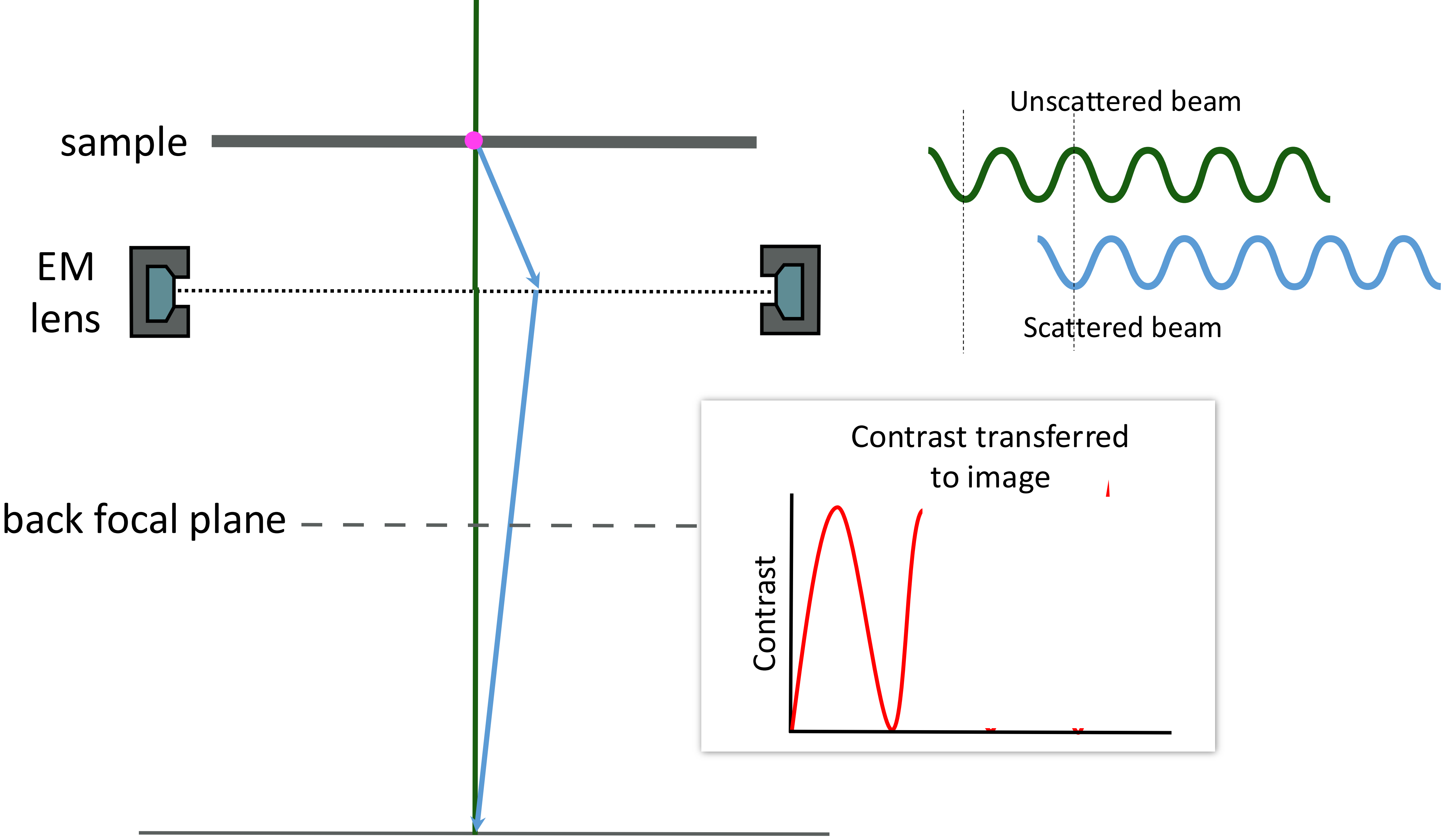
inner potential of biological samples ~20 eV (vs. 300,000 eV)


sample
EM
lens
back focal plane
Unscattered beam
Scattered beam
Contrast transferred
to image
Contrast

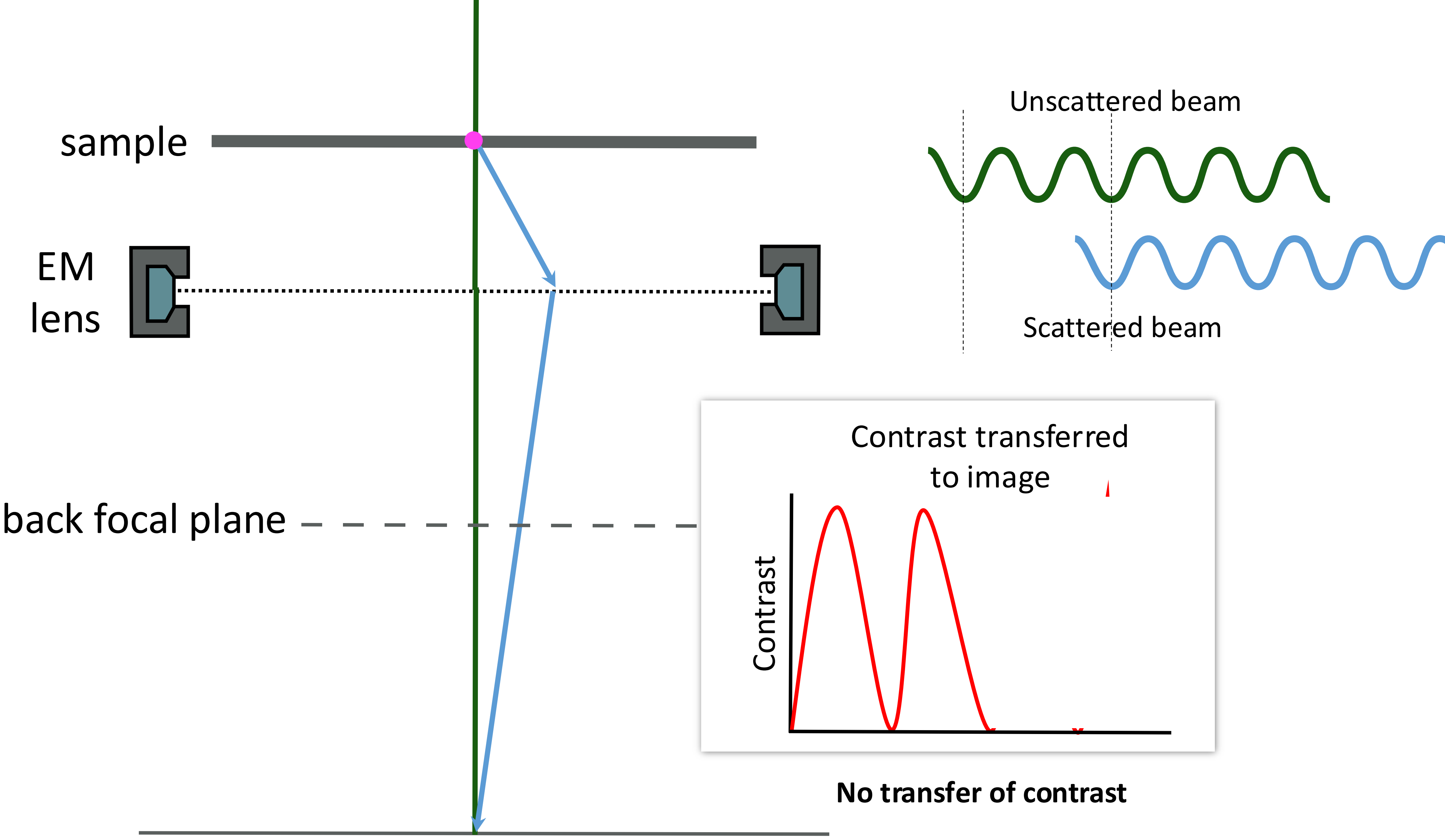

sample
EM
lens
back focal plane
Unscattered beam
Scattered beam
Contrast transferred
to image
Contrast

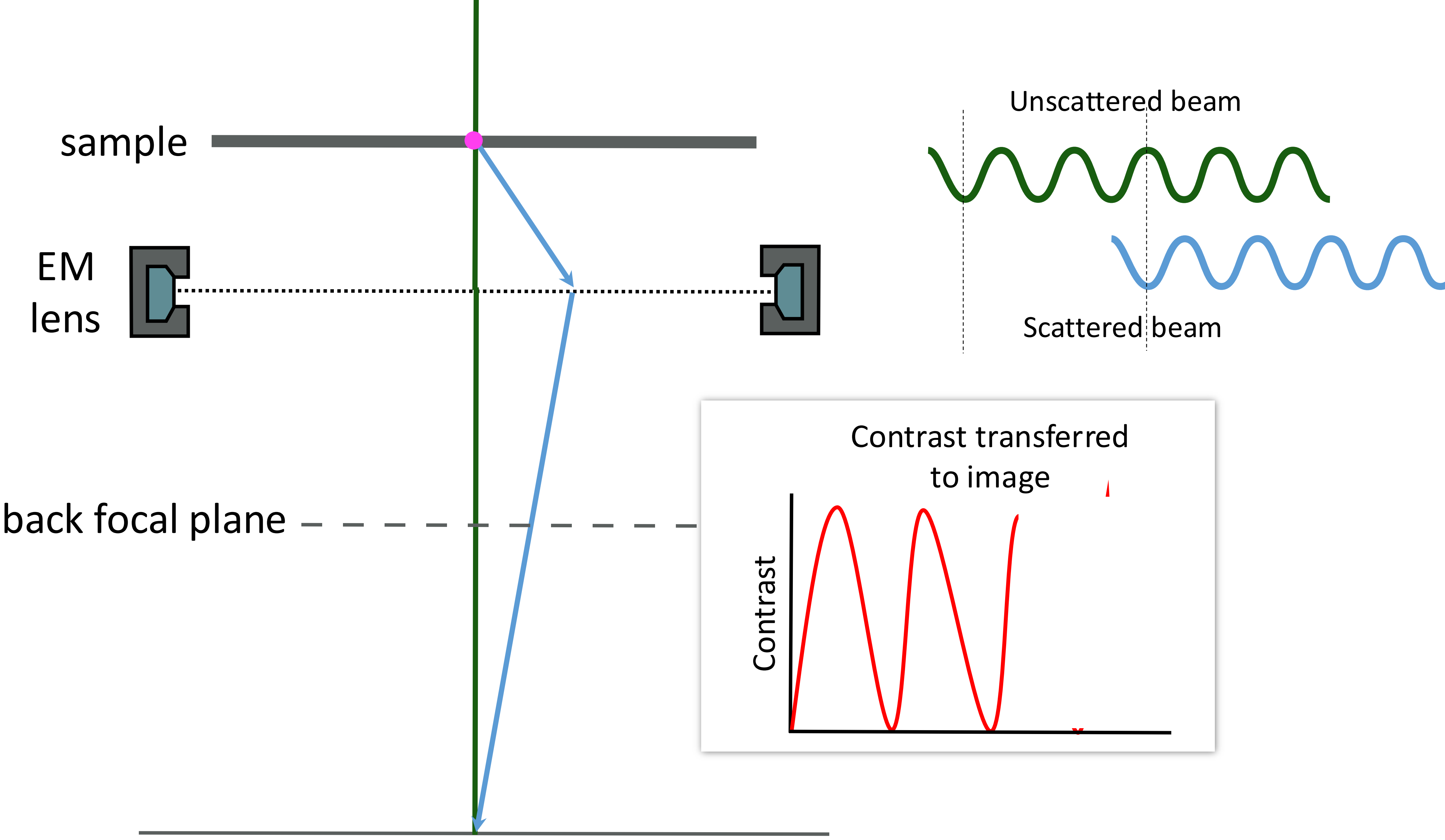

sample
EM
lens
back focal plane
Unscattered beam
Scattered beam
Contrast transferred
to image
Contrast

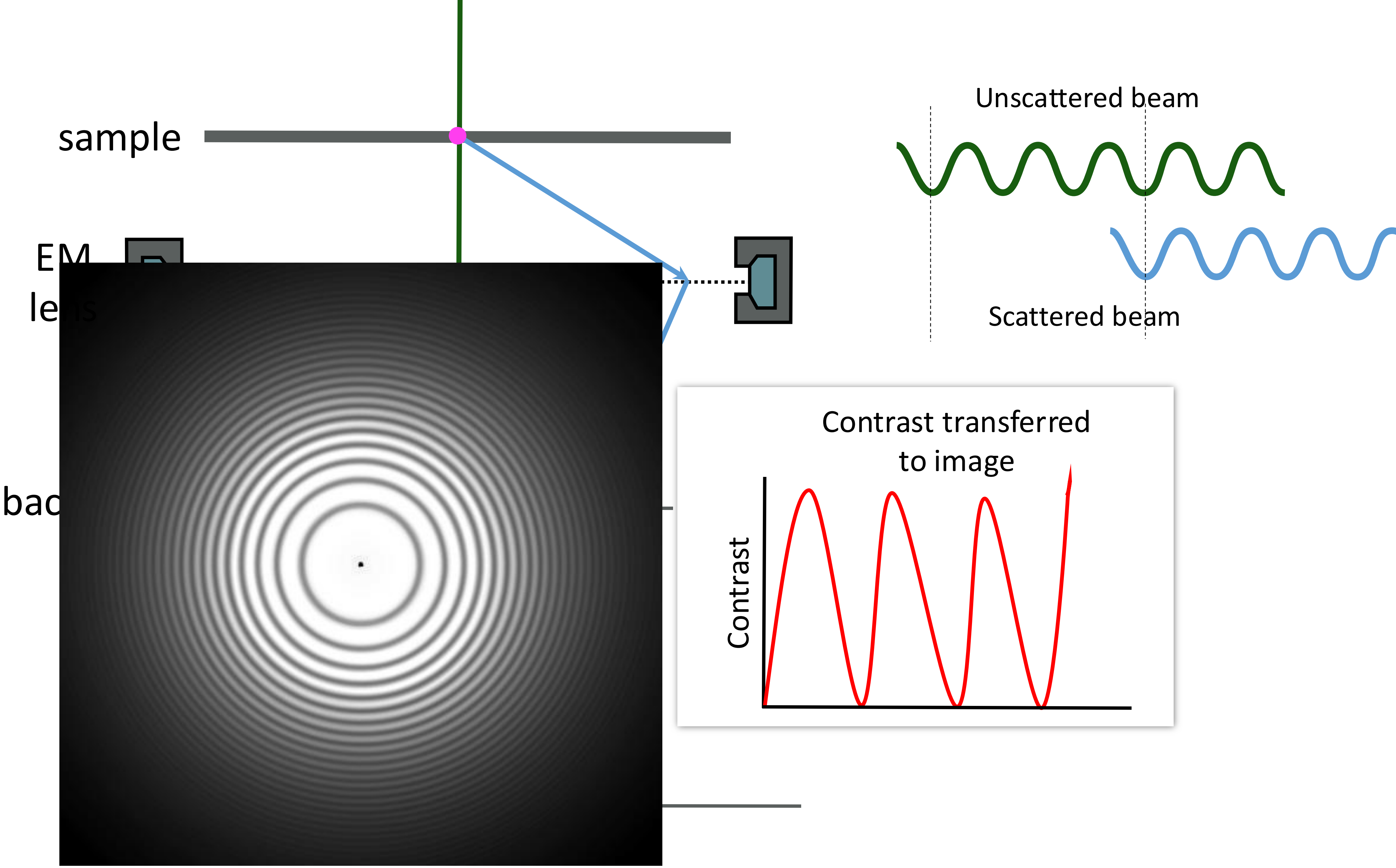

sample
EM
lens
back focal plane
Unscattered beam
Scattered beam
Contrast transferred
to image
Contrast
No transfer of contrast



sample
EM
lens
back focal plane
Unscattered beam
Scattered beam
Contrast transferred
to image
Contrast

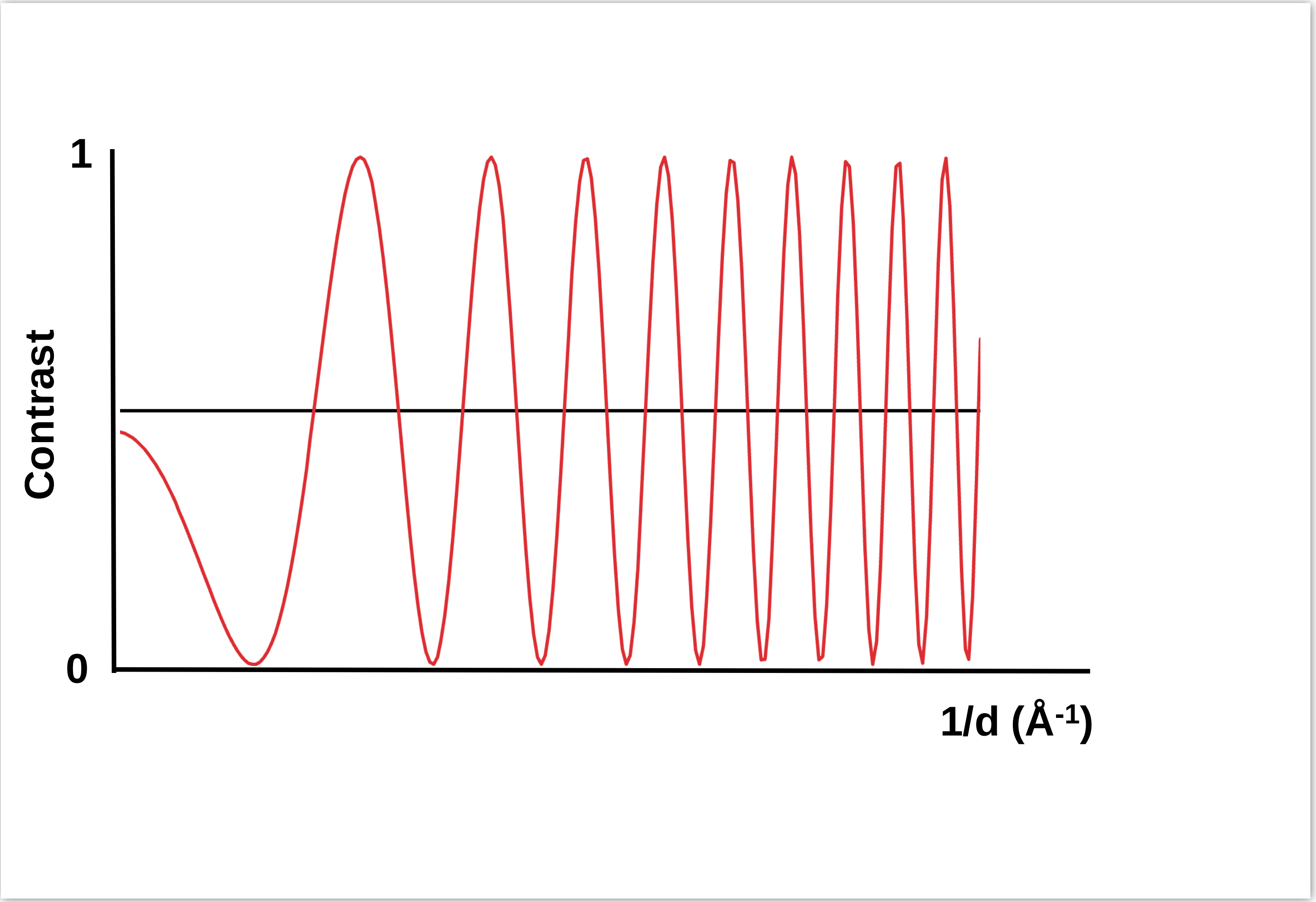

sample
EM
lens
back focal plane
Unscattered beam
Scattered beam
Contrast transferred
to image
Contrast
No transfer of contrast


sample
EM
lens
back focal plane
Unscattered beam
Scattered beam
Contrast transferred
to image
Contrast


sample
EM
Unscattered beam
Scattered beam
Contrast transferred to image
Contrast

Contrast transfer function (CTF)

The phase-contrast transfer function (PCTF) is a function to express what extent the amplitudes converted from the phase changes of the diffracted waves contribute (are transferred) to the TEM image



Expressed as a function of **spatial frequency** (**1/distance**)

Contrast transfer function (CTF)

The phase-contrast transfer function (PCTF) is a function to express what extent the amplitudes converted from the phase changes of the diffracted waves contribute (are transferred) to the TEM image

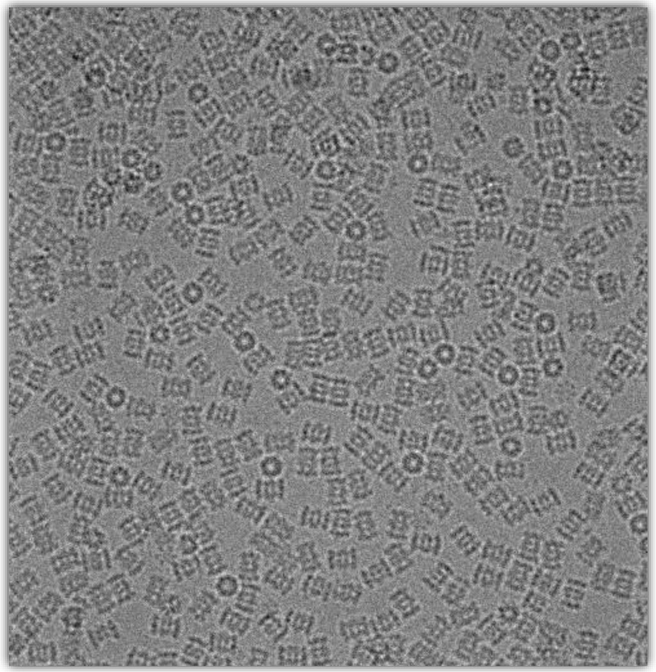
$$\mathrm{CTF}(\vec{s}) = \sqrt{1 - A^2} \cdot \sin\big(\gamma(\vec{s})\big) + A \cdot \cos\big(\gamma(\vec{s})\big)$$

$$\gamma(\vec{s}) = -\frac{\pi}{2} C_s \lambda^3 s^4 + \pi \lambda z(\theta) s^2$$

s = spatial frequency
A = amplitude contrast
C_s = spherical aberration
λ = wavelength of electrons
z(θ) = defocus

CTF is used to "correct" micrographs

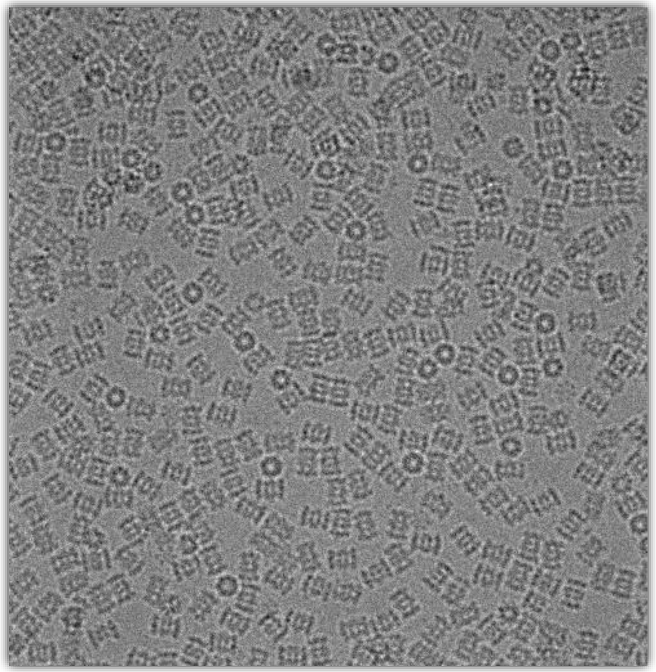
Raw micrograph
(real space)



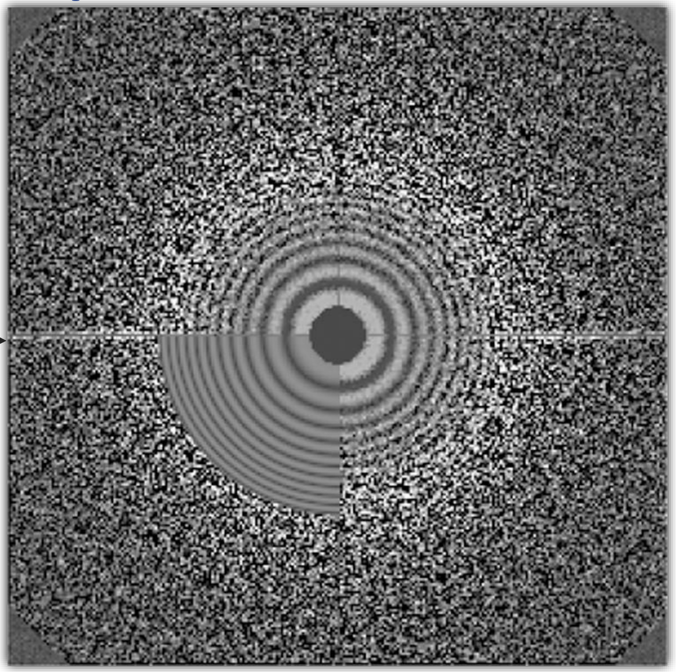
Rohou and Grigorieff, JSB 2015

CTF is used to “correct” micrographs

Raw micrograph
(real space)



FFT

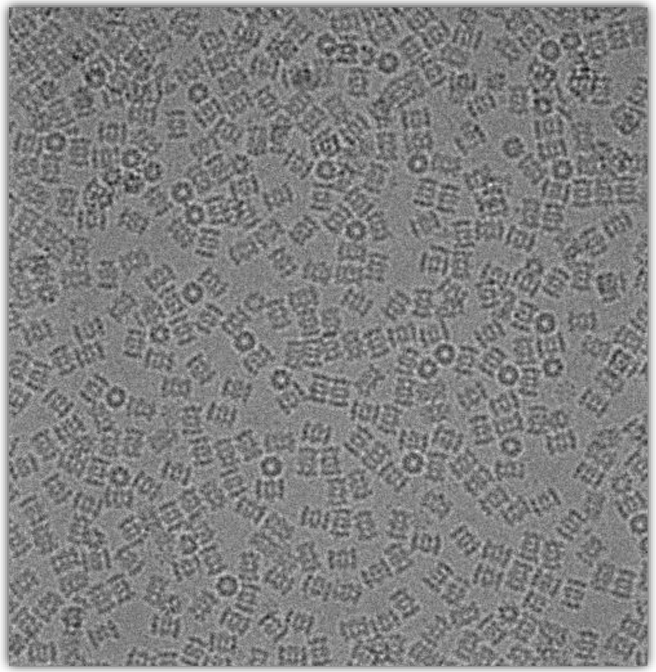


Fourier transform
(power spectrum)

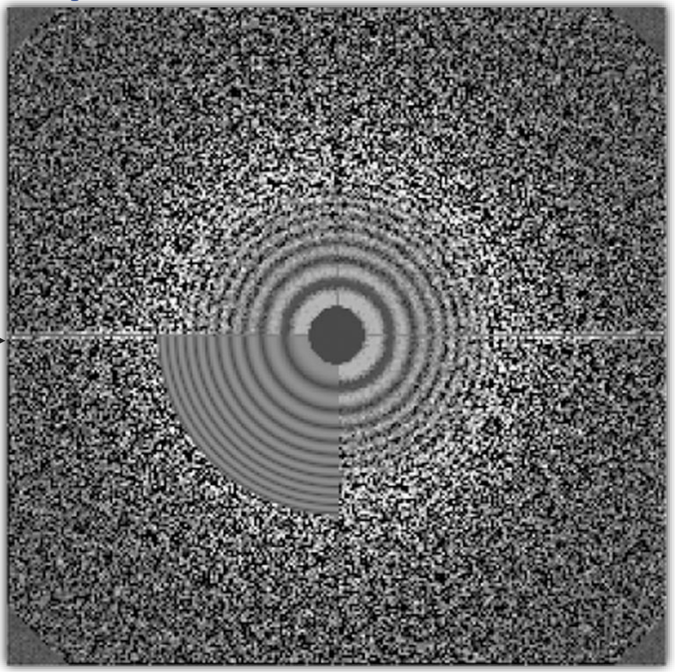
Rohou and Grigorieff, JSB 2015

CTF is used to “correct” micrographs

Raw micrograph (real space)

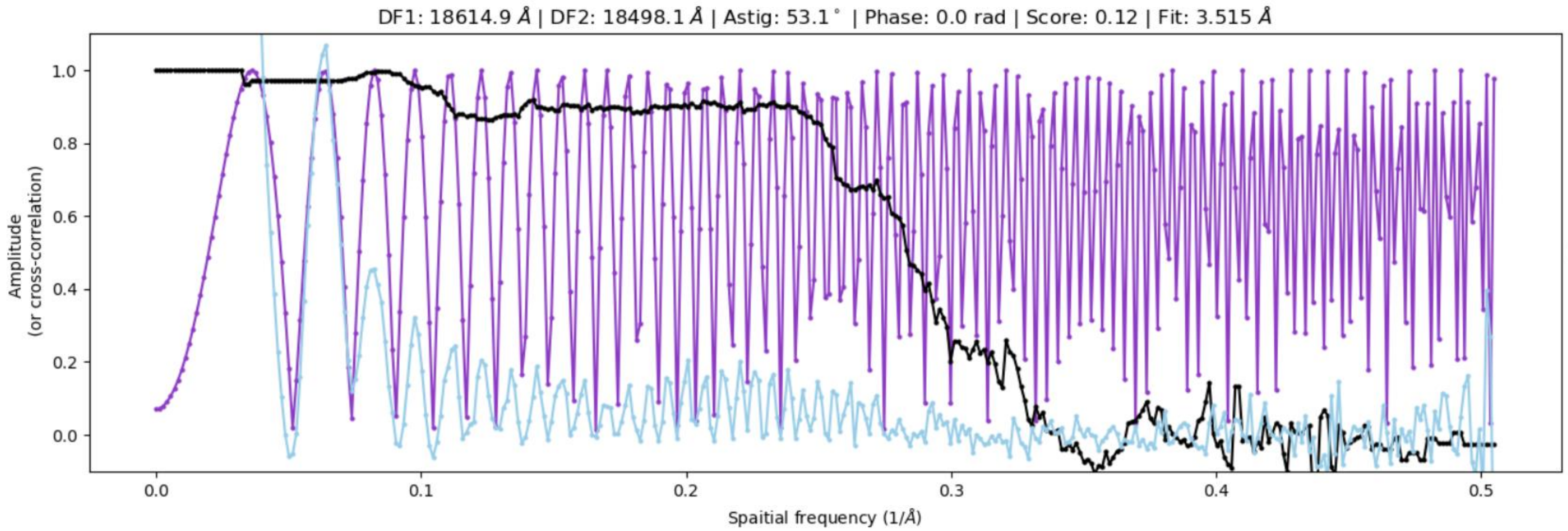


FFT



Fourier transform (power spectrum)

CTF fitting



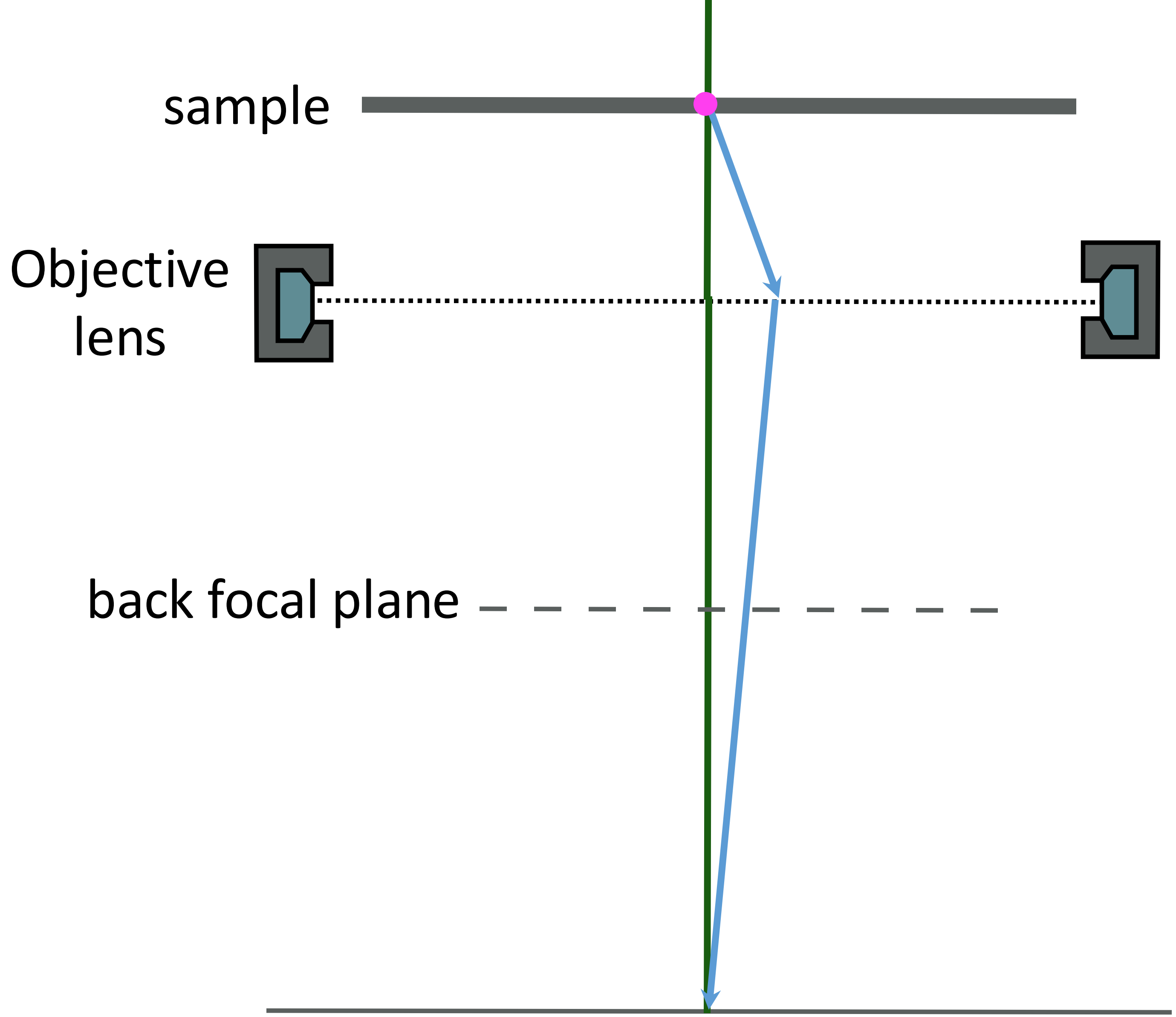
Real data

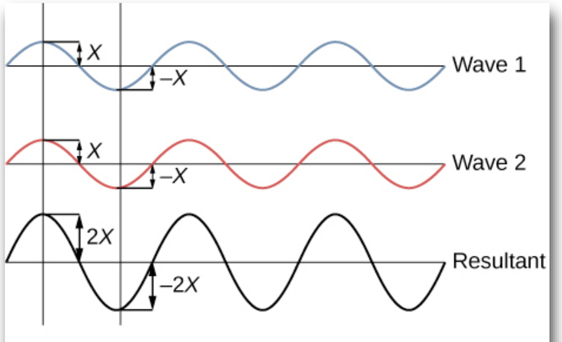
CTF fit

Correlation

Rohou and Grigorieff, JSB 2015

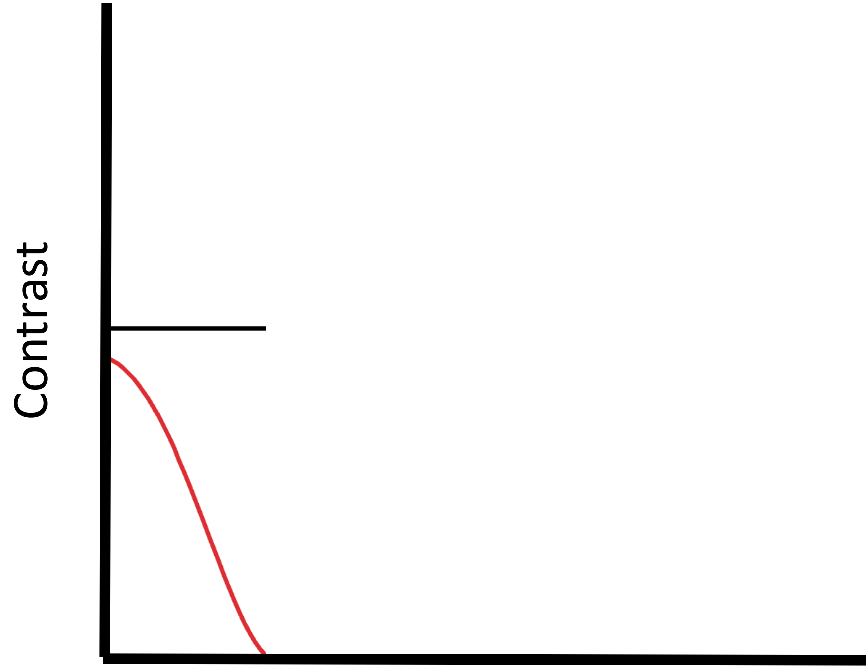
Defocus in EM imaging



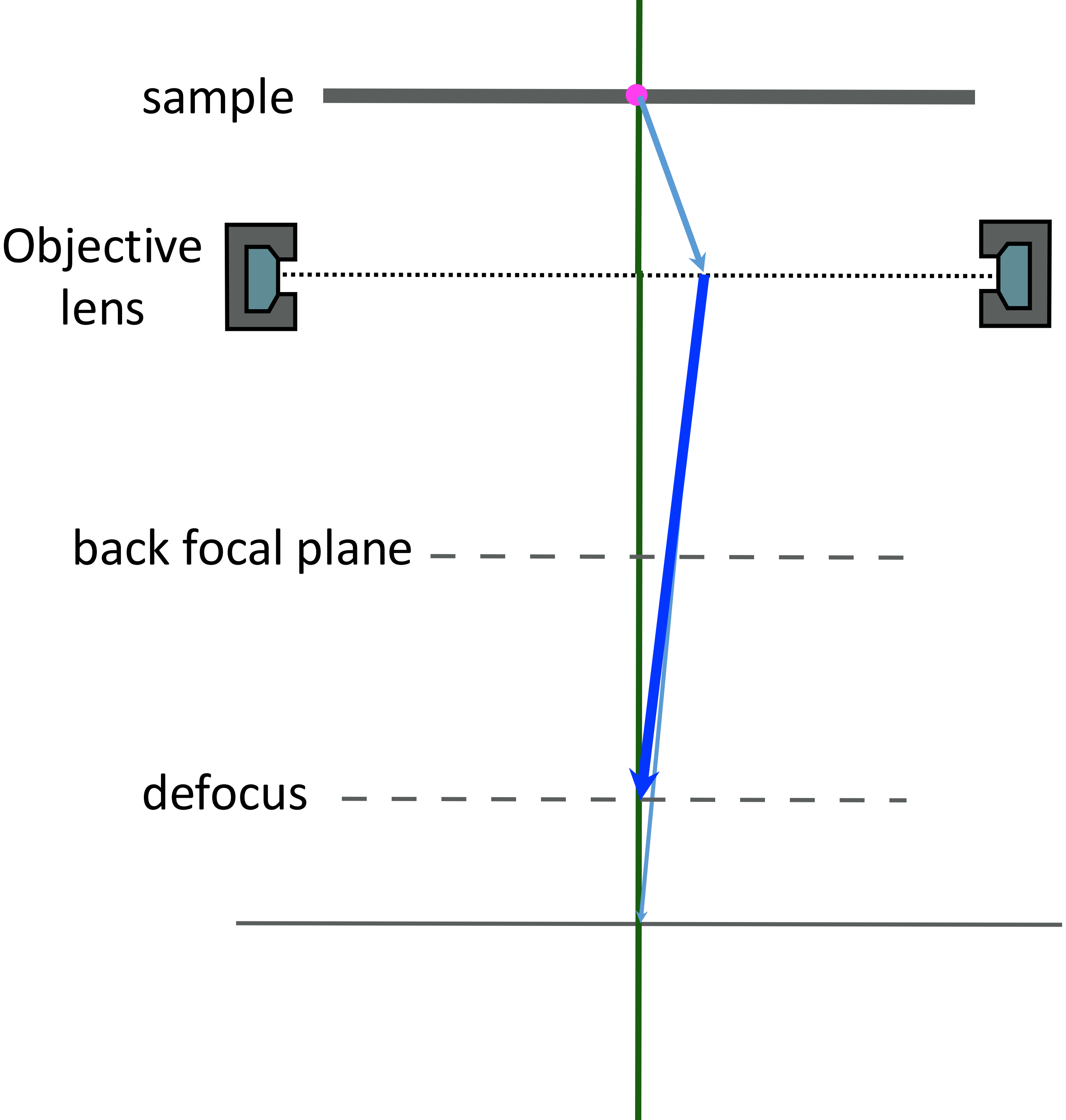


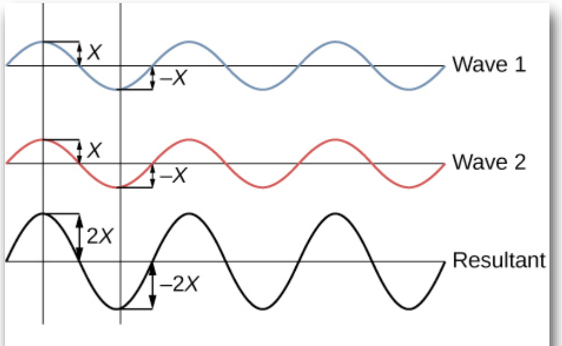
(a) Constructive interference

Contrast transferred to image



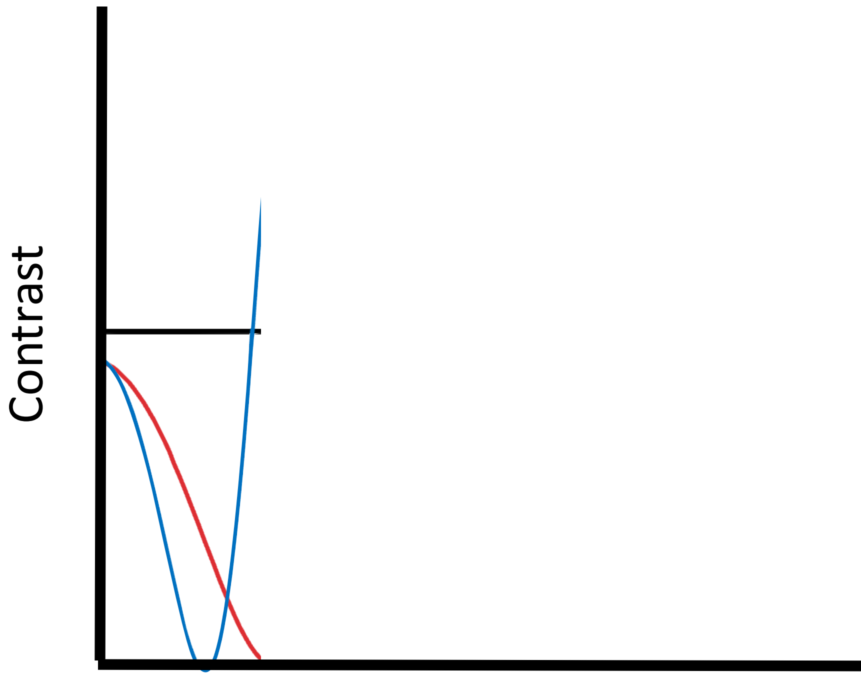
Defocus is controlled by the current in objective lens

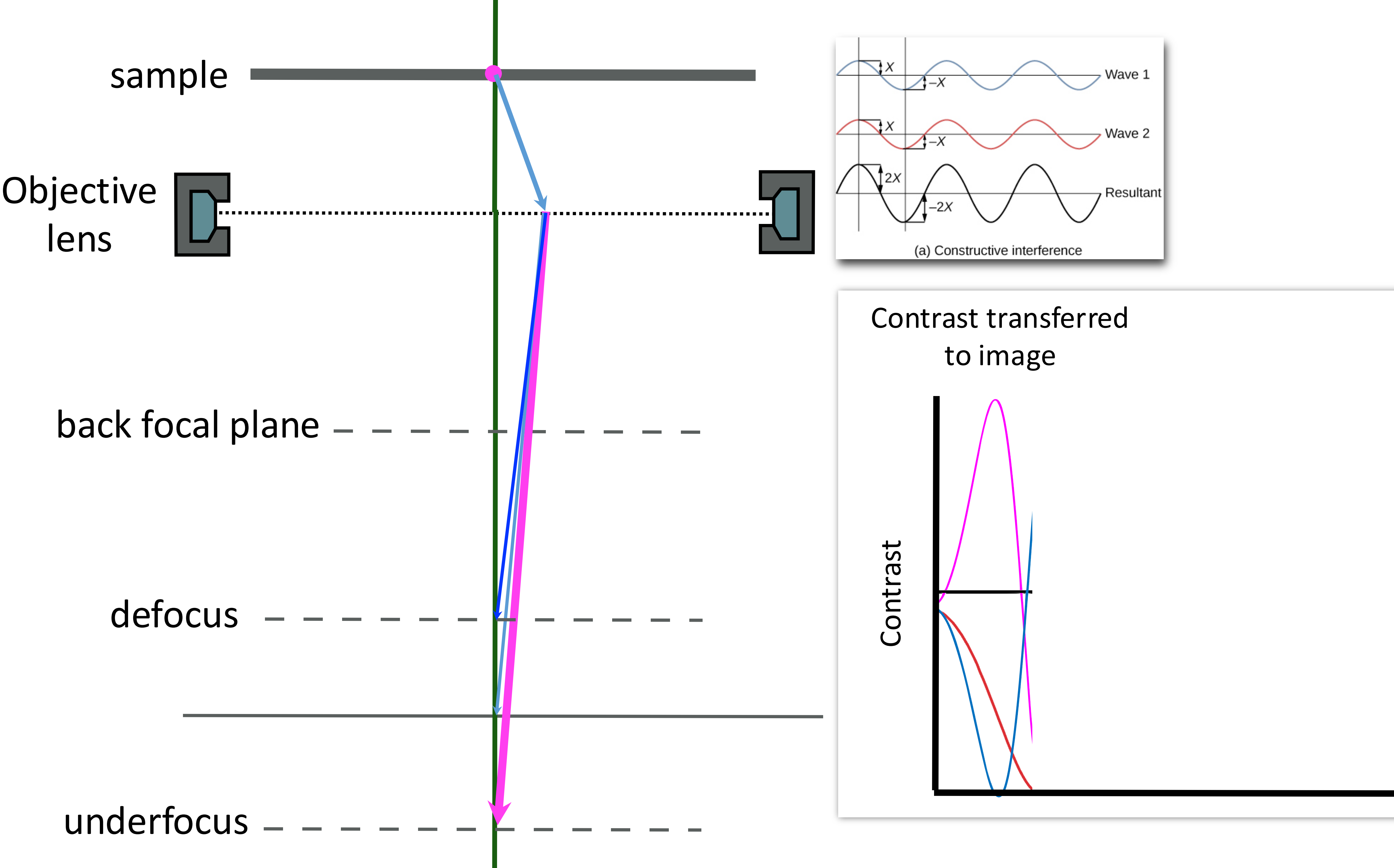




(a) Constructive interference

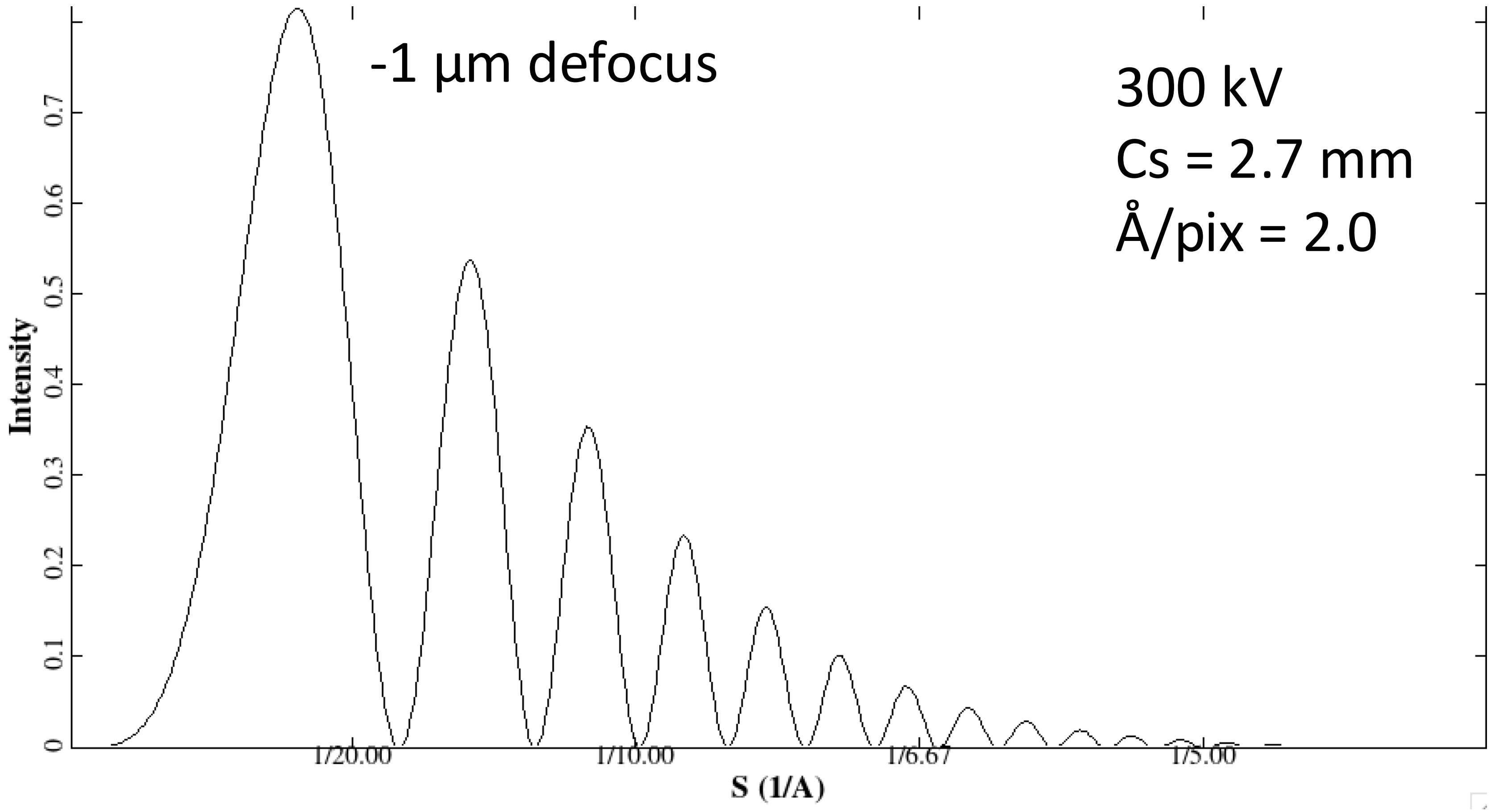
Contrast transferred to image



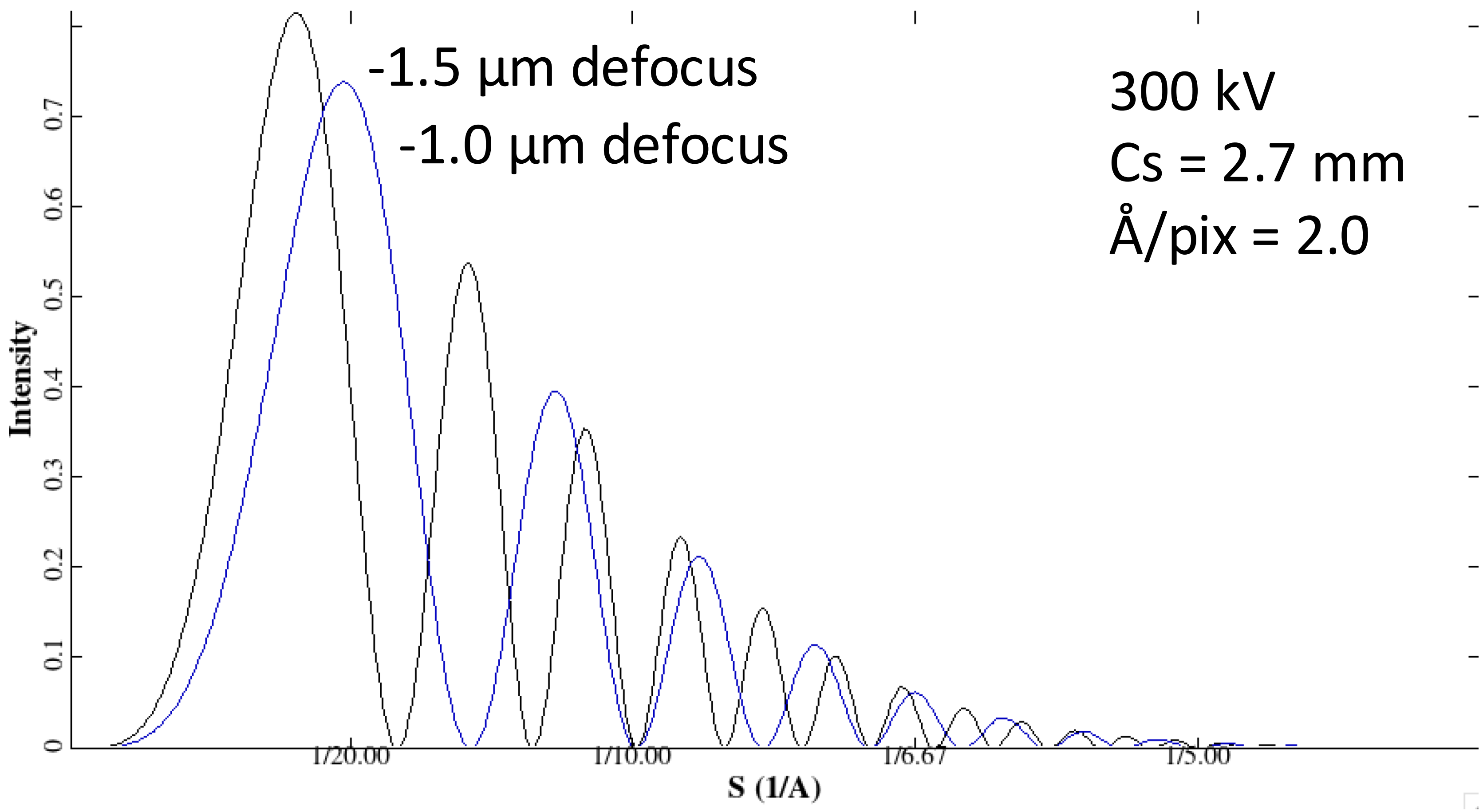


sample
Objective
lens
back focal plane
defocus
underfocus
X
−X
Wave 1
X
−X
Wave 2
2X
−2X
Resultant
(a) Constructive interference
Contrast transferred
to image
Contrast

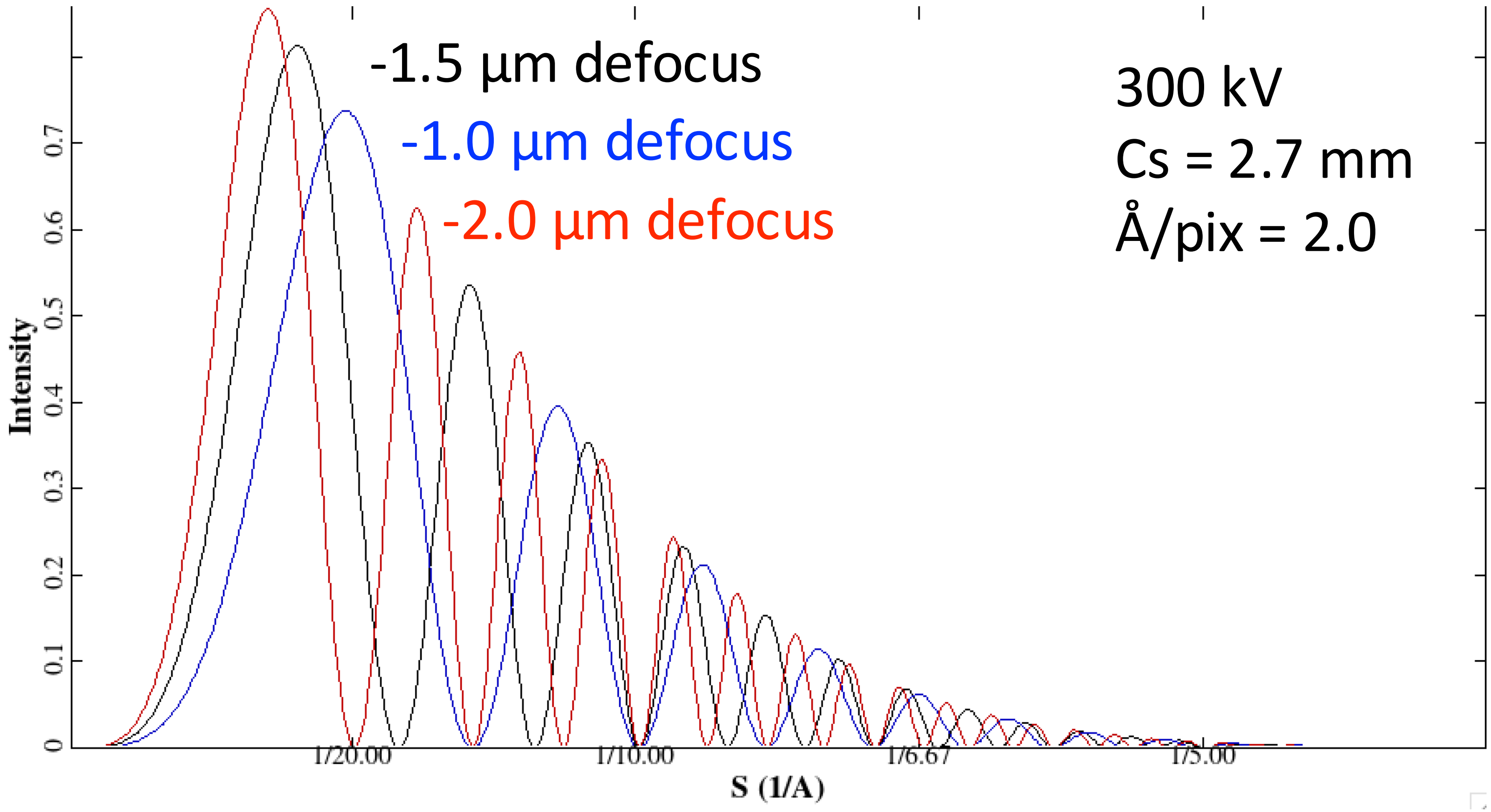
Defocus & Data Acquisition



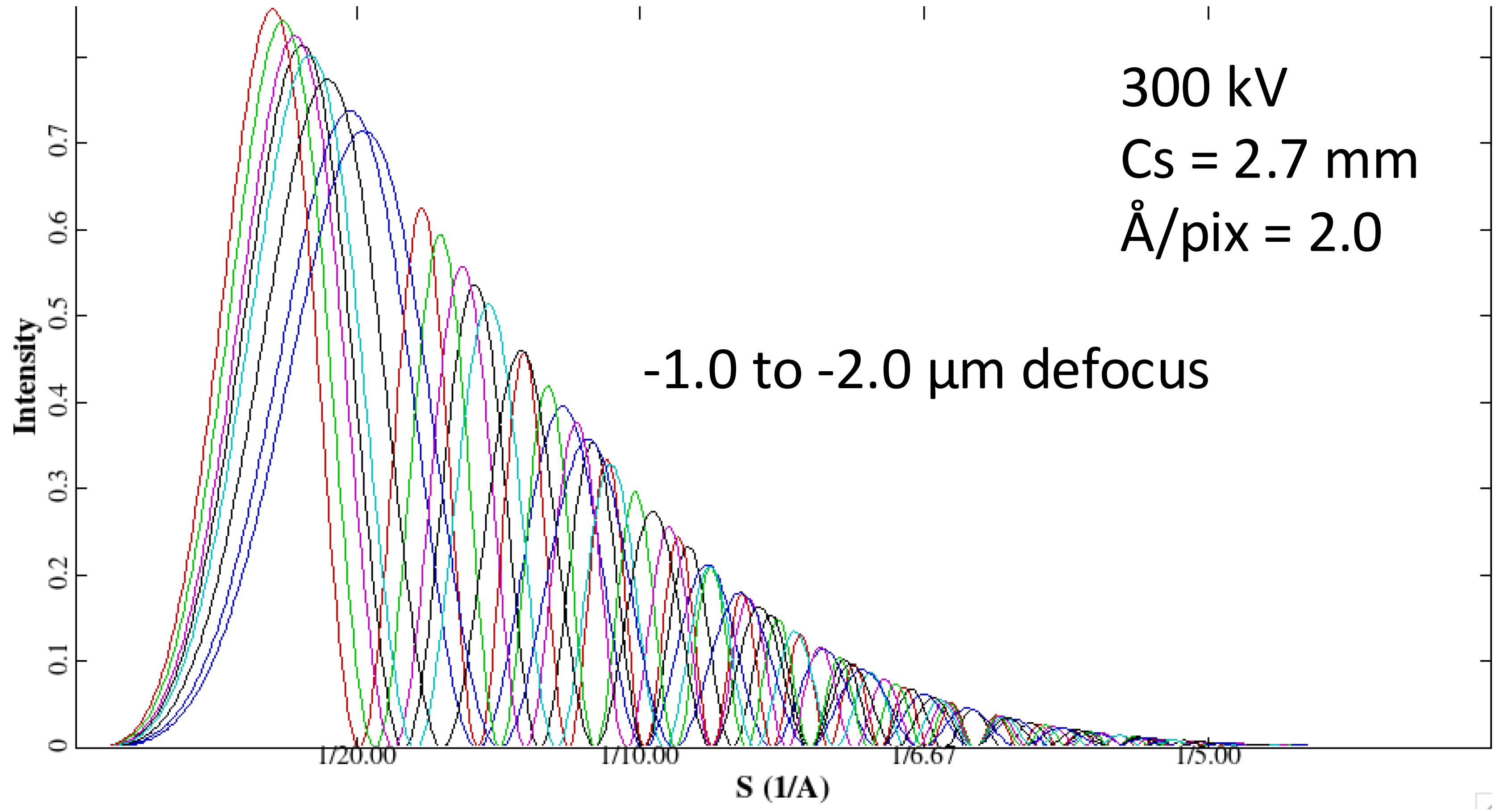
Defocus & Data Acquisition



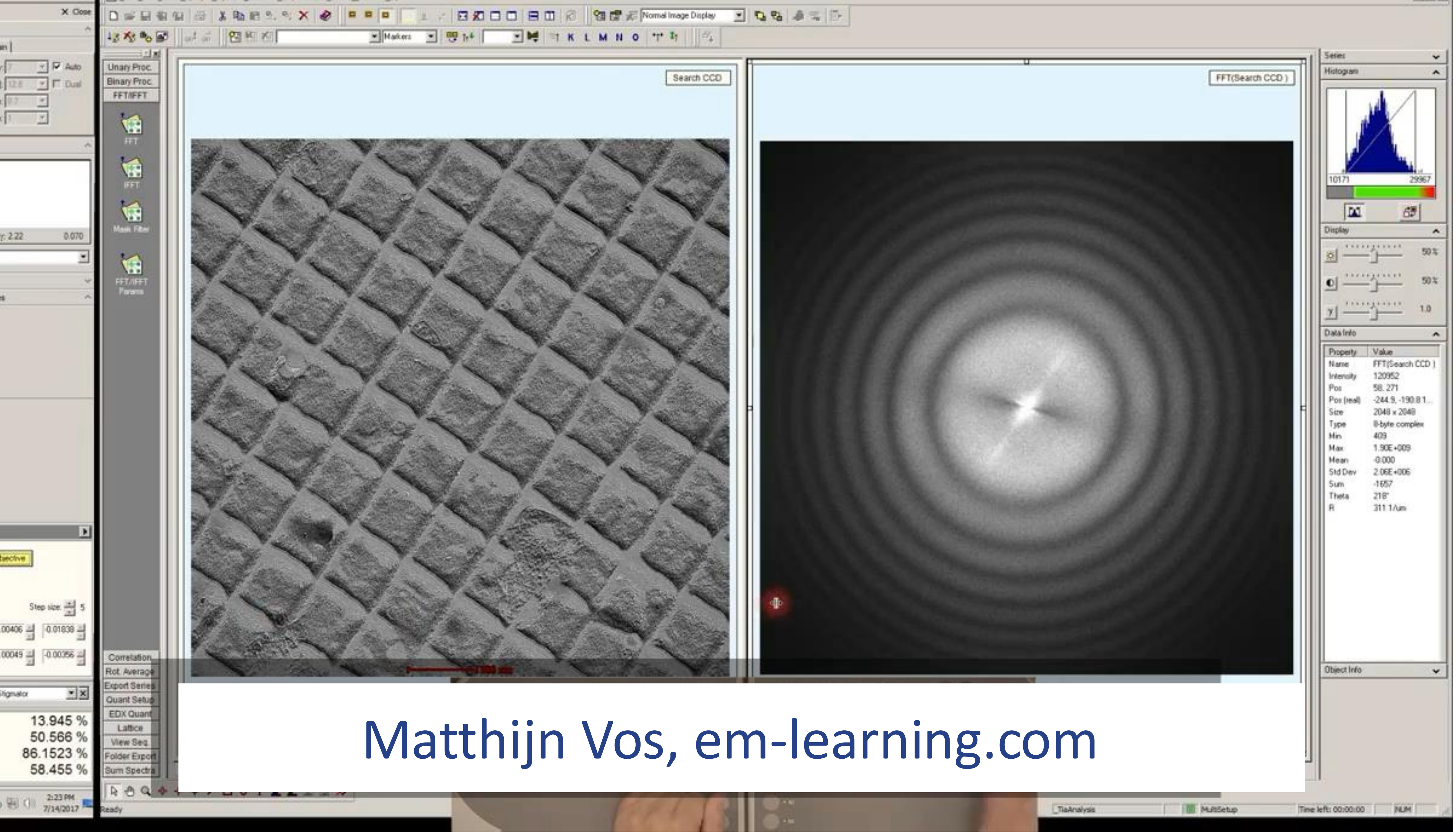
Defocus & Data Acquisition



Defocus & Data Acquisition



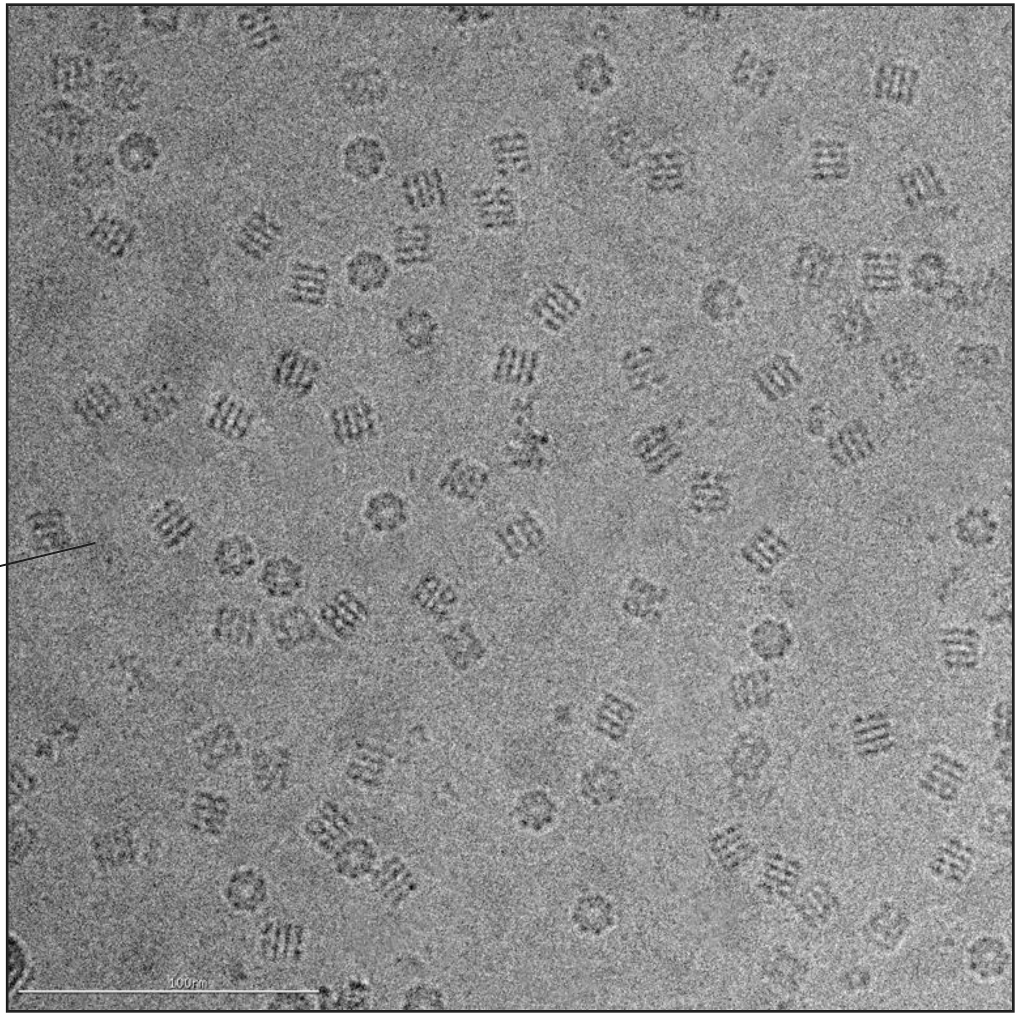
- Varying defocus values within a range of 1-2µm during imaging allows to collect micrographs with different CTF parameters and processing such data assures contrast transfer across all resolution shells for the particle projection images



Matthijn Vos, em-learning.com

Intro to EM data processing

cryoEM-grid imaging



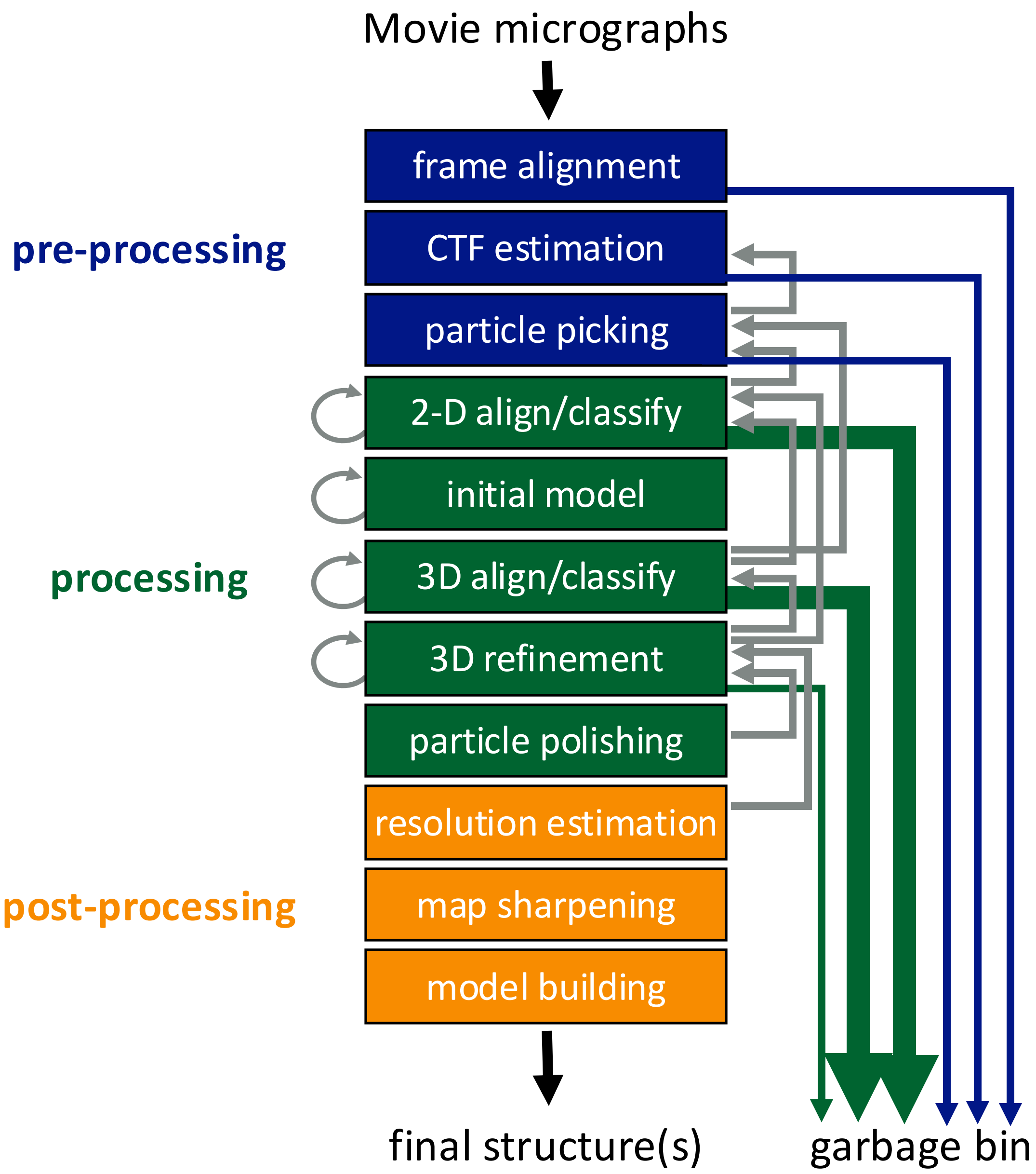
Microscope settings:

- Select optimal magnification
- Adjust microscope alignments
- Typical defocus range: -0.5 to -3µm
- Typical e^- dose: ~20-60 e^-/Å2

Data collection:

- Collecting couple thousand movies
- Assuring good contrast in images
- Assuring optimal particle density
- Inspecting particle orientations
- Assessing particle heterogeneity

Most facilities have staff to help with microscope alignment and imaging



many of the downstream steps feed back to earlier steps (not a linear path)

in silico pruning: images can be rejected at multiple points

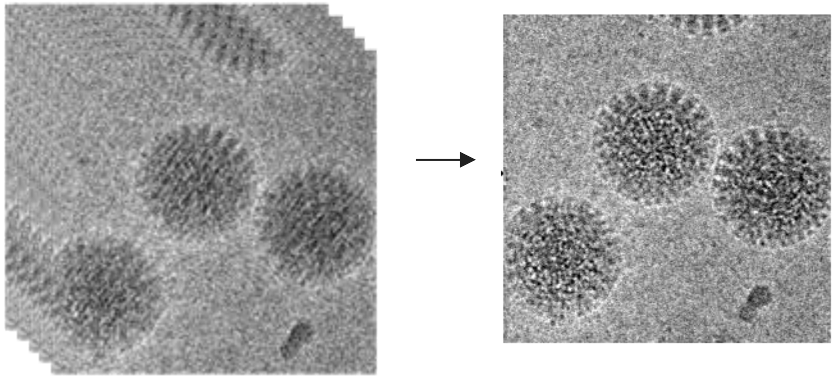
Correct for motion, radiation damage, scope & camera defects

- **frame alignment**
- CTF estimation
- particle picking
- 2-D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- resolution estimation
- map sharpening
- model building

poor alignment
excessive drift

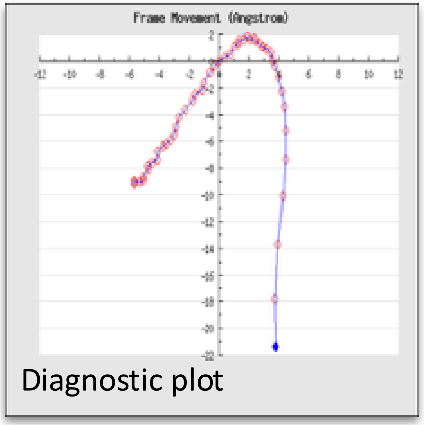
General concept

Frames



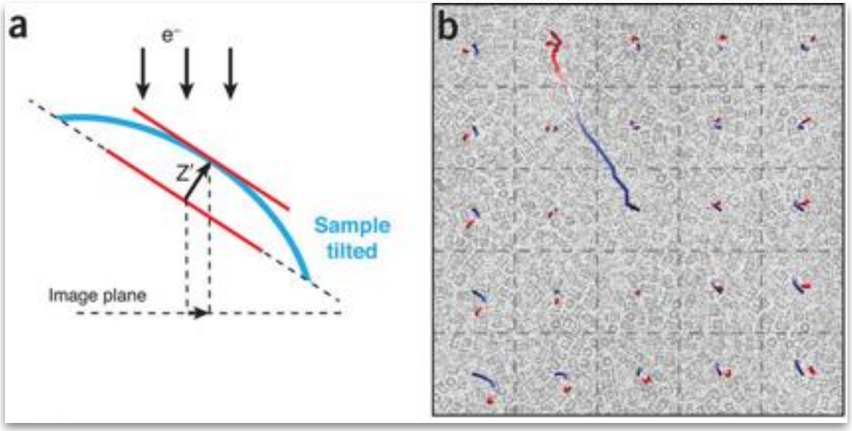
Corrected micrograph

Motioncorr - per frame



Li et al. Nat. Methods 2013

MotionCor2 - model local deformation



Zheng et al. Nat. Methods 2017

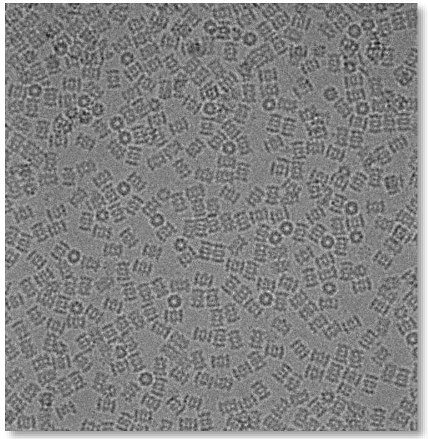
output: aligned, dose-weighted, mag-corrected micrographs

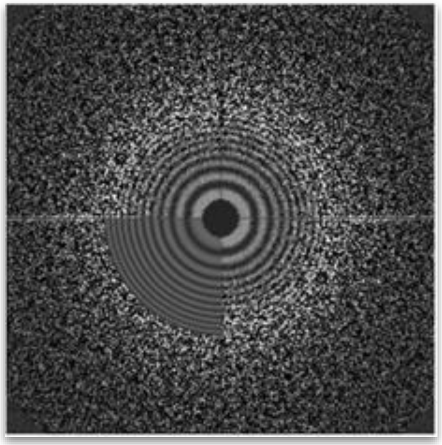
Determine CTF parameters for correction during processing

- frame alignment
- **CTF estimation**
- particle picking
- 2-D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- resolution estimation
- map sharpening
- model building

poor determination
low micrograph resolution
defocus out of range

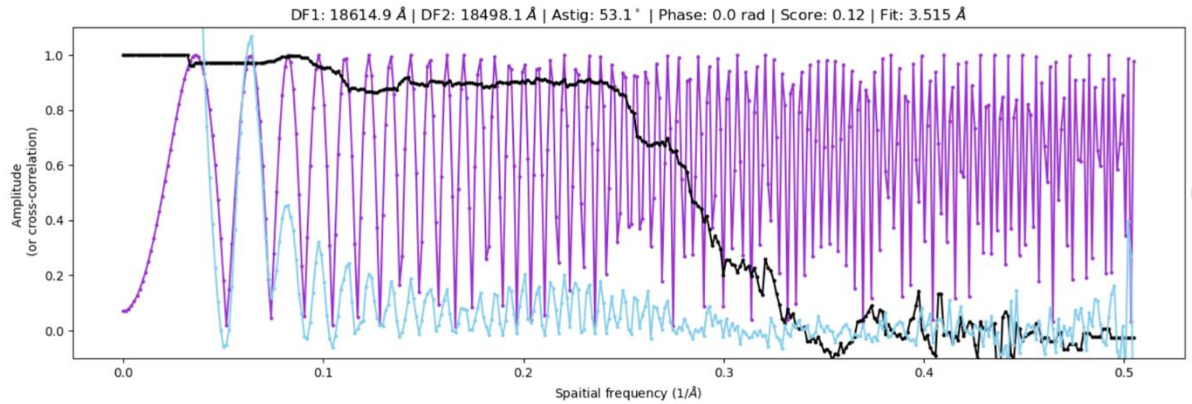
Parameters:
Defocus,
astigmatism
(phase shift)





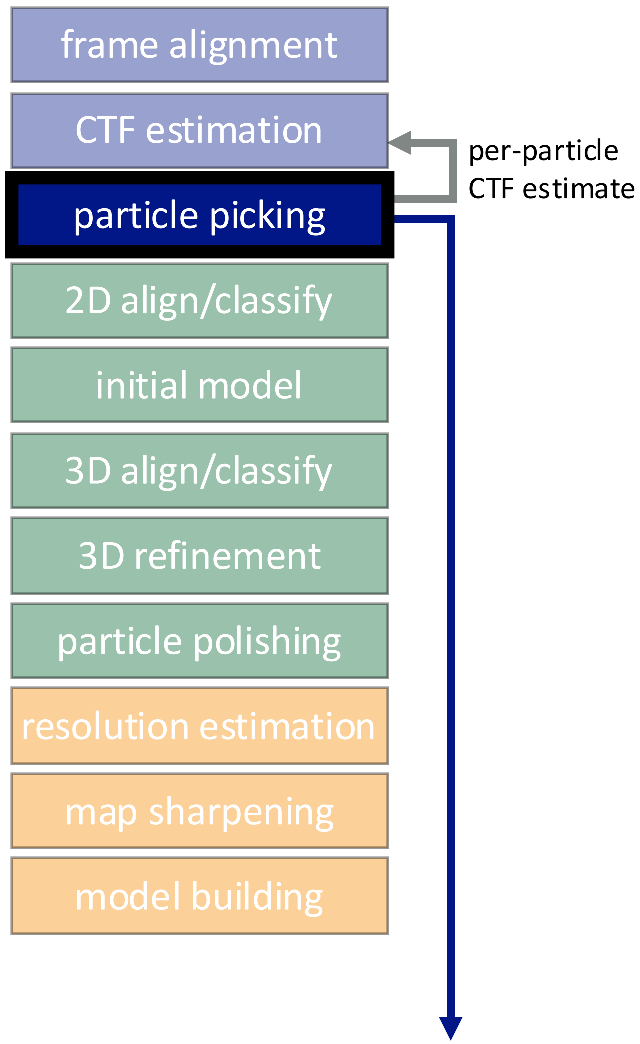
Rohou & Grigorieff. J Struct Biol. 2015

assess image quality



output: CTF parameters for each micrograph

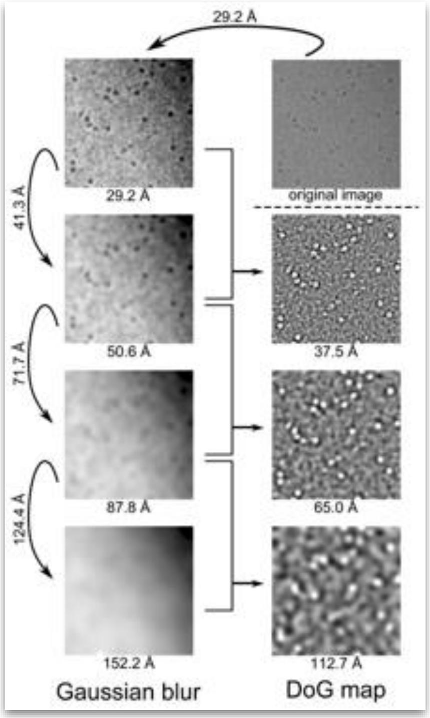
Identify and extract every particle in the dataset



thresholding
manual evaluation

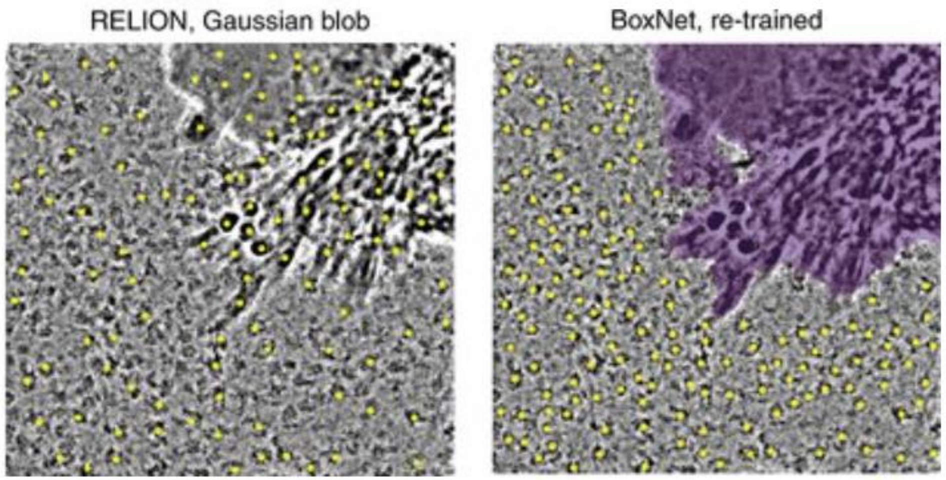
- Manual picking
- Feature-based (e.g., blob, ring)
- Template-based (2D averages)
- Machine learning

Gaussian blob picker



Voss J., Struct. Biol. 2009

Deep learning based



- Less selective
- More bad picks

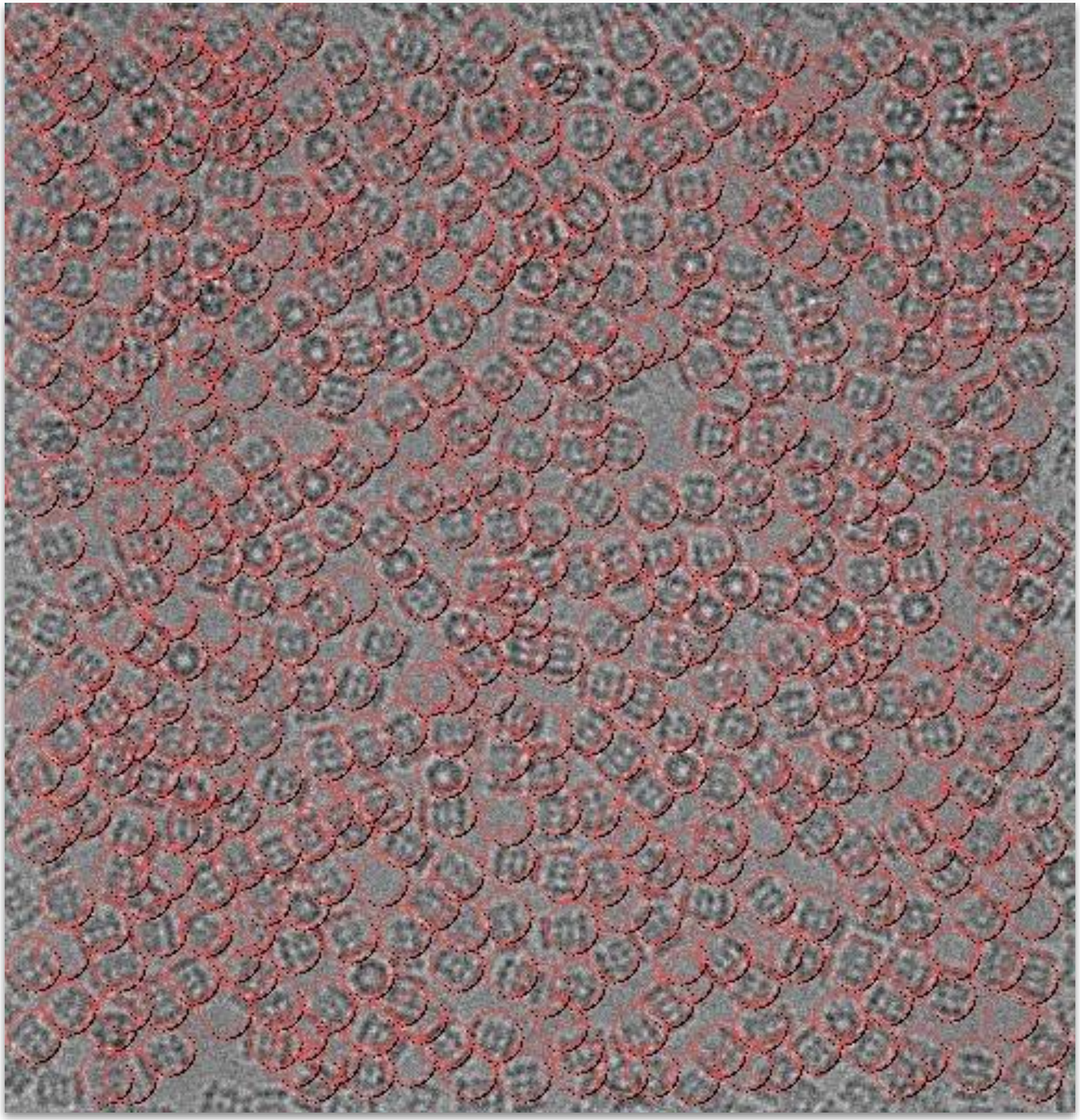
- More selective
- Computationally heavy

Tegunov D., Nature Methods 2019

output: particle images extracted to stacks

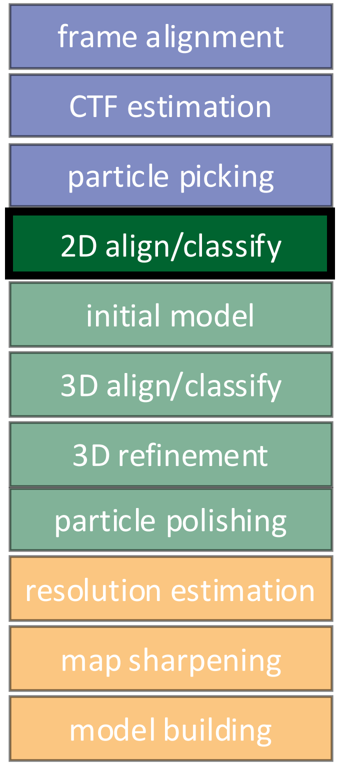
Many, many particles to boost signal to noise

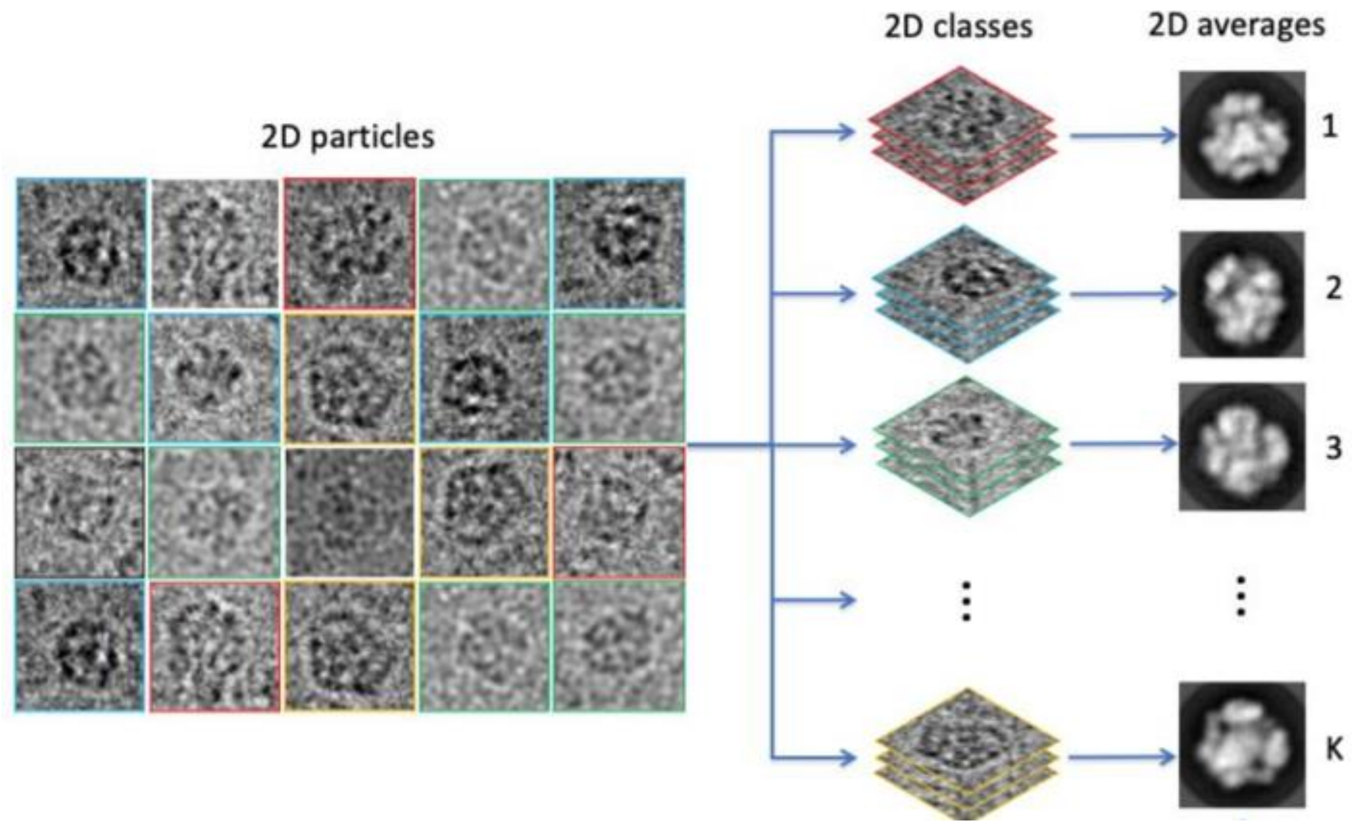
- Target particle number depends on the purpose (e.g., screening or collection) and desired resolution
- In simpler cases ~10k particles are sufficient to achieve near-atomic resolution (~3-4Å)
- Well-behaved samples >100kDa, featuring high homogeneity and high symmetry are very optimal for analysis by cryoEM
- In more complex (e.g., heterogeneous) cases you may need many million particles and extensive sorting to achieve desired resolution.
- Sometimes it is necessary to go back and collect more data



Use 2D classification to assess quality of particles

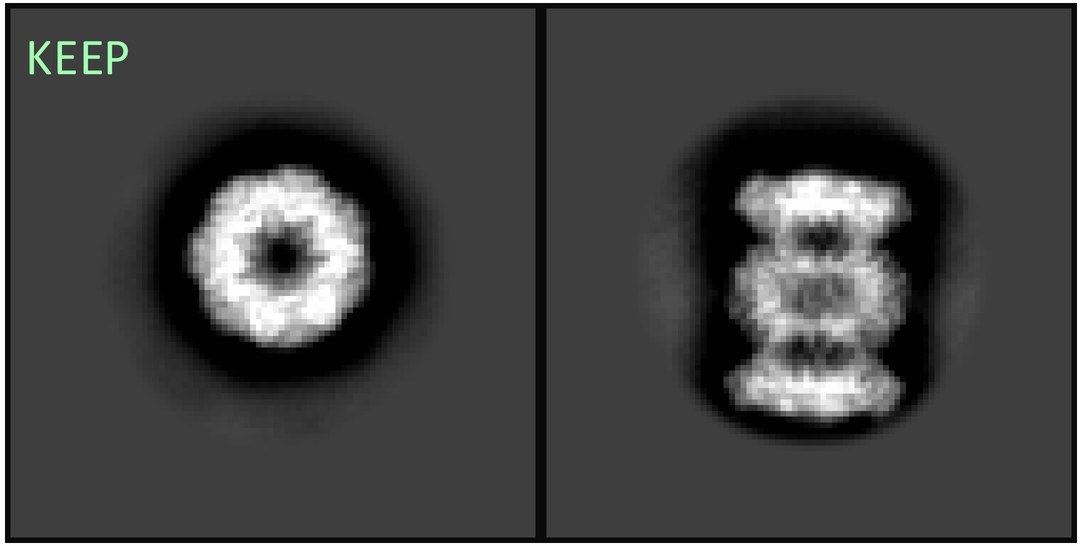
- Defined subregions around each particle are extracted from micrographs and classified using 2D alignment algorithms

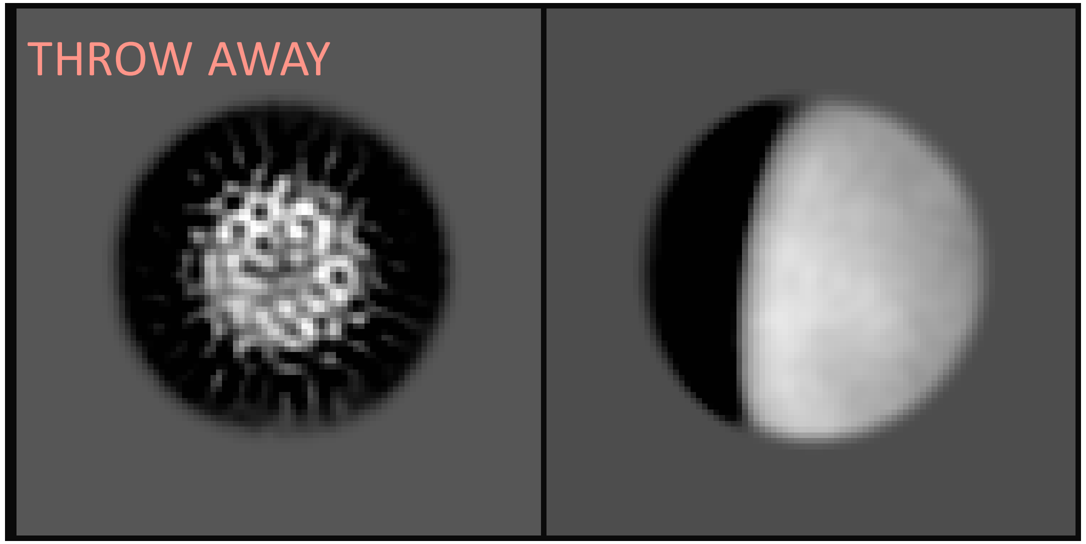




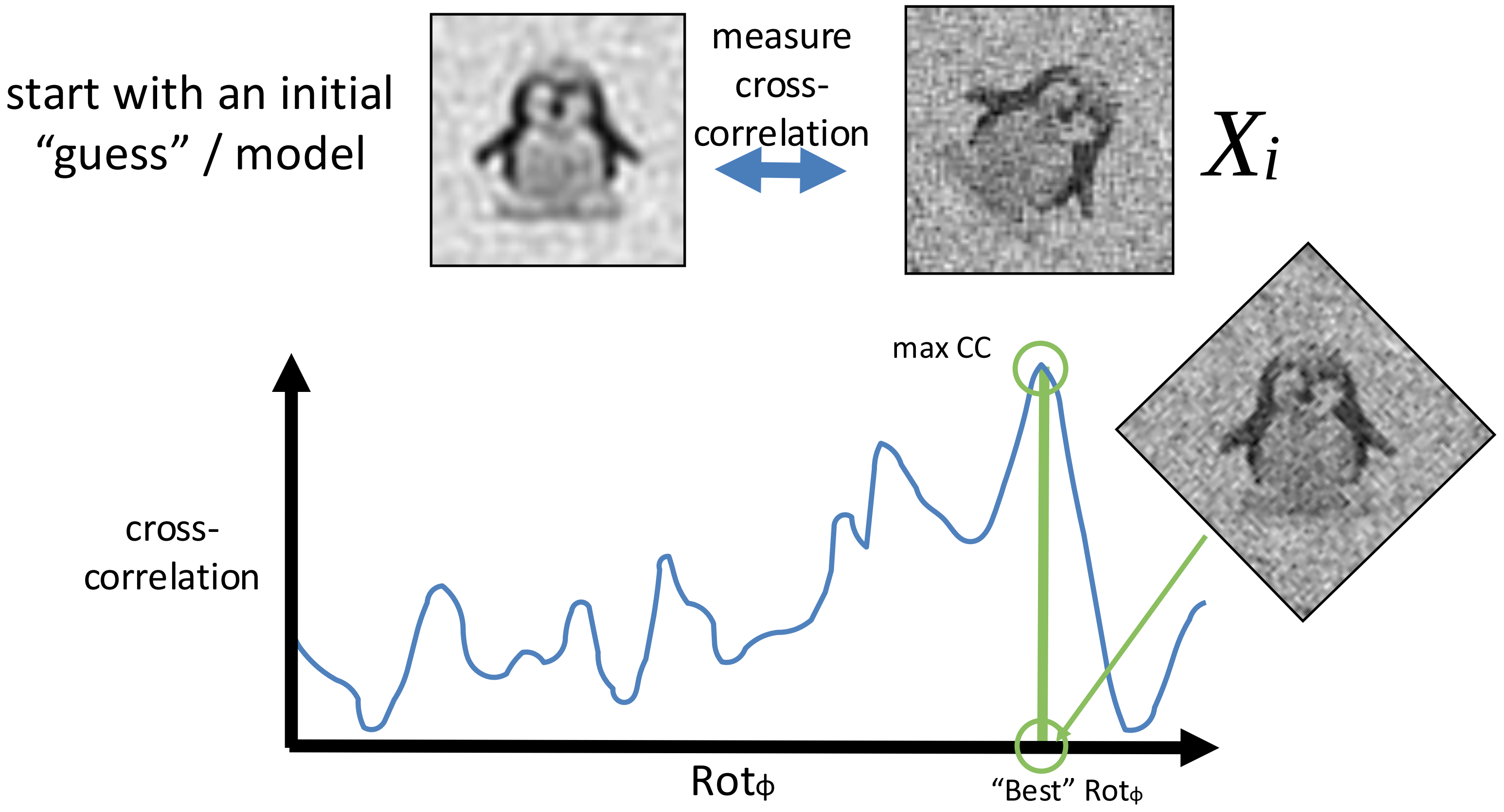
Use 2D classification to assess quality of particles

- frame alignment
- CTF estimation
- particle picking
- **2D align/classify**
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- resolution estimation
- map sharpening
- model building



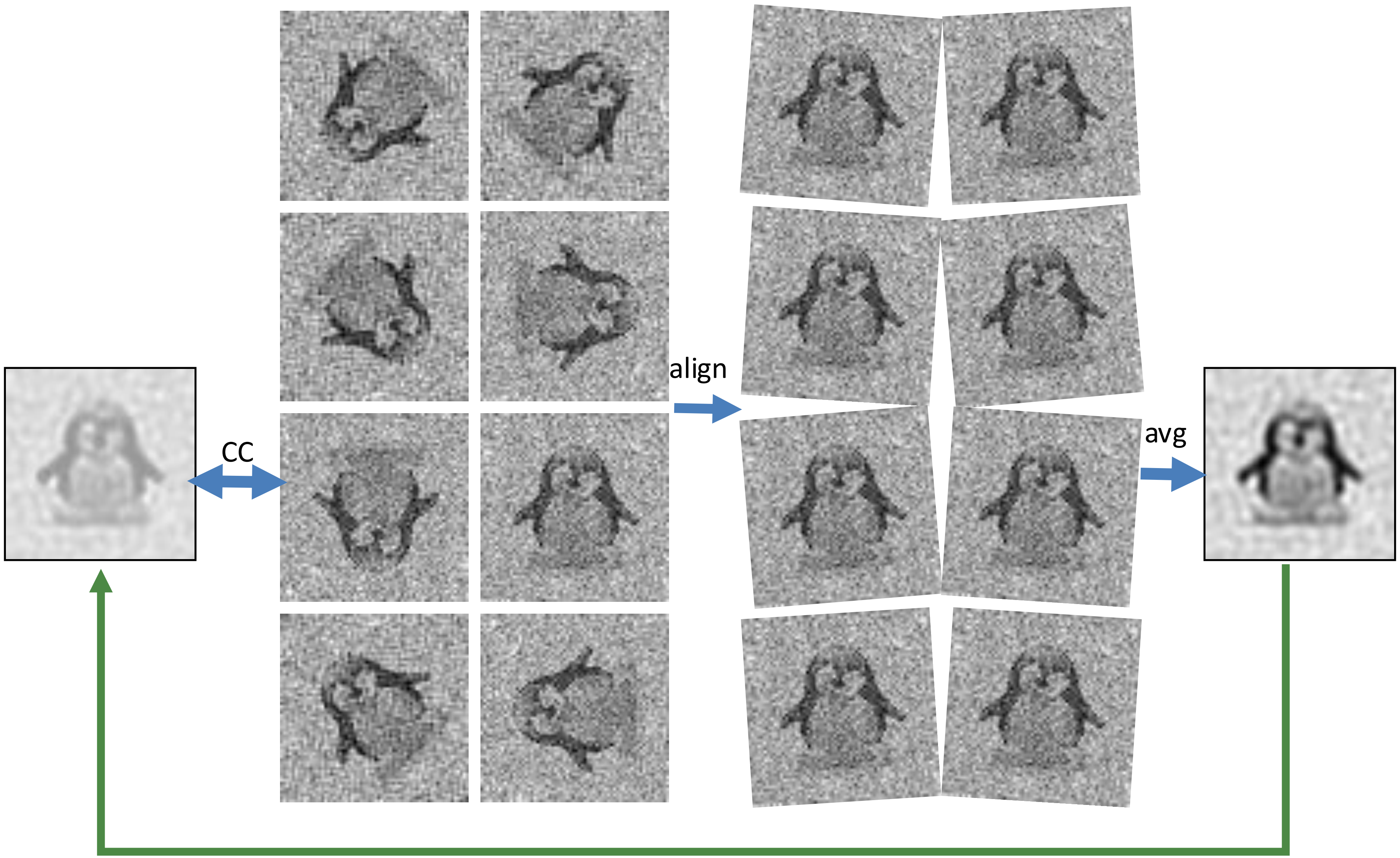


The maximum cross-correlation / least squares approach

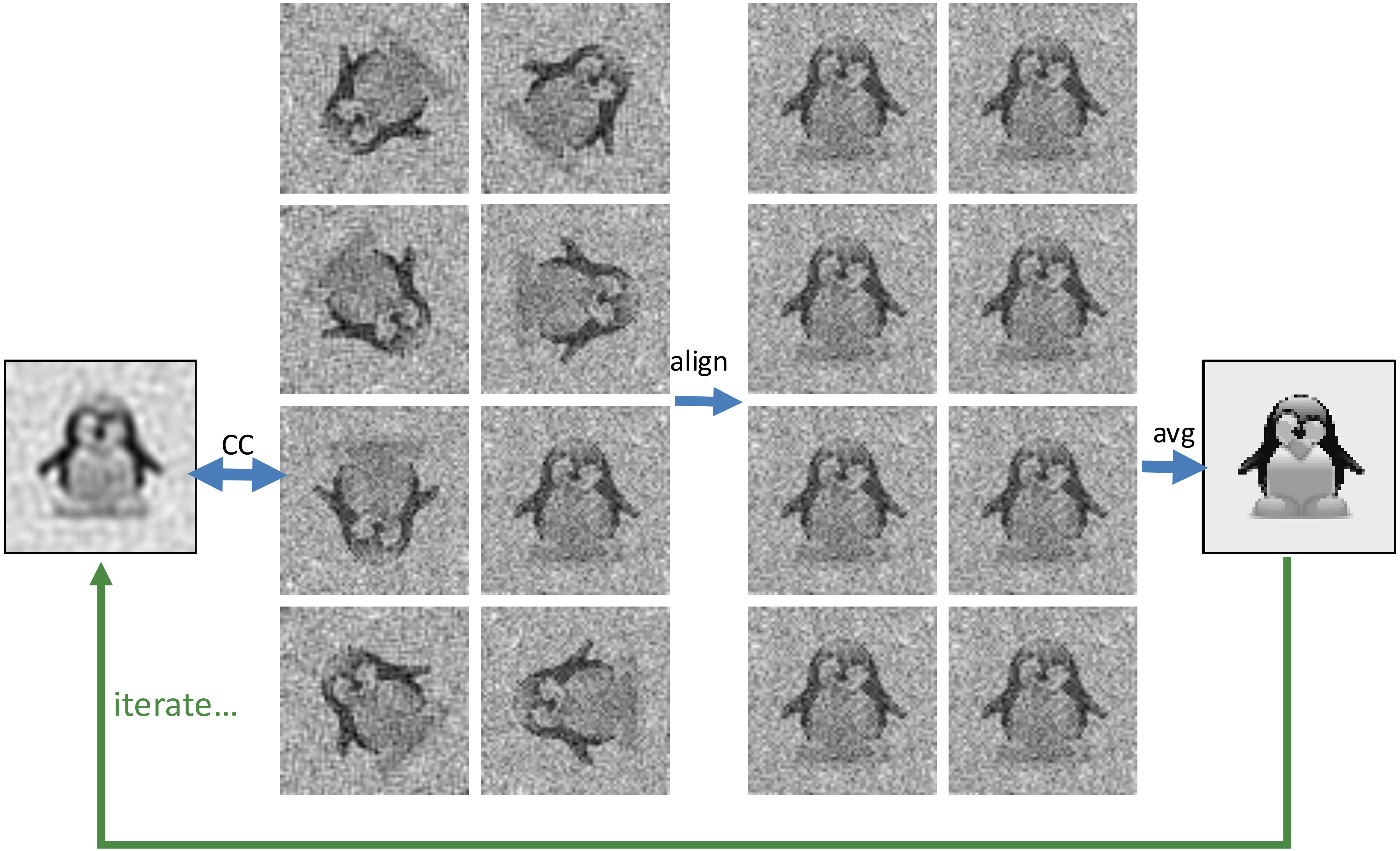


Sigworth, Doerschuk, Carazo, Scheres. "An introduction to maximum-likelihood methods in cryo-EM"
Methods in Enzymology 2010

Align & average…. iterate

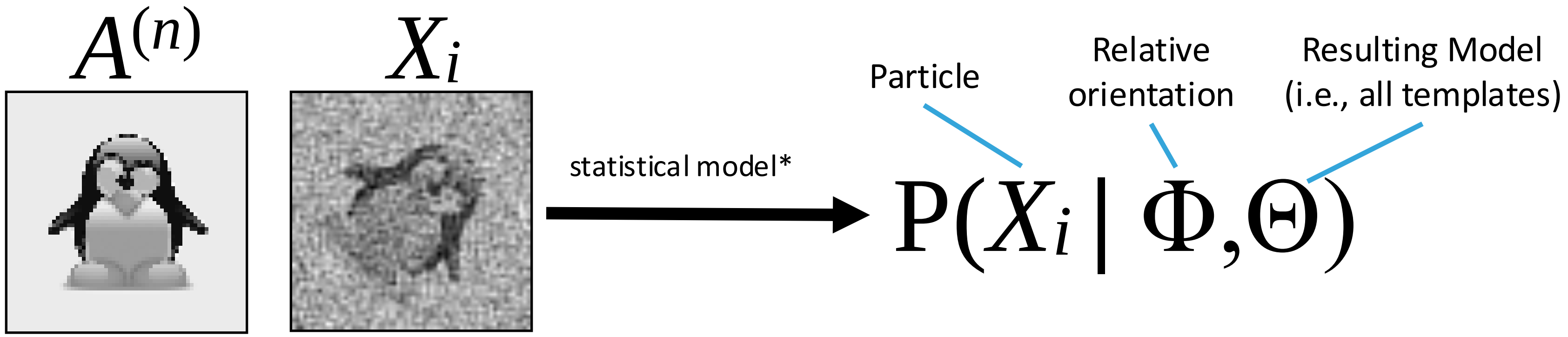


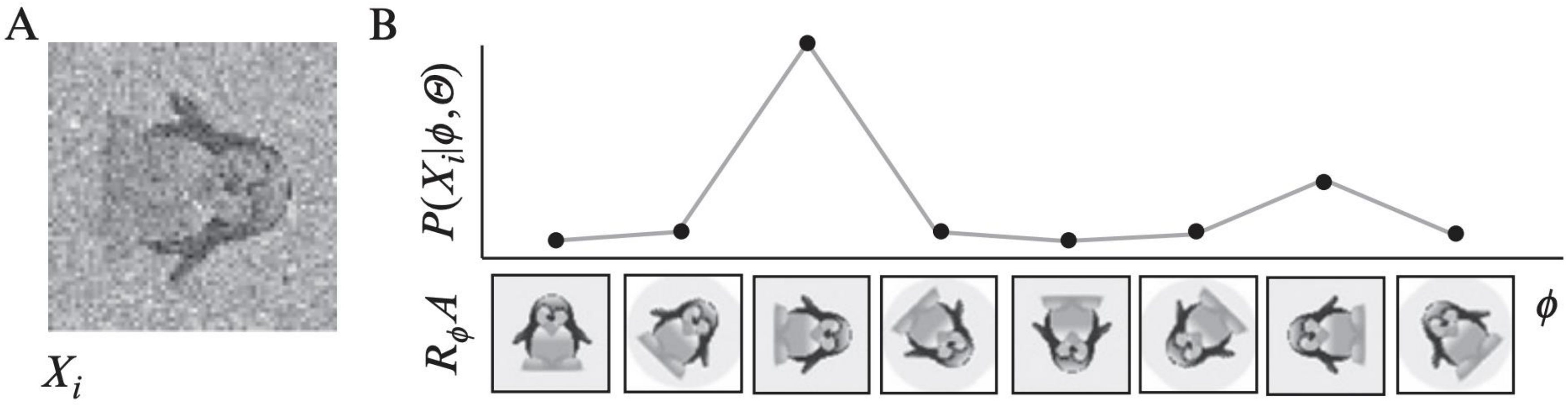
Align & average…. iterate



Maximum Likelihood – A Bayesian approach

- Considers the likelihood of the current model is correct given the data
- Therefore, must be able to assign a probability that a projection describes an image.
- The program searches for a set of 2D classes and particle positions with **combined highest likelihood**

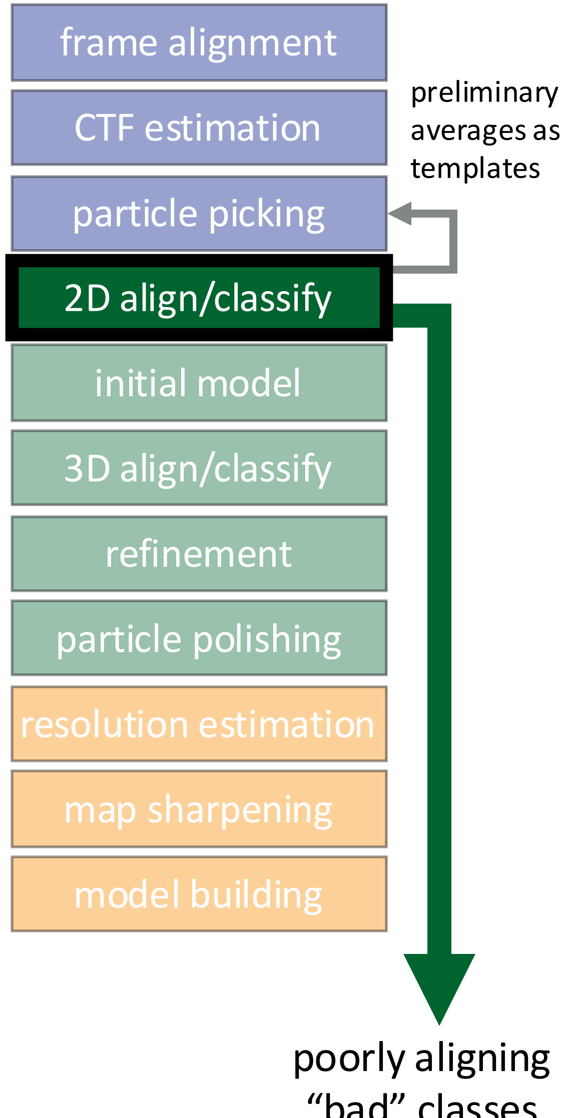


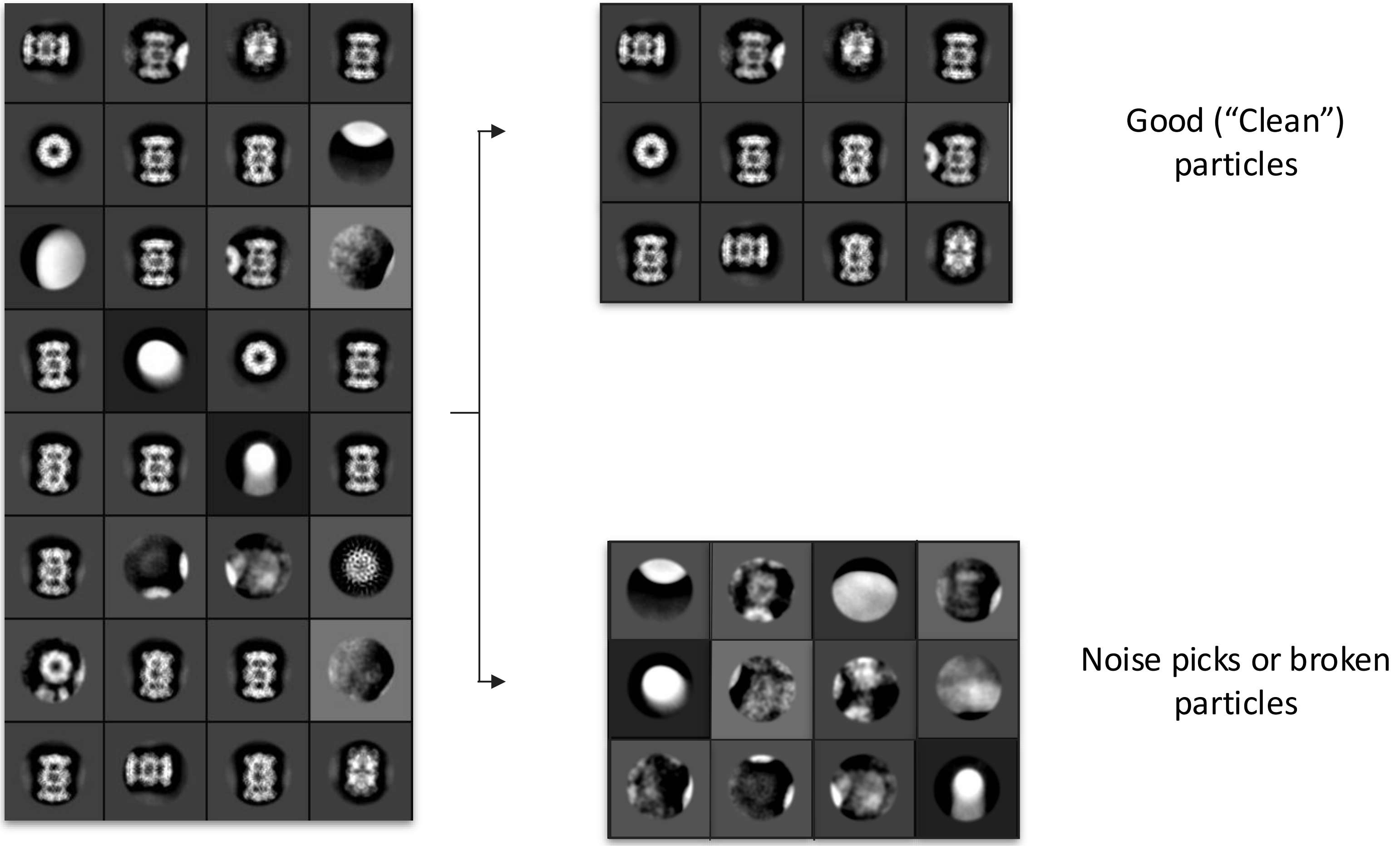


Sigworth, Doerschuk, Carazo, Scheres. “An introduction to maximum-likelihood methods in cryo-EM” Methods in Enzymology 2010

What is 2D classification used for?

Identify poorly-behaved "particles", compositional/conformational heterogeneity

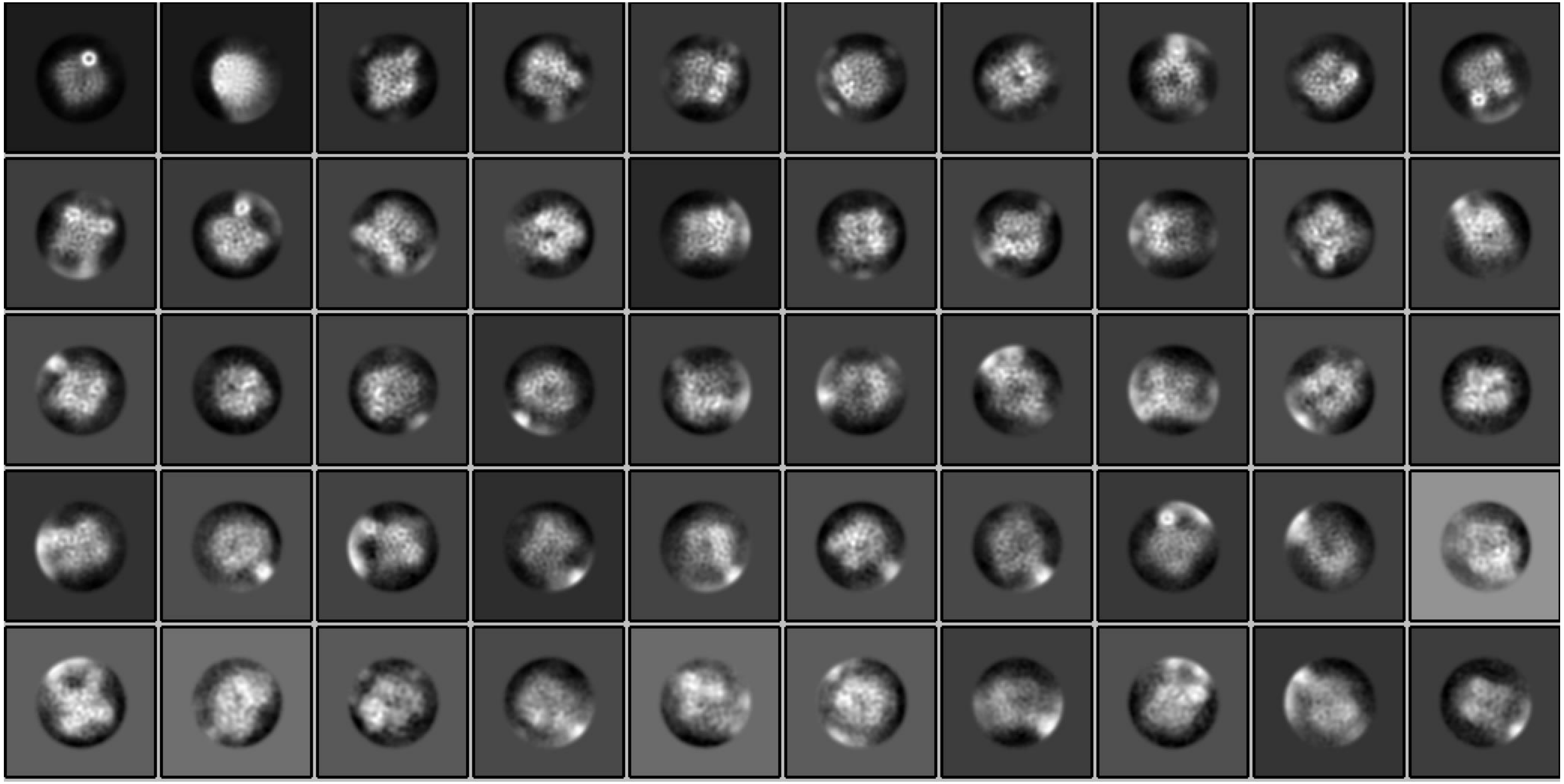




output: aligned class averages, "cleaned"dataset

What is 2D classification used for?

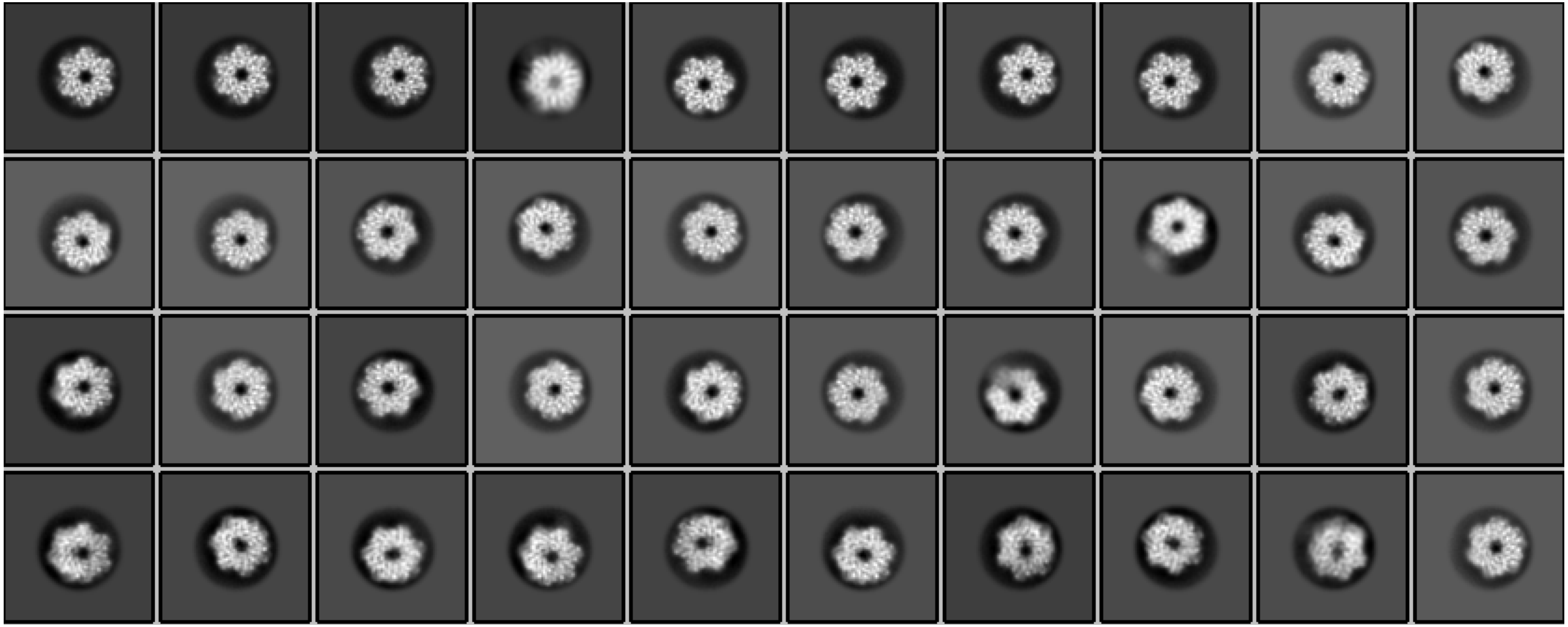
Diagnosing problematic datasets: **Excessive noise in the data**



What is 2D classification used for?

Diagnosing problematic datasets: **Preferential particle orientation problem**

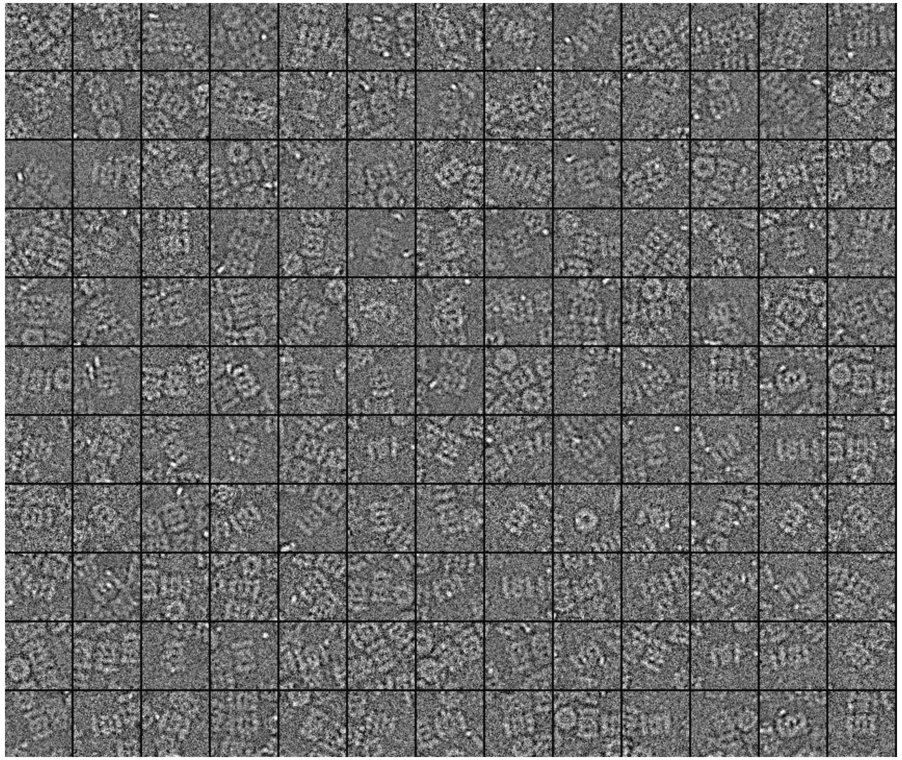
1-2 Dominant views



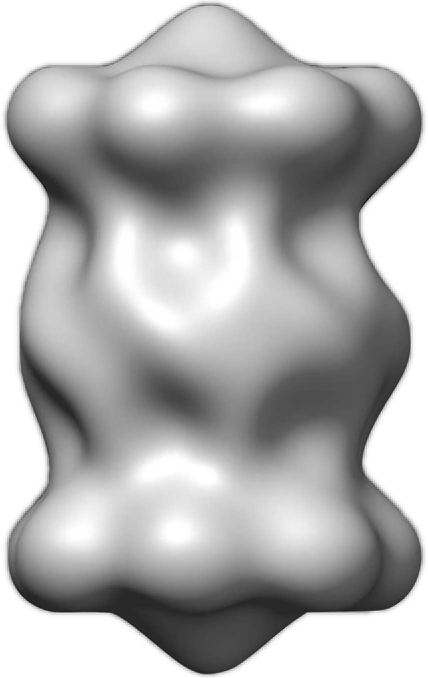
3D structure determination from 2D projection images

- frame alignment
- CTF estimation
- particle picking
- 2D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- resolution estimation
- map sharpening
- model building

particles



initial model



The problem:

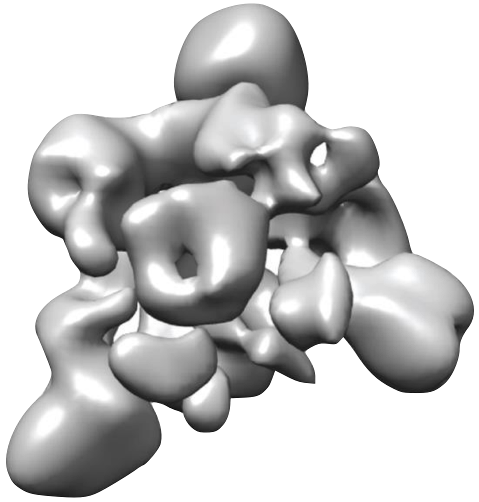
- **have to determine 5 parameters for each 2D particle image (3 Euler angles and X/Y shifts)**

Generating an initial model for 3D refinement

- frame alignment
- CTF estimation
- particle picking
- 2D align/classify
- **initial model**
- 3D align/classify
- 3D refinement
- particle polishing
- resolution estimation
- map sharpening
- model building

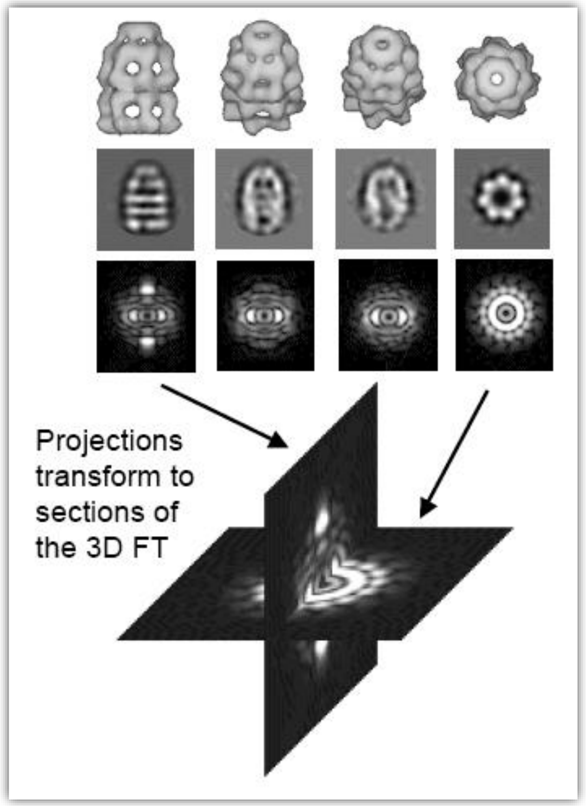
User-provided map

Low-pass filtered to ~20-50Å



- 3D map of the "expected" molecule (can come from AlphaFold)
- Biased approach which can be problematic if there is no prior knowledge

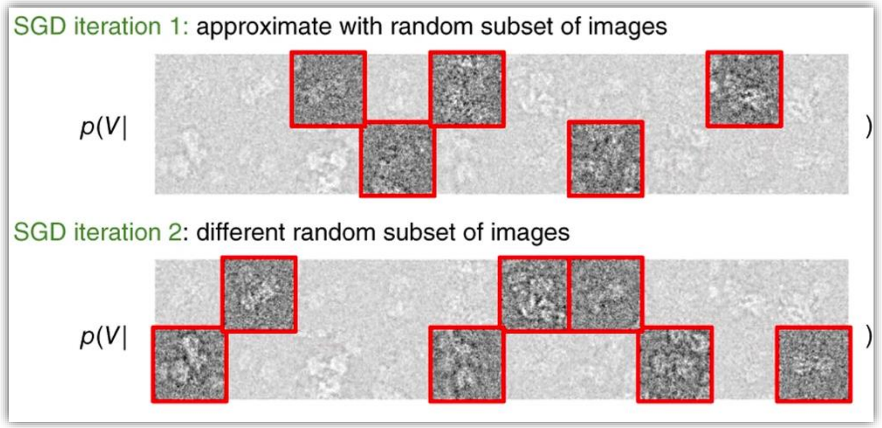
Common lines approach

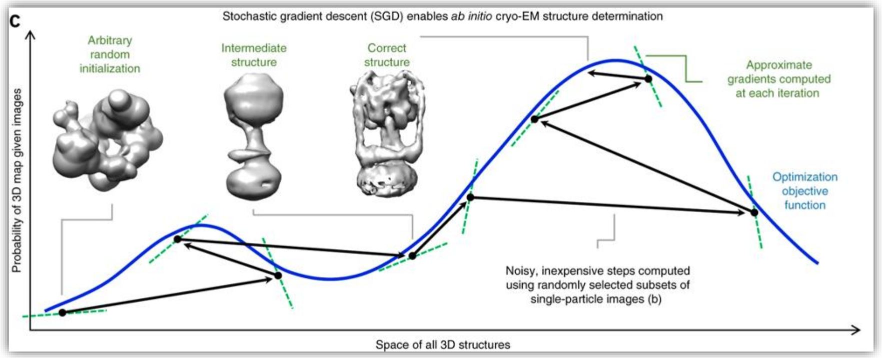


- Approach based on finding the common line shared by 2D projections in Fourier space
- Unbiased but works poorly due to low signal to noise in EM images

Saibil et al., Acta Cryst D (2000)

Ab initio reconstruction

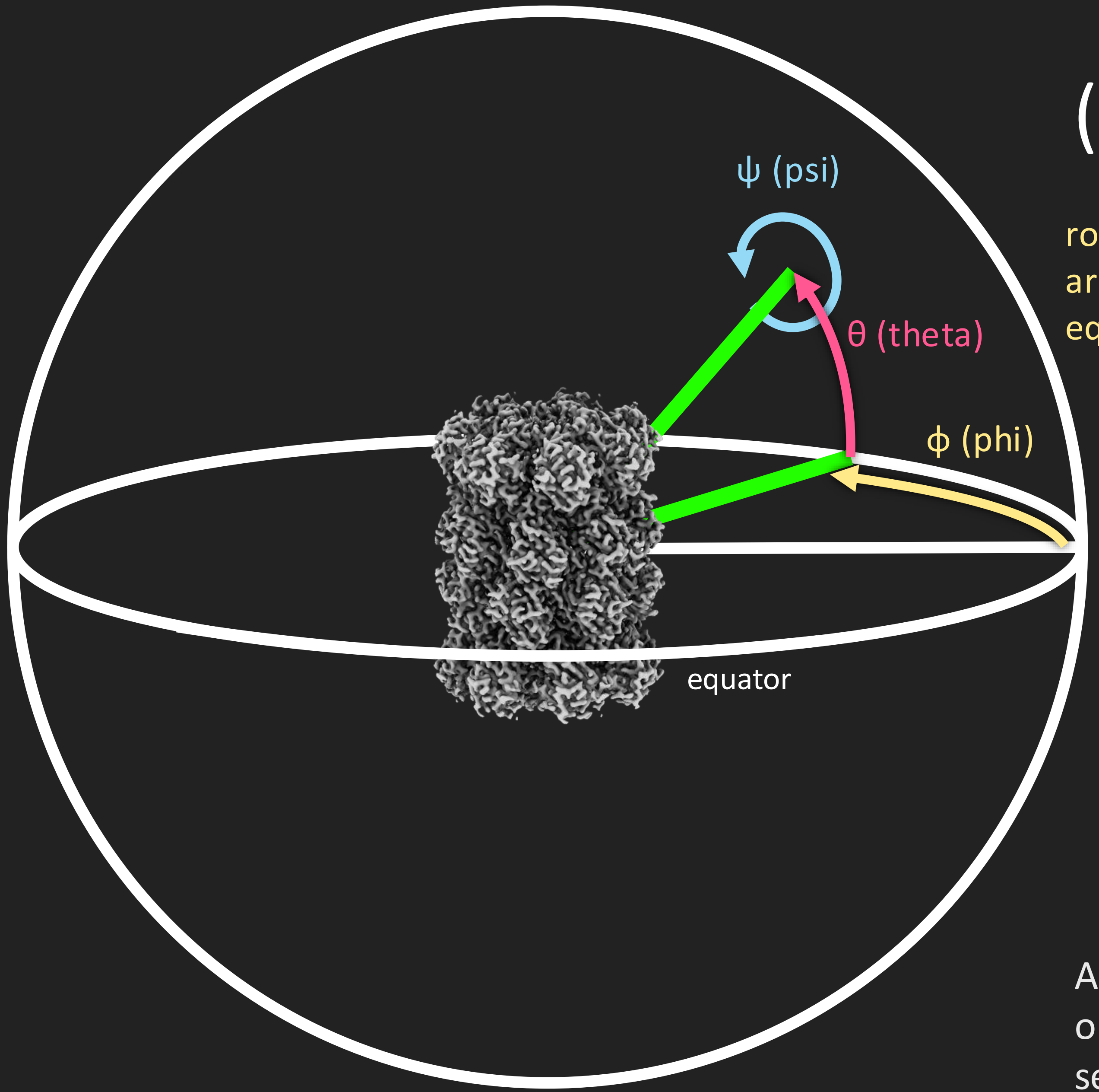




- Model generation using Stochastic Hill Climbing (Gradient Descent)
- Unbiased approach but can produce non-sensical maps

Punjani et al., Nat Methods (2017)

Defining particle orientation with respect to the 3D object

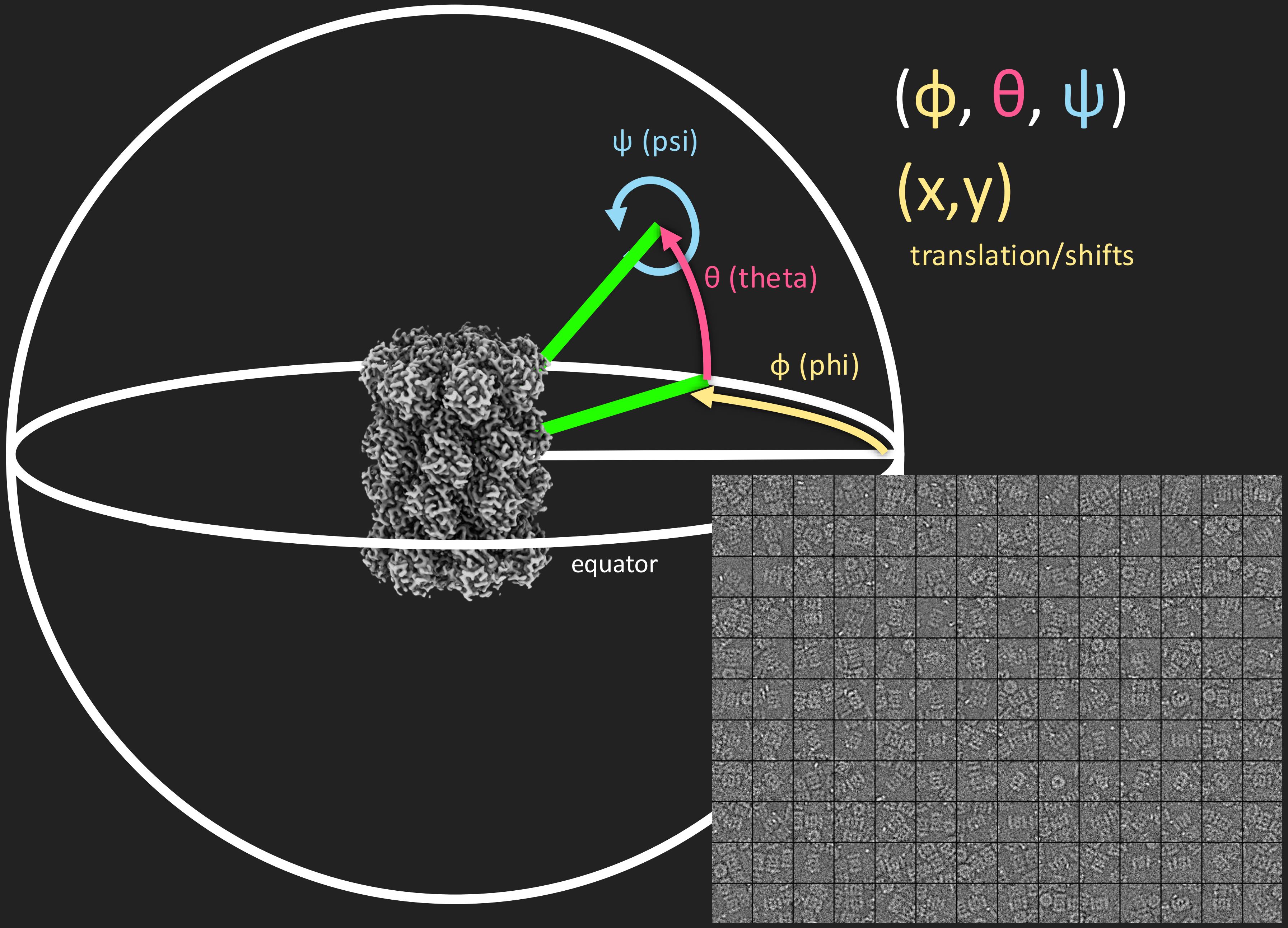


(ϕ, θ, ψ)

- ϕ: rotation around equator
- θ: elevation from equator
- ψ: in-plane rotation

According to the Euler definition, any orientation can be achieved by a maximum set of three successive rotations.

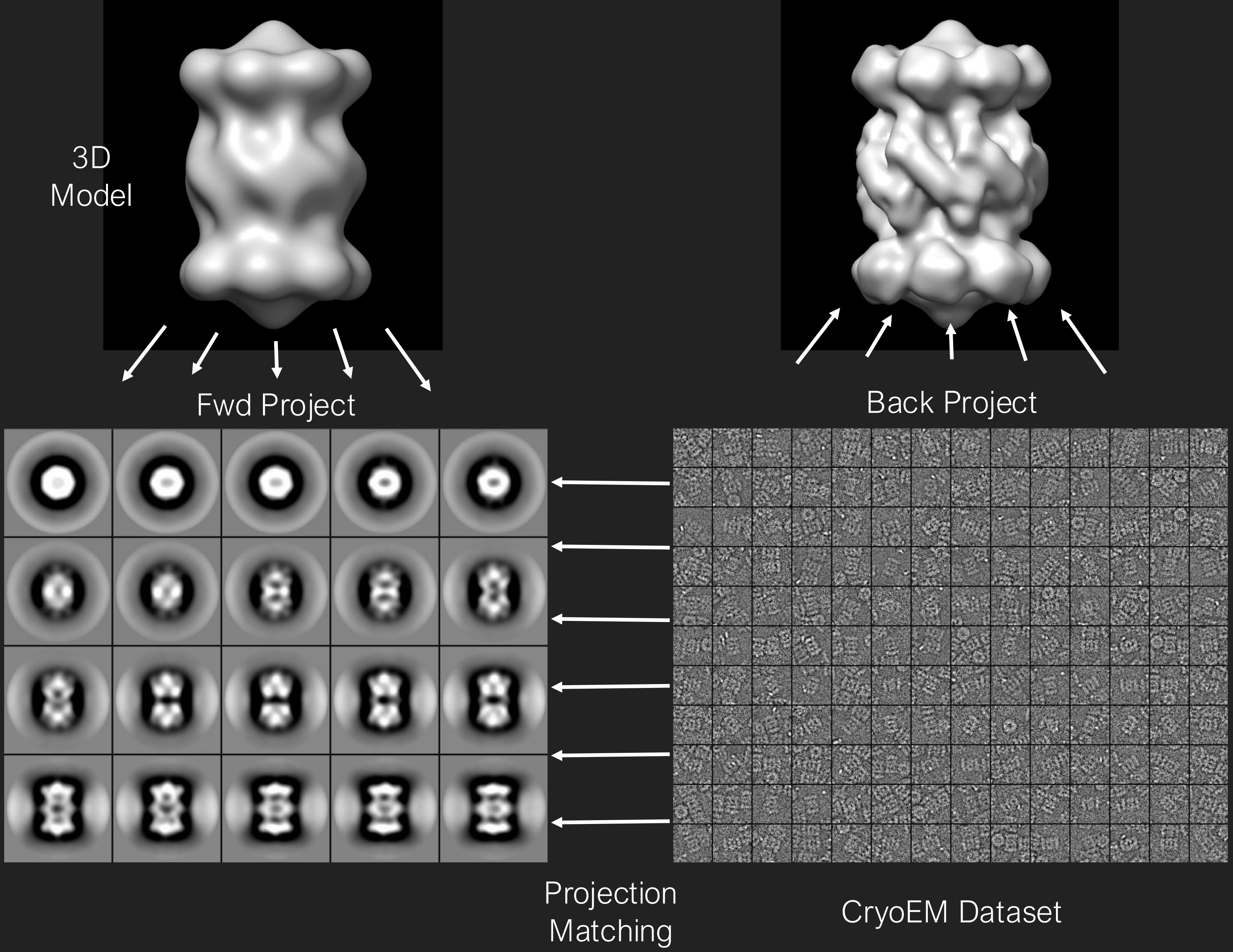
Defining particle orientation with respect to the 3D object

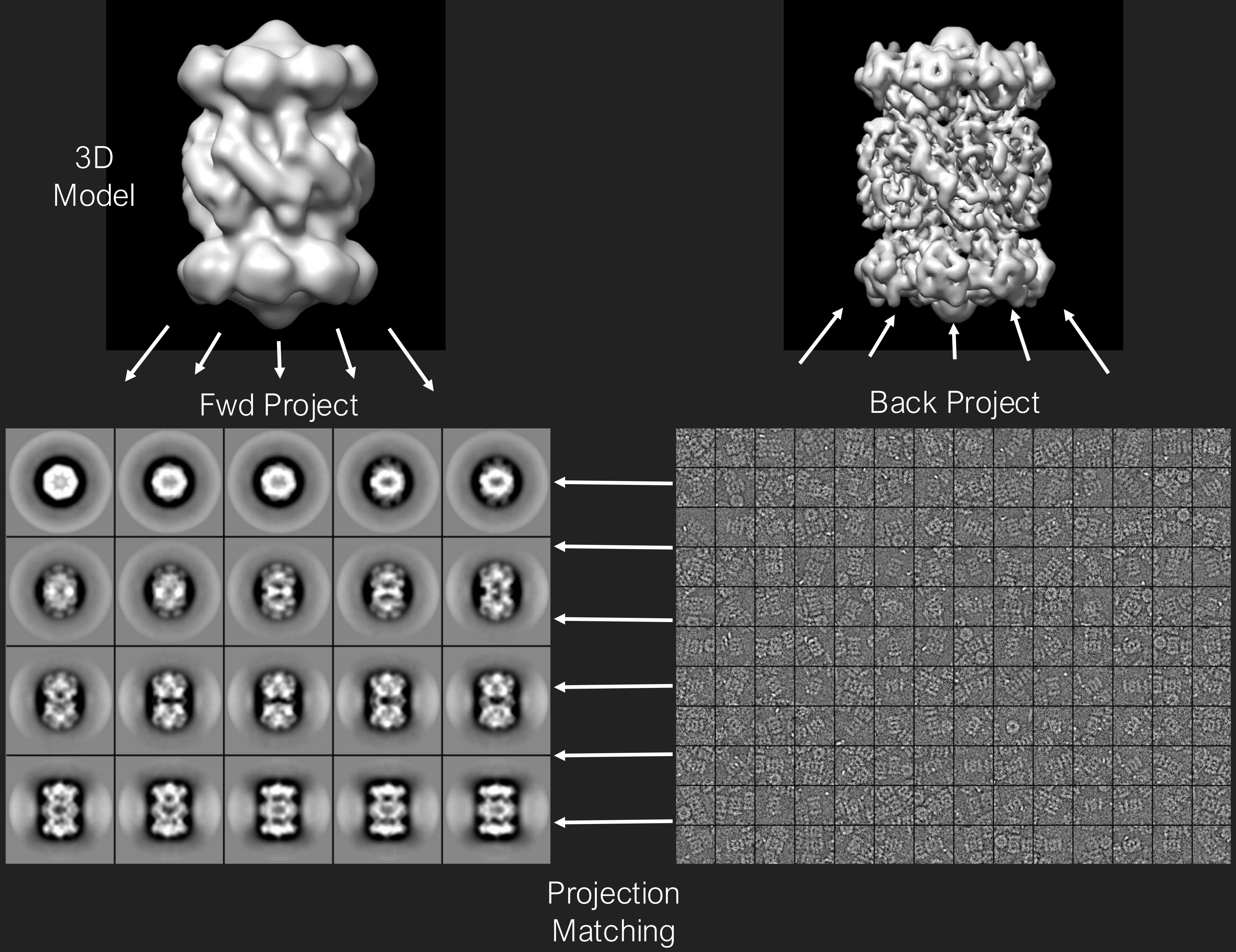


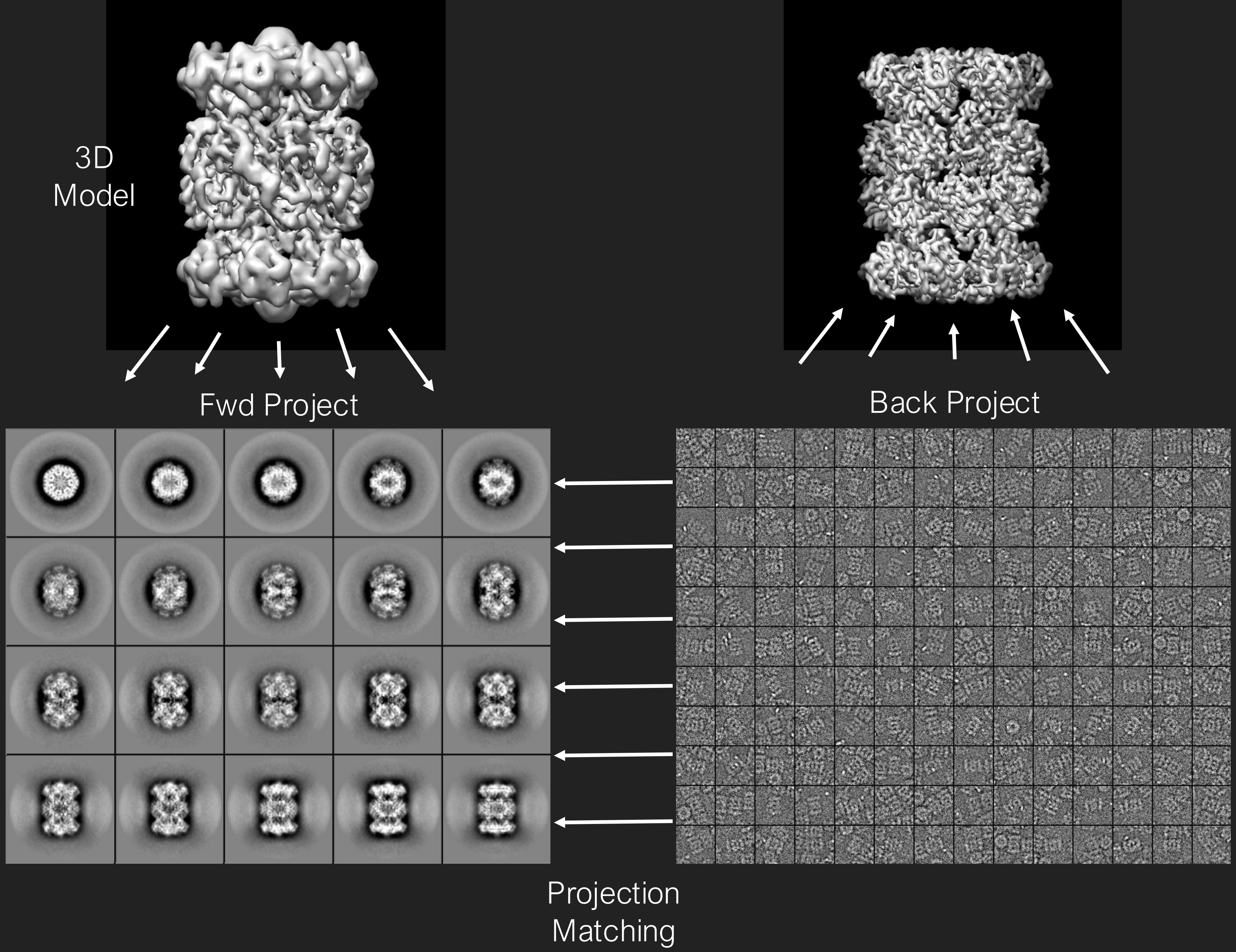
(ϕ, θ, ψ)

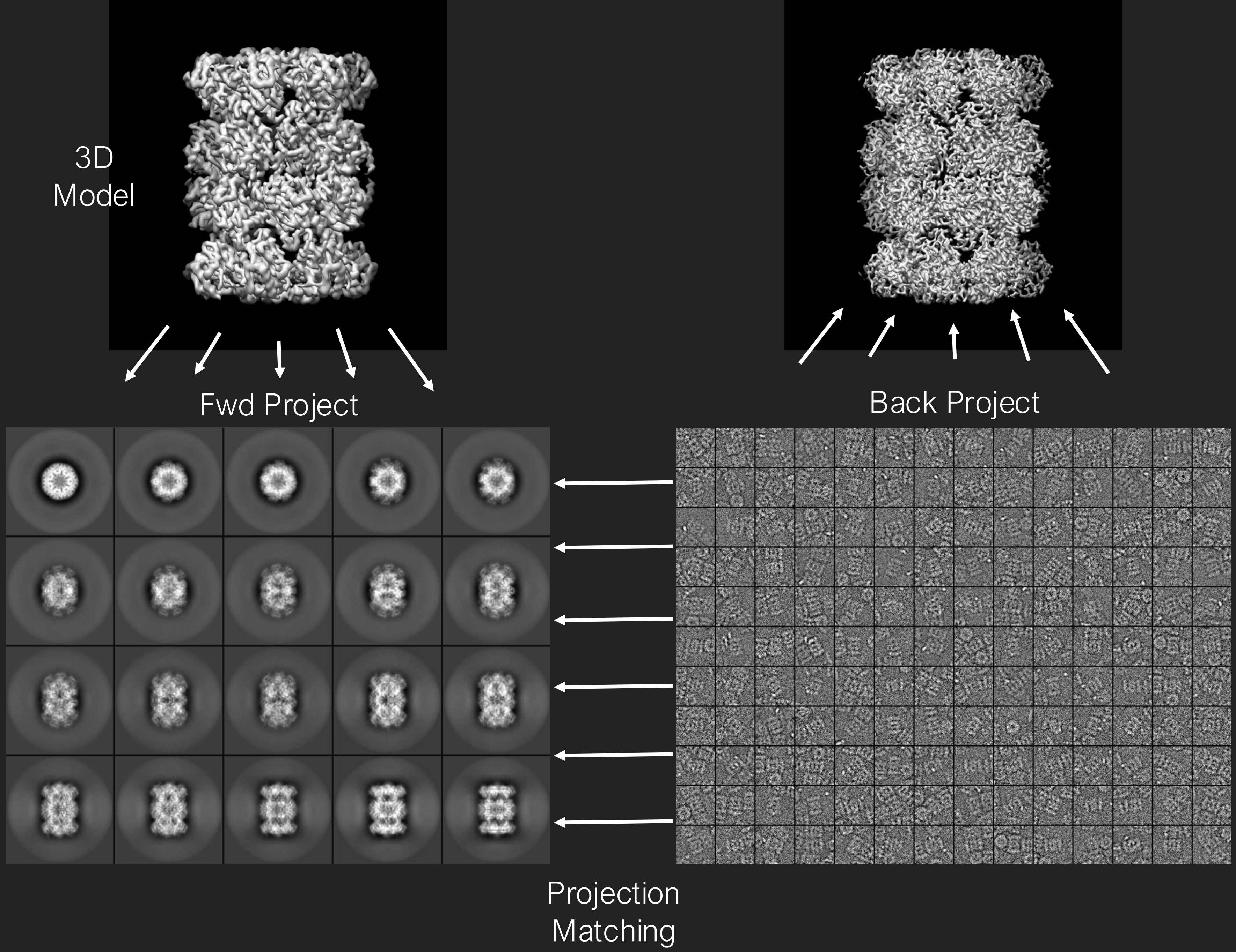
(x,y)

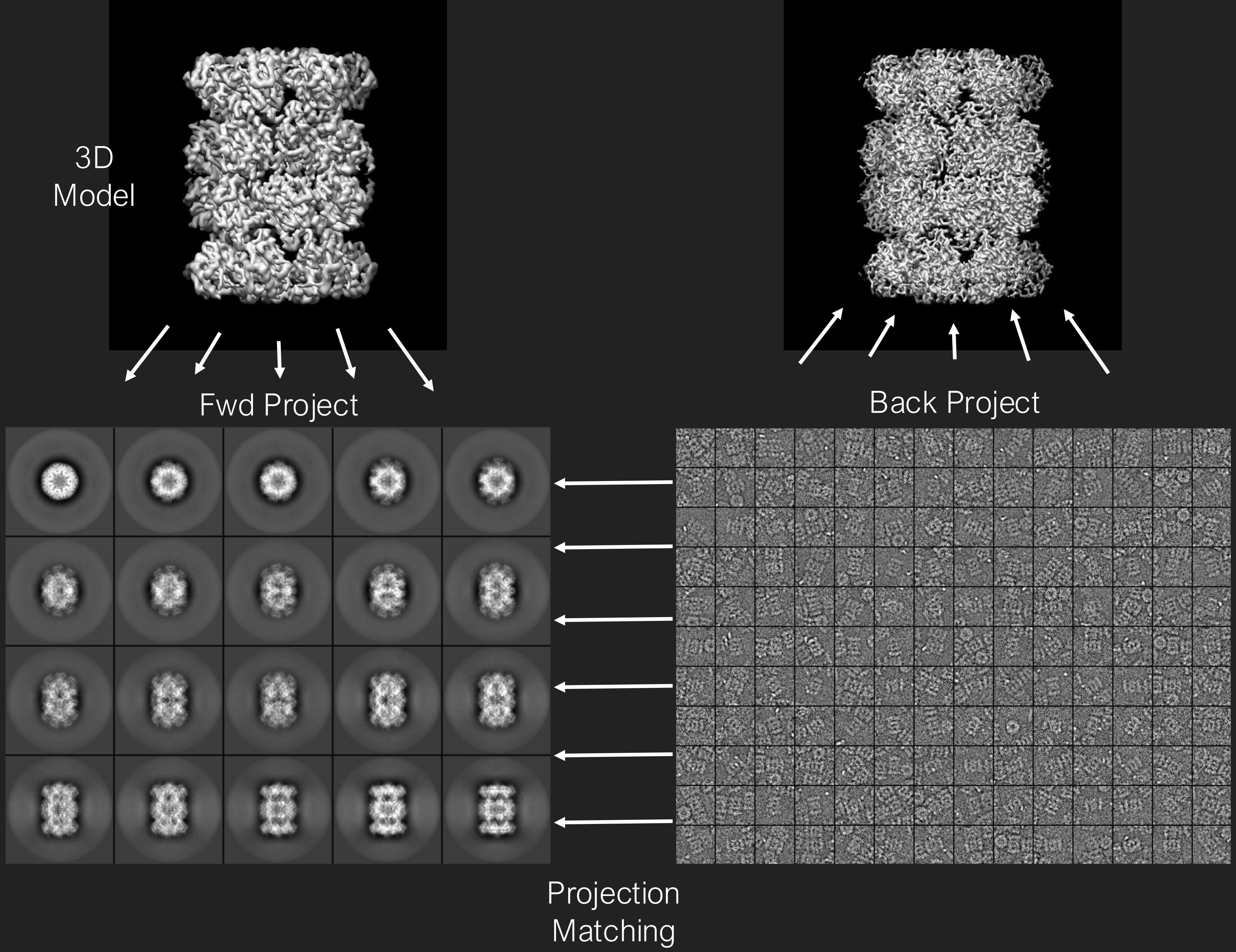
translation/shifts

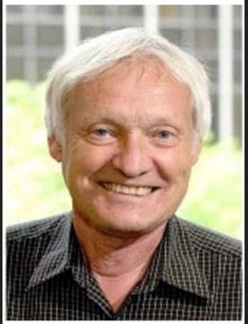

3D
Model
Fwd Project
Back Project
Projection
Matching
CryoEM Dataset


3D
Model
Fwd Project
Back Project
Projection
Matching


3D
Model
Fwd Project
Back Project
Projection
Matching

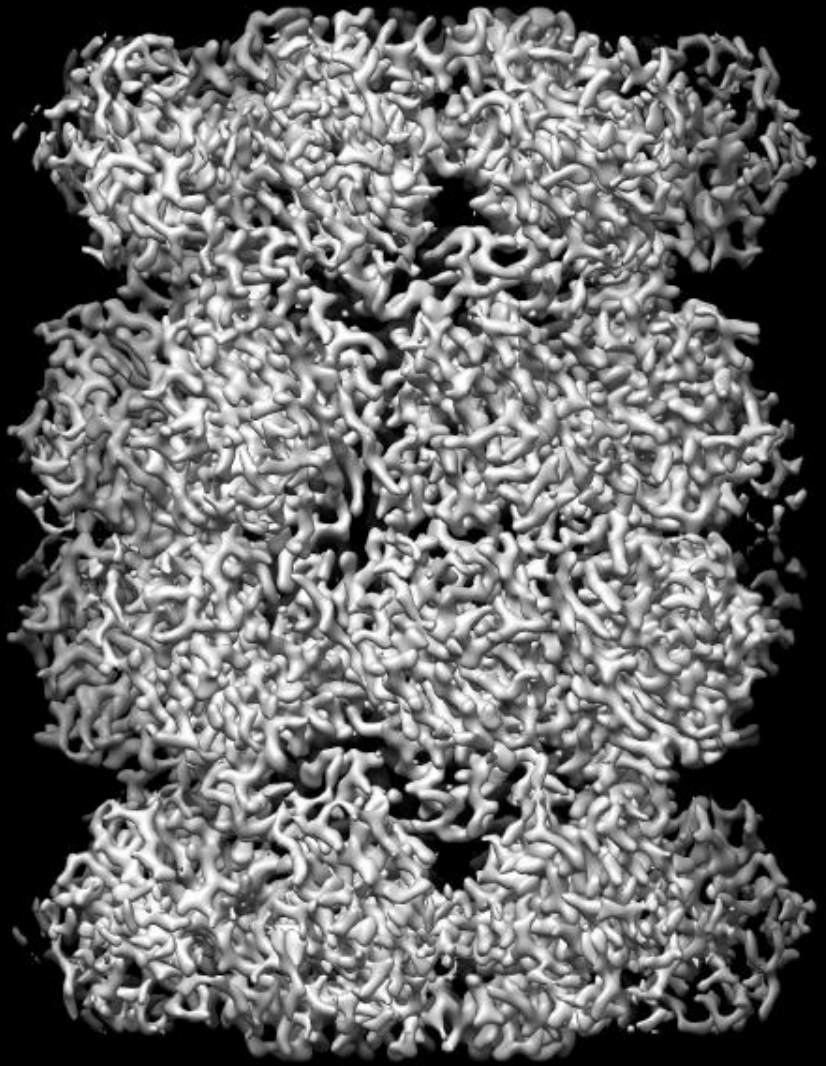

3D
Model
Fwd Project
Back Project
Projection
Matching





Joachim Frank
(Columbia University,
New York)

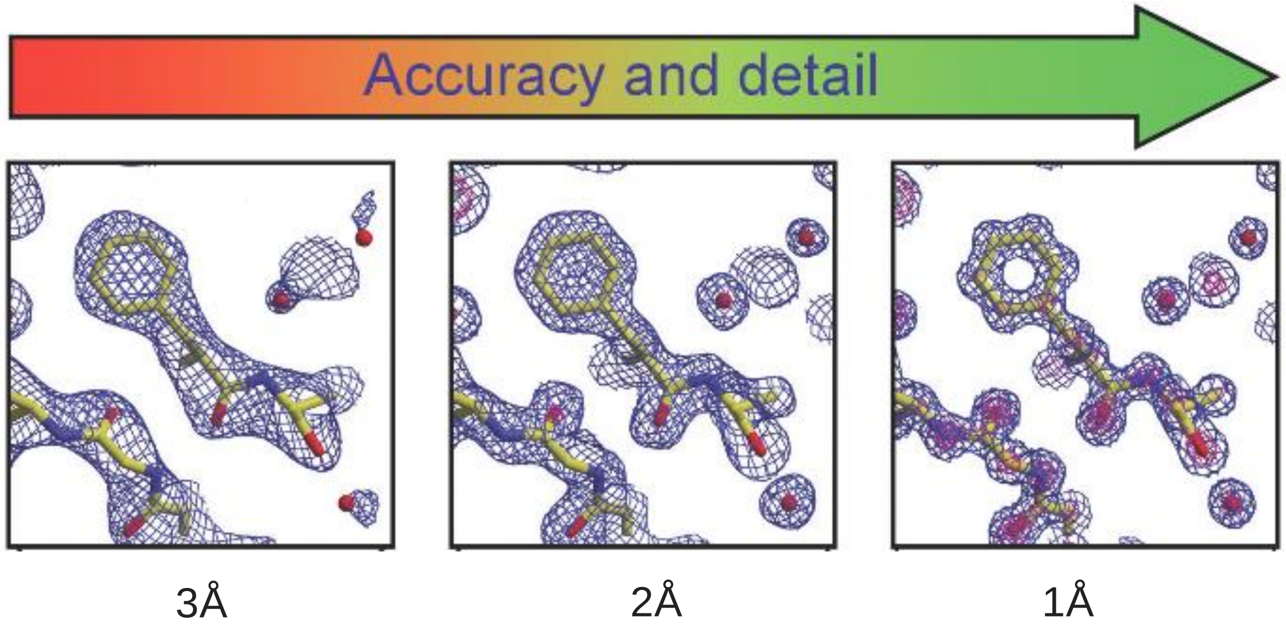
The final product: EM map of the biomolecule



Determining the resolution of EM density map

- frame alignment
- CTF estimation
- particle picking
- 2-D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- **resolution estimation**
- map sharpening
- model building

- Resolution, in structure determinations, **is the distance corresponding to the smallest observable feature**: if two objects are closer than this distance, they cannot be discerned

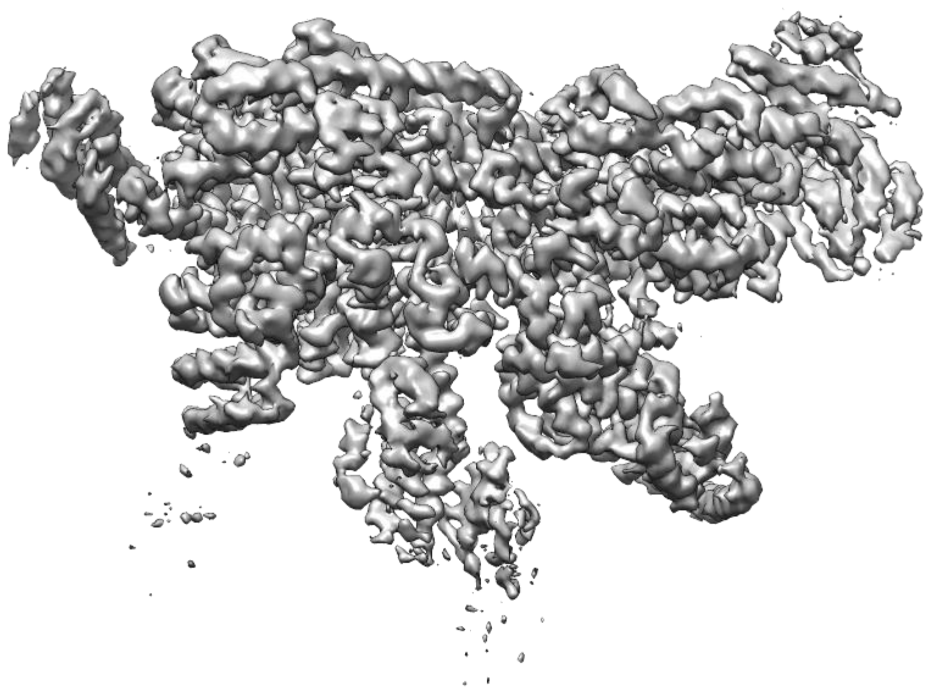


- Usually expressed in Å (10^{-10}m), the resolution is used as **a measure of map quality and its' interpretability** (i.e., you cannot build atoms in a 20Å map)

But how can we assign a resolution of a map?

- frame alignment
- CTF estimation
- particle picking
- 2-D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- **resolution estimation**
- map sharpening
- model building

- Resolution, in structure determinations, **is the distance corresponding to the smallest observable feature**: if two objects are closer than this distance, they cannot be discerned



- Usually expressed in Å (10^{-10}m), the resolution is used as **a measure of map quality and its' interpretability** (i.e., you cannot build atoms in a 20Å map)

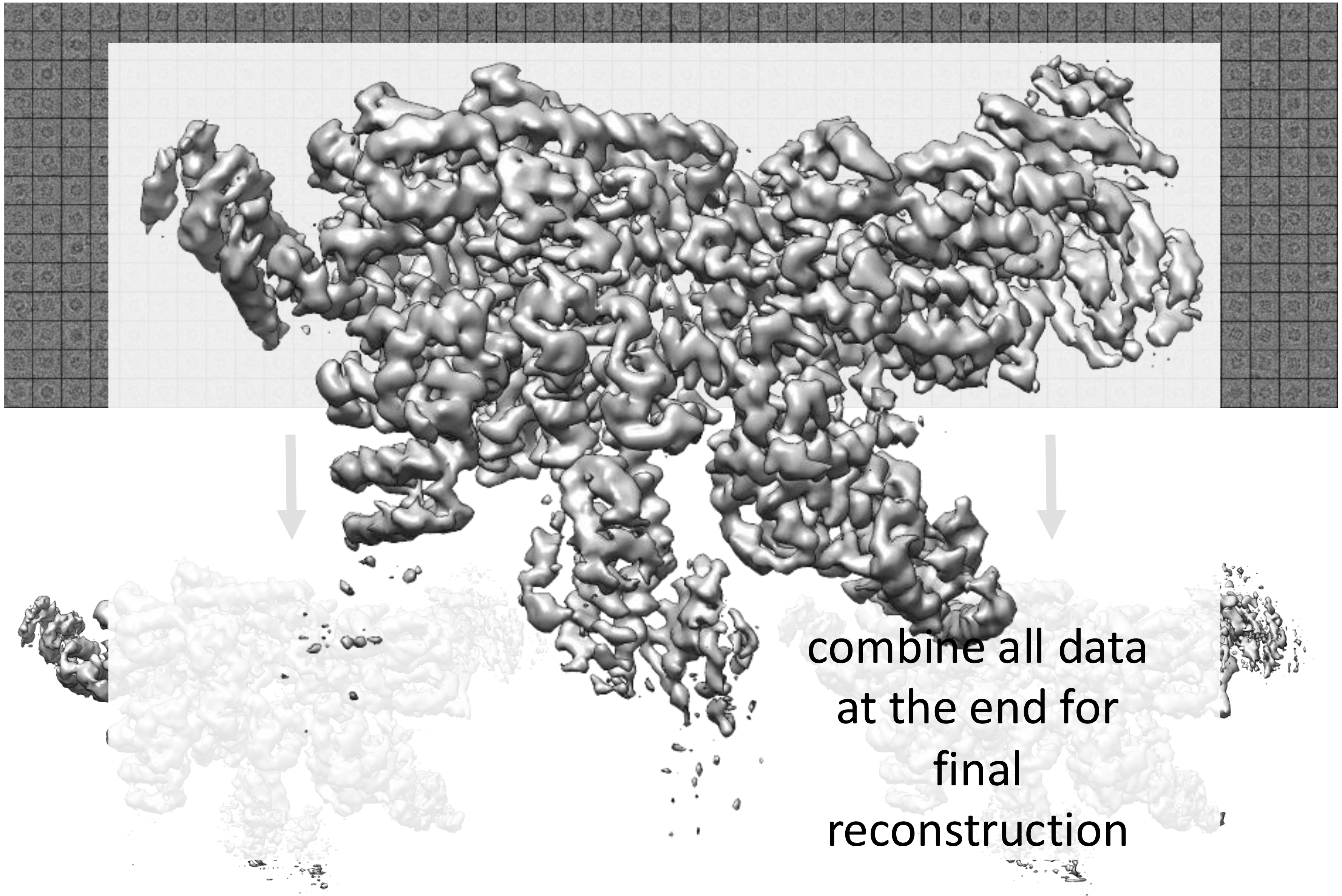
How to assess resolution?

- frame alignment
- CTF estimation
- particle picking
- 2-D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- **resolution estimation**
- map sharpening
- model building

Fourier Shell Correlation: a function plotted versus resolution (1/d $Å^{-1}$) whose values are correlation coefficients computed between the Fourier Transforms of two volumes over shells of approximately equal resolution (Penczek, Methods in Enzymology, 2010)

$$FSC(r) = \frac{\sum_{r_i \epsilon r} F_1(r_i) \cdot F_2(r_i)^*}{\sqrt[2]{\sum_{r_i \epsilon r} |F_1(r_i)|^2 \cdot \sum_{r_i \epsilon r} |F_2(r_i)|^2}}$$

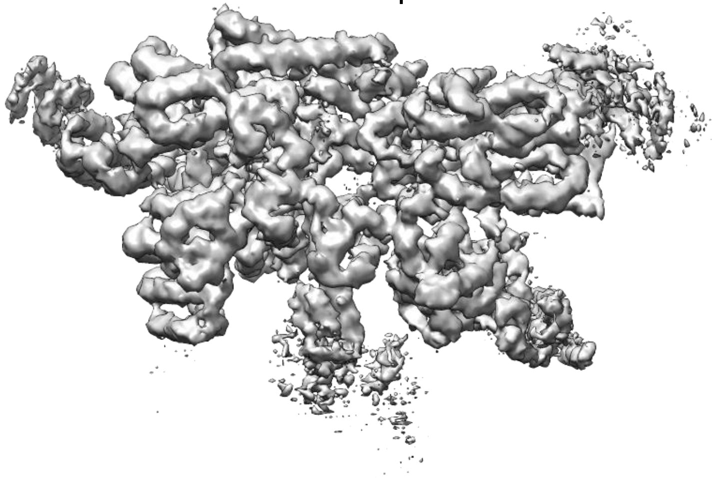
Fourier Shell Correlation



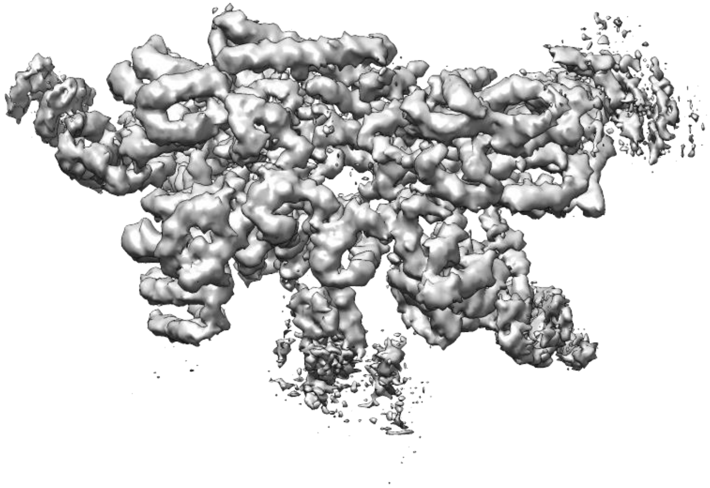
"Gold Standard" FSC

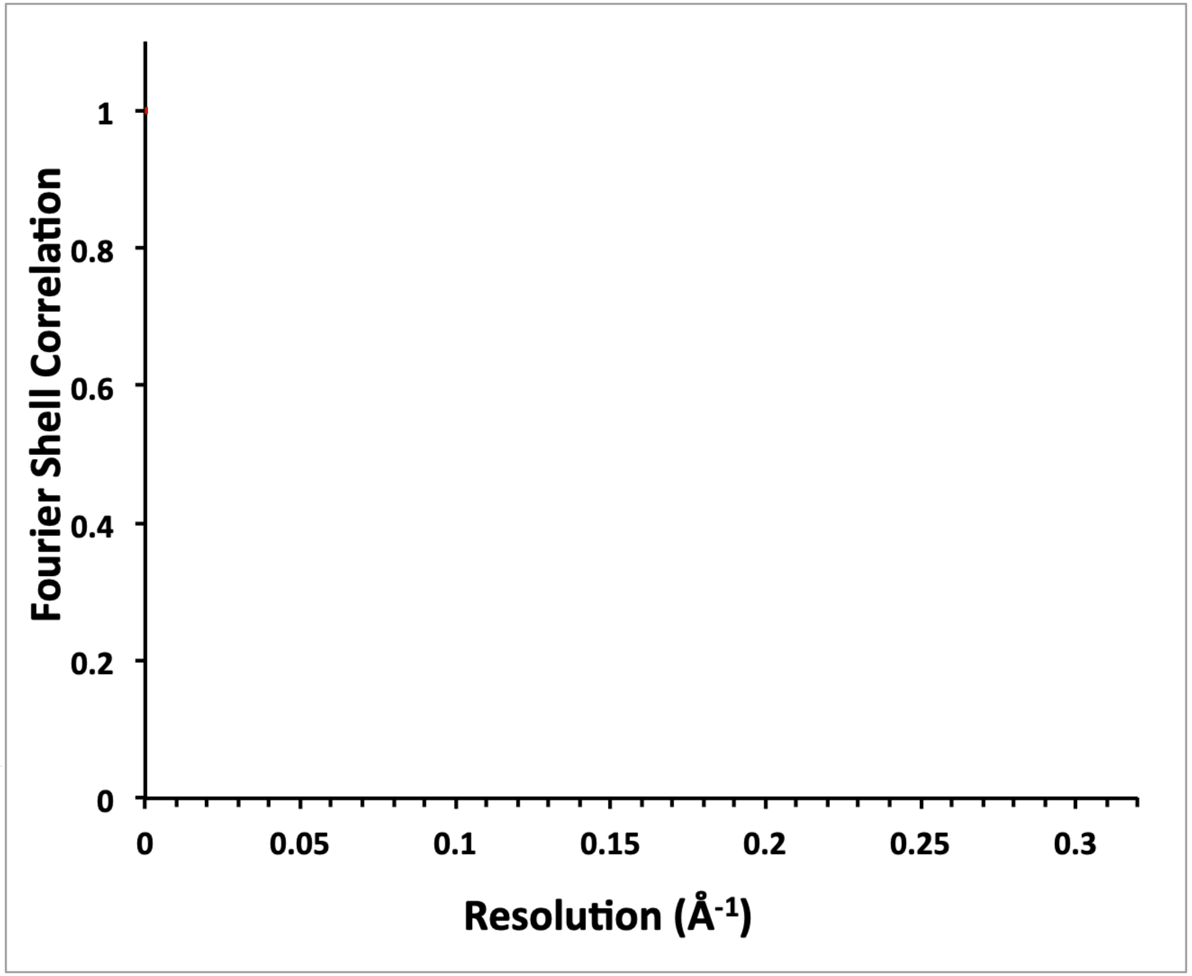
Fourier Shell Correlation

half-map 1



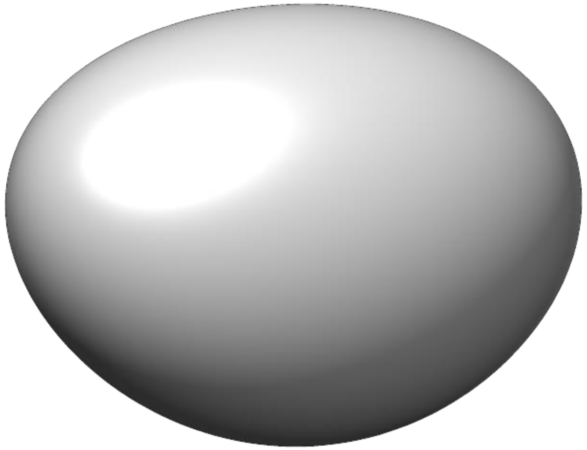
half-map 2



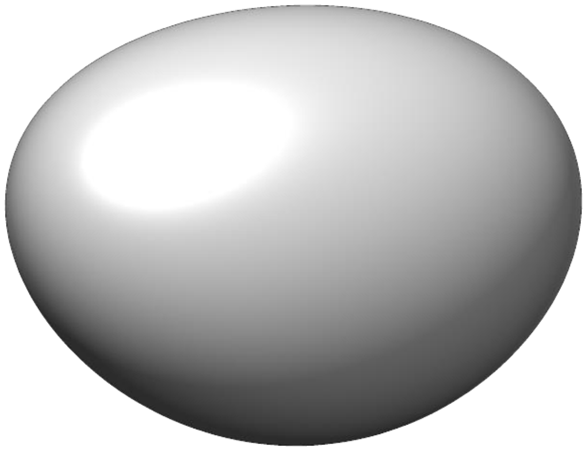


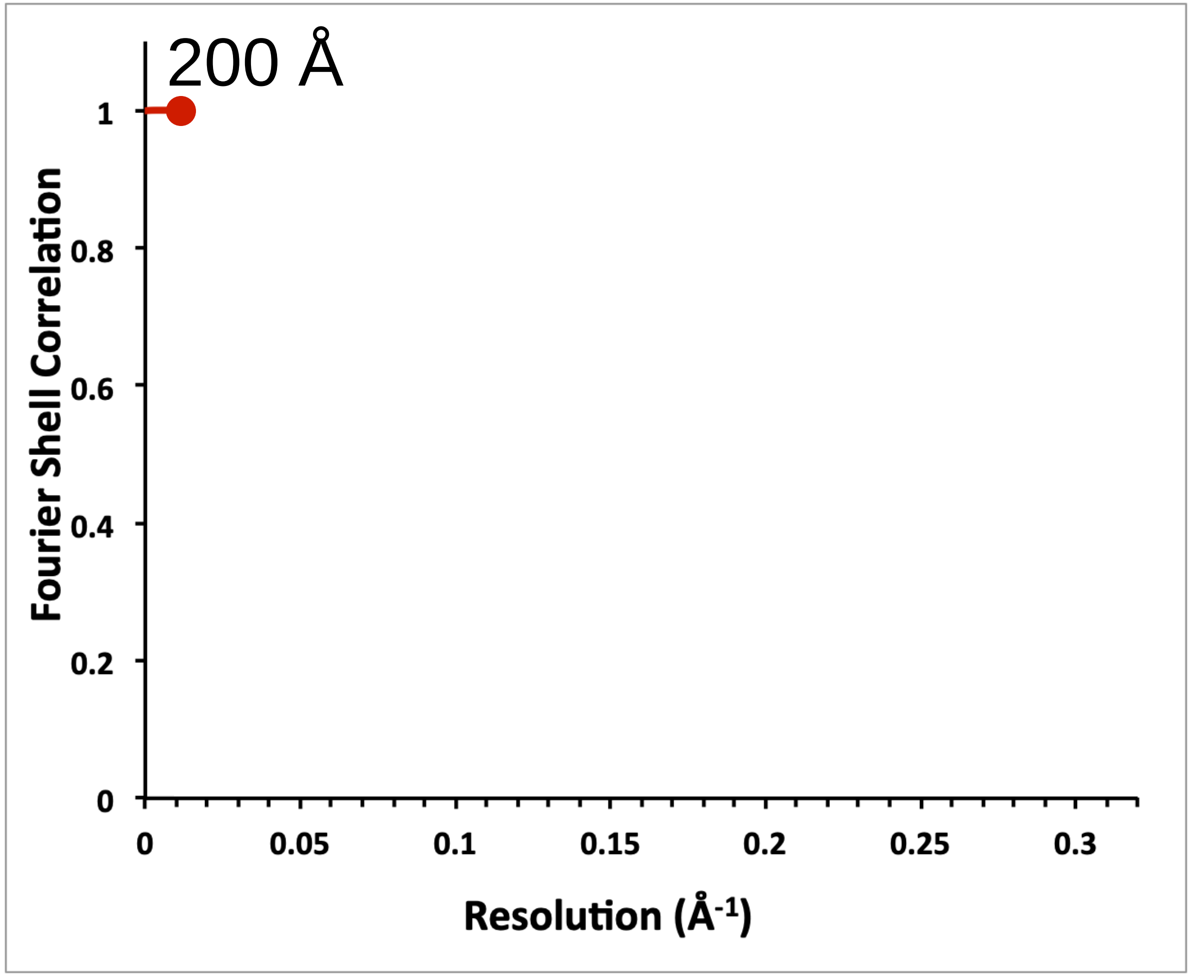
Fourier Shell Correlation

half-map 1



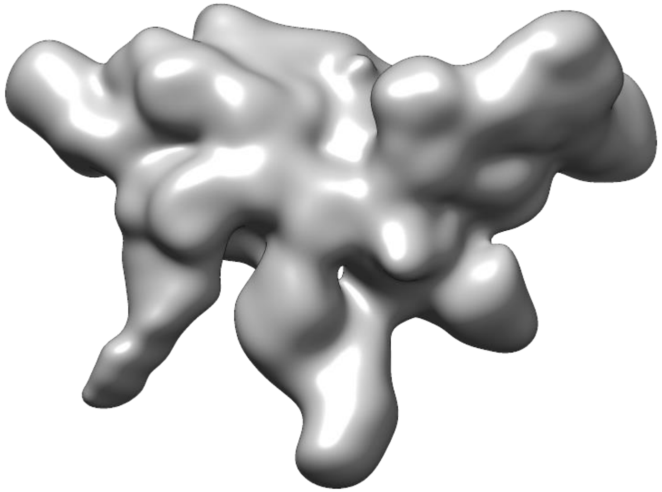
half-map 2



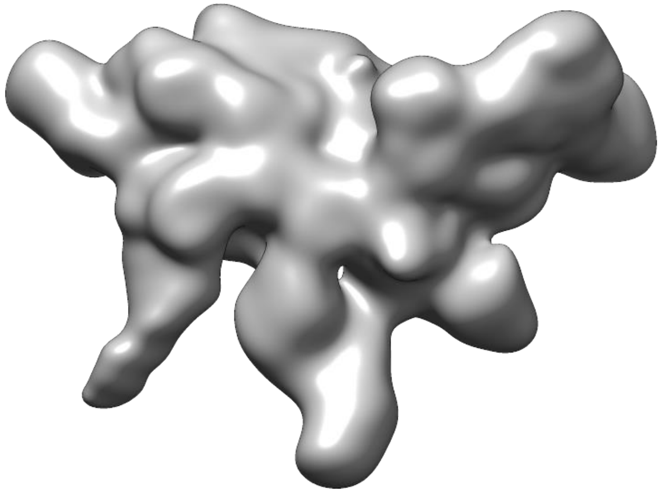


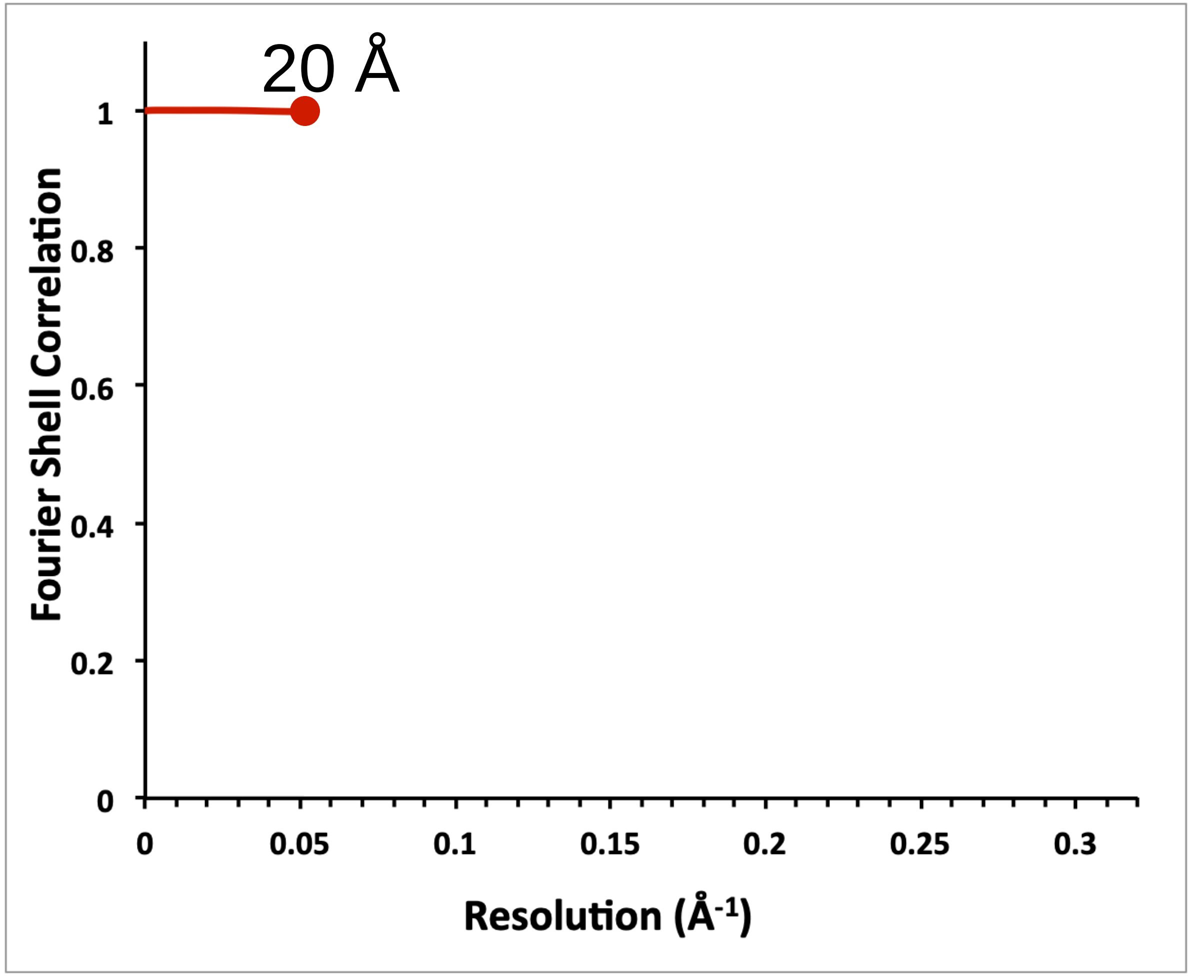
Fourier Shell Correlation

half-map 1



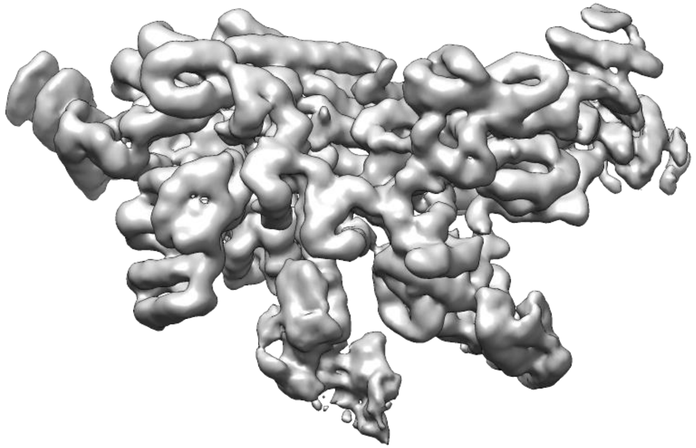
half-map 2



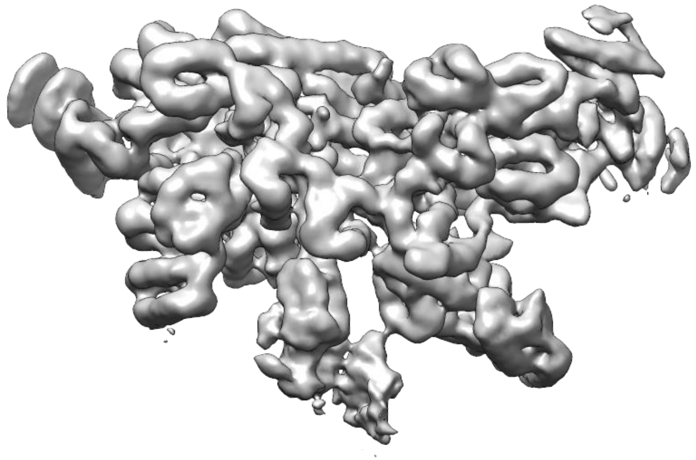


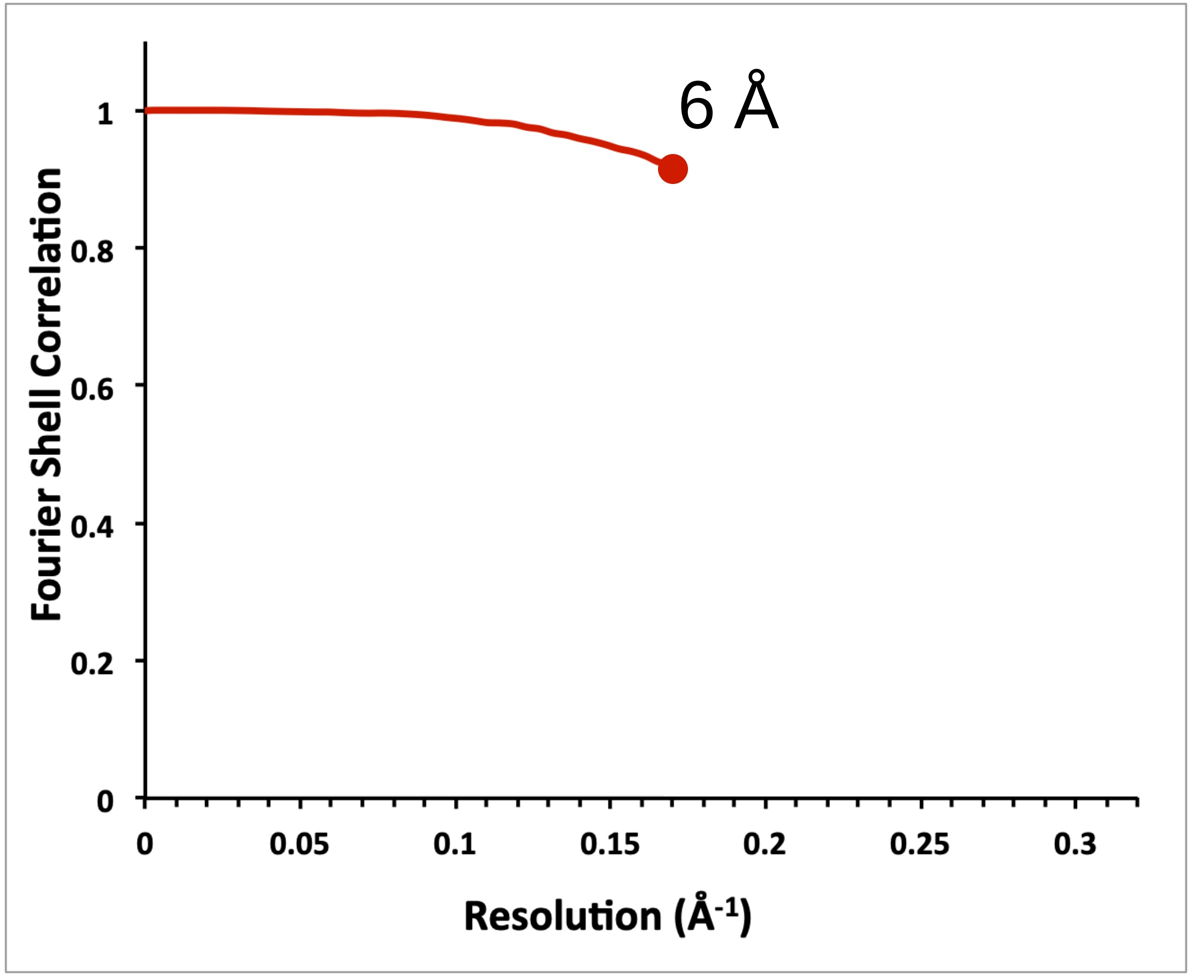
Fourier Shell Correlation

half-map 1



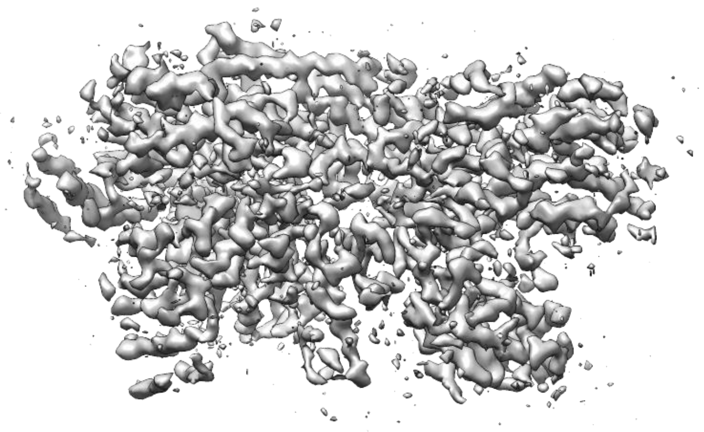
half-map 2



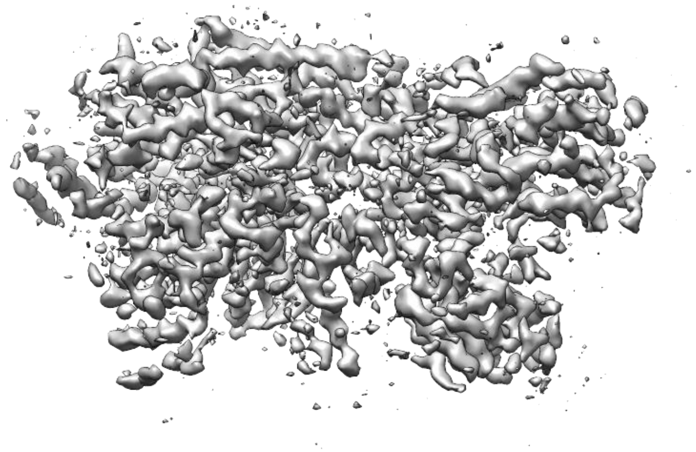


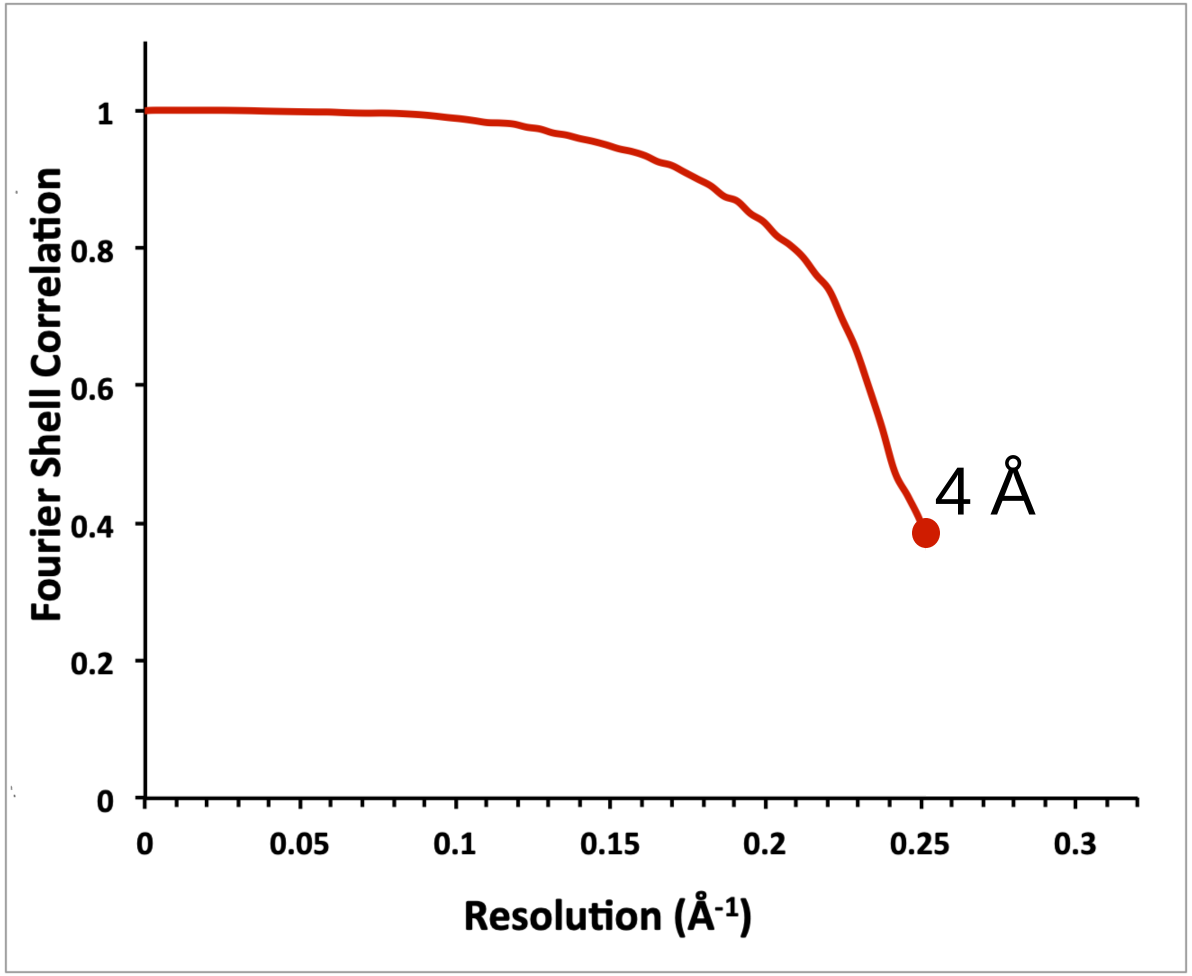
Fourier Shell Correlation

half-map 1

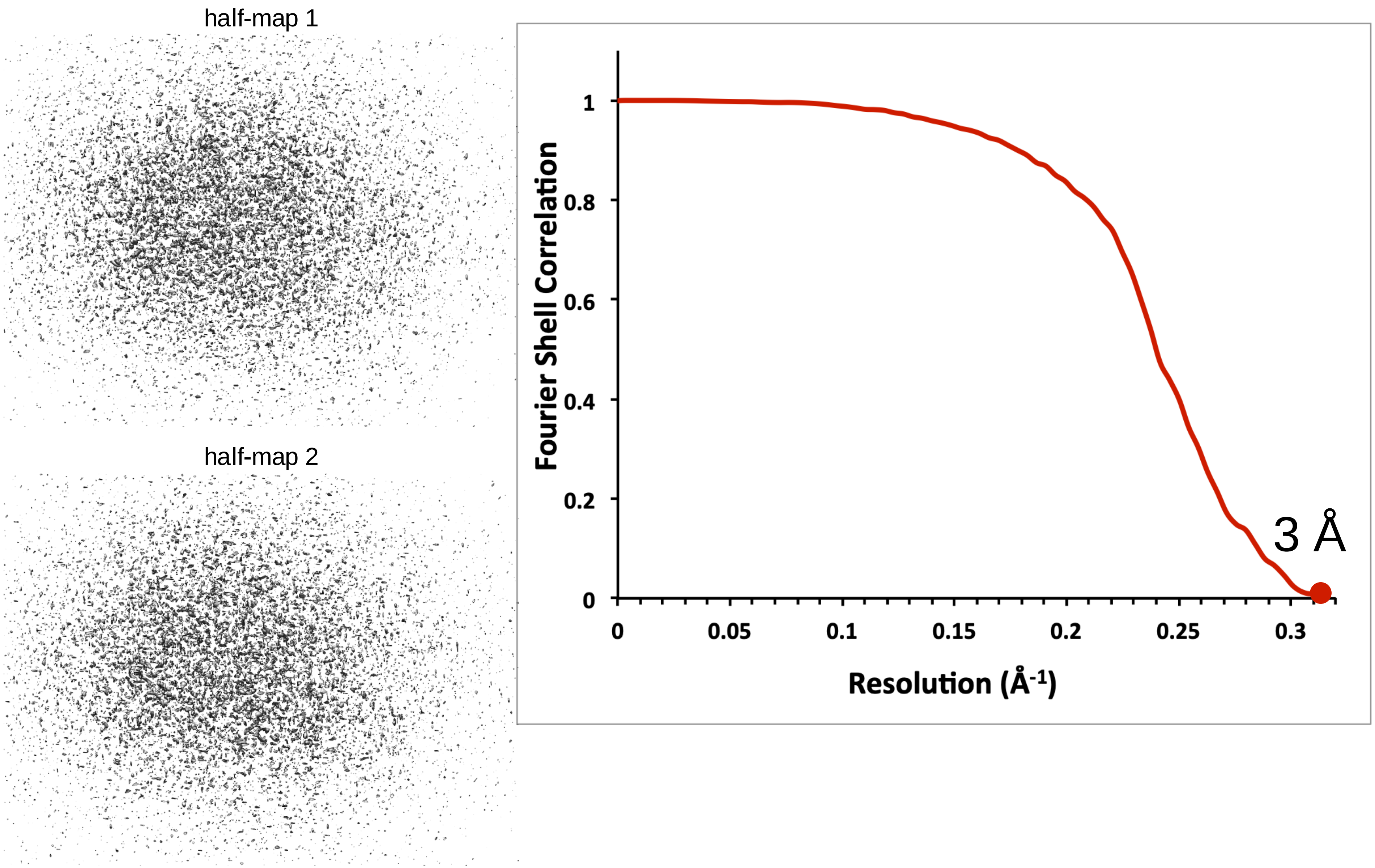


half-map 2





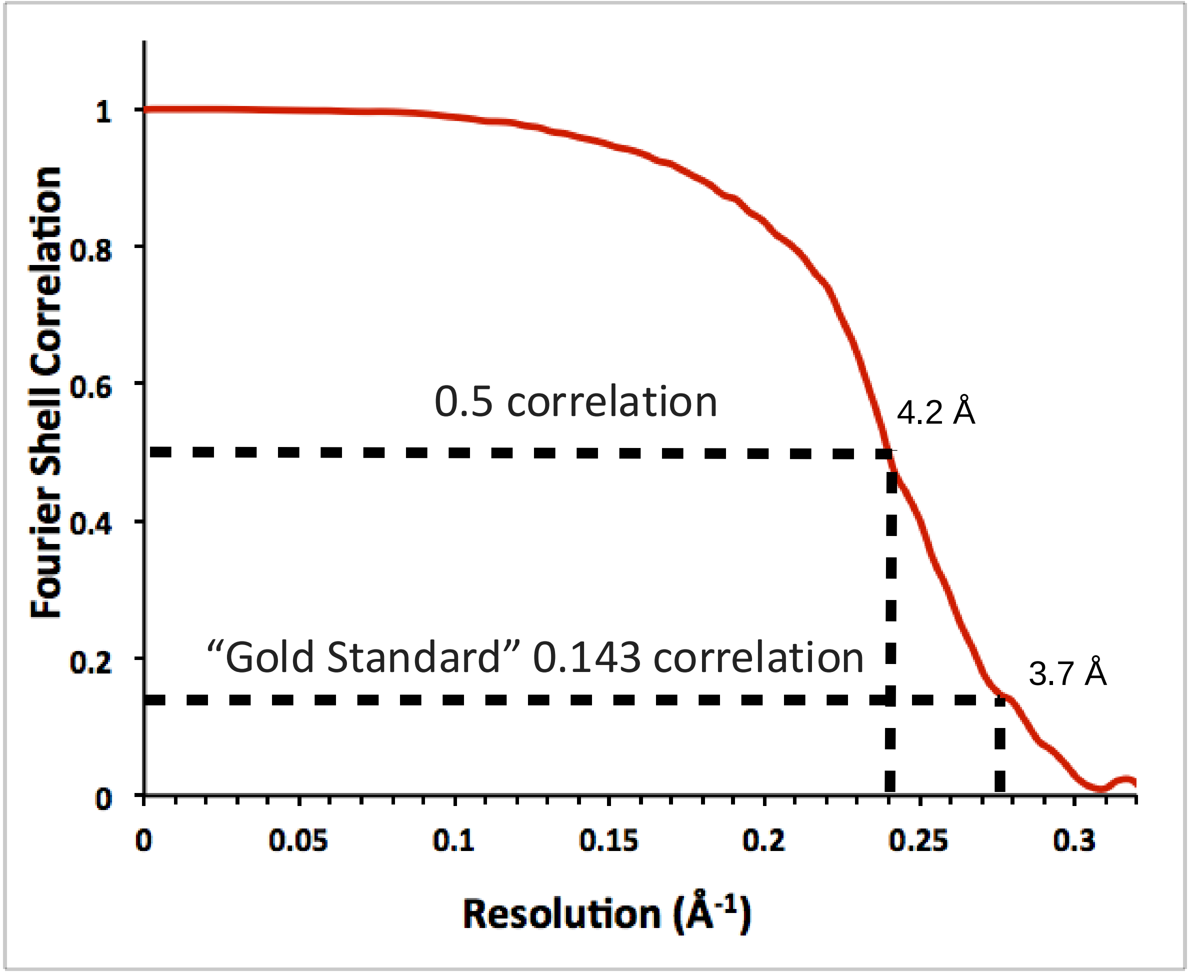
Fourier Shell Correlation



How should an FSC curve look?

"Proper FSC curve remains one at low frequencies, which is followed by a semi-Gaussian fall-off and a drop to zero, in high frequencies oscillates around zero."

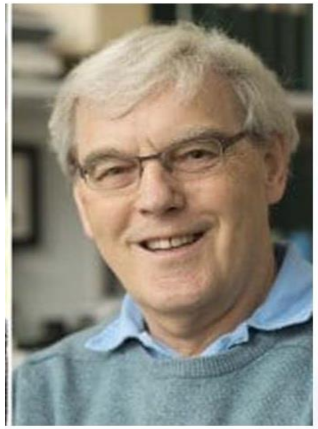
Penczek et al., 2011



$$FSC(r) = \frac{\sum_{r_i \epsilon r} F_1(r_i) \cdot F_2(r_i)^*}{\sqrt[2]{\sum_{r_i \epsilon r} |F_1(r_i)|^2 \cdot \sum_{r_i \epsilon r} |F_2(r_i)|^2}}$$

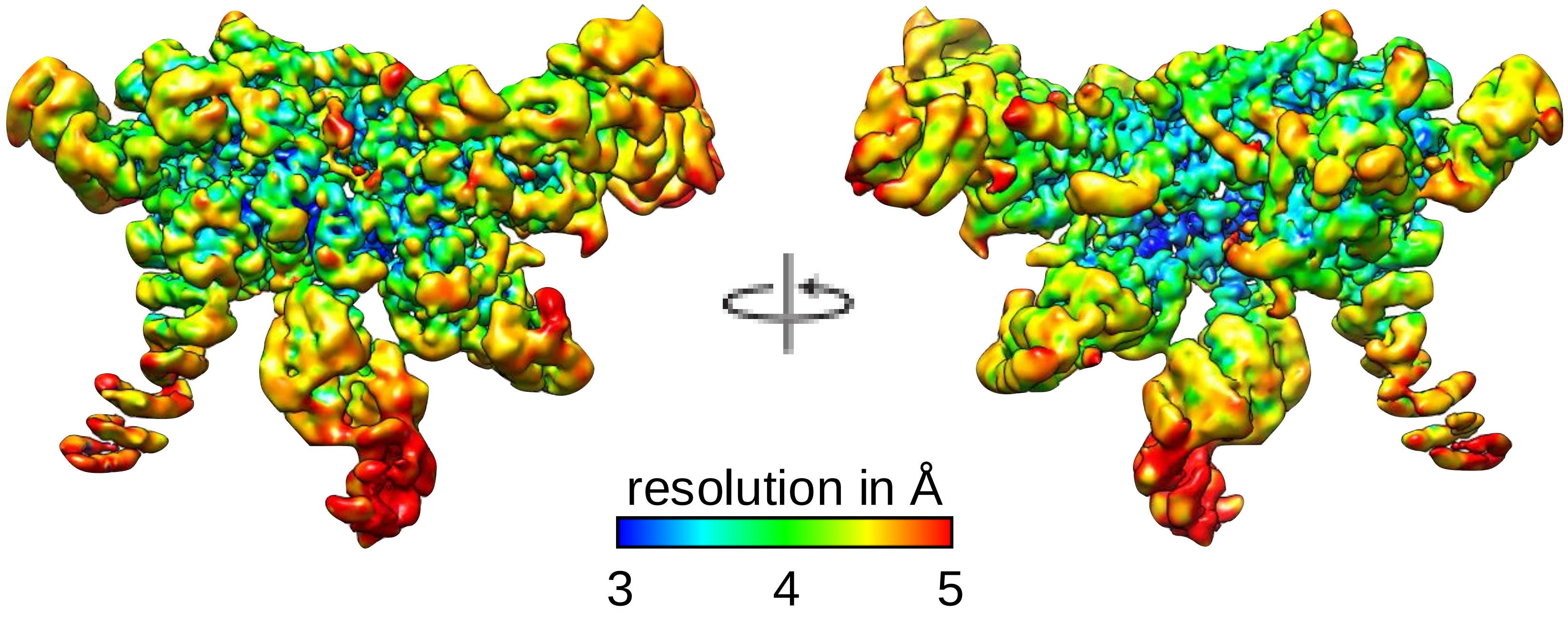
- Much debate in the 90s and early 2000s over the threshold used to assign global resolution.

- See for example: https://pubmed.ncbi.nlm.nih.gov/16125414/



Richard Henderson
(MRC Laboratory of Molecular Biology, Cambridge, U.K.)

Local Resolution Plots



Resmap - compares power of Fourier components
Bsoft - calculates windowed FSCs
Relion - calculates windowed FSCs
Sparx - calculates local variance from 2D images

- Useful for visualization of local resolution across the entire map
- Can be used for map flitering to improve interpetability
- Can be used to inform data processing (e.g., positioning of sorting masks)
- Can be used to inform model building (e.g., which areas to build and where to stop)

B-factor to balance attenuation of amplitudes at high resolution

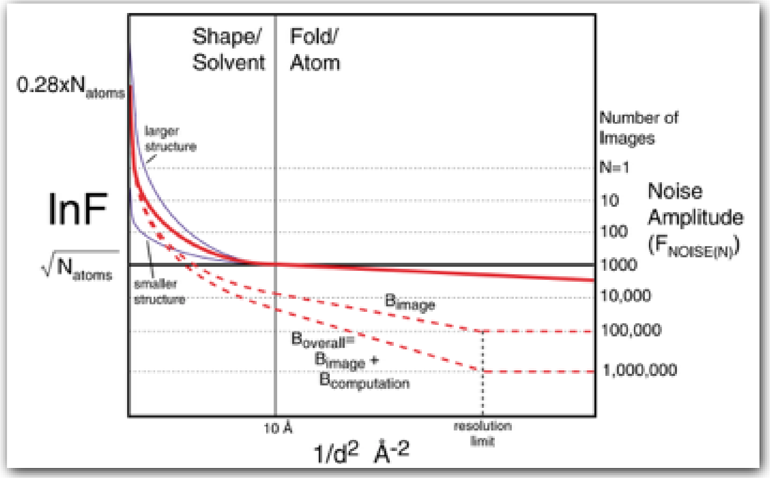
- frame alignment
- CTF estimation
- particle picking
- 2-D align/classify
- initial model
- 3D align/classify
- refinement
- particle polishing
- resolution estimation
- **map sharpening**
- model building

The problem:
combined effects of imaging and processing reduces observed high-frequency amplitudes

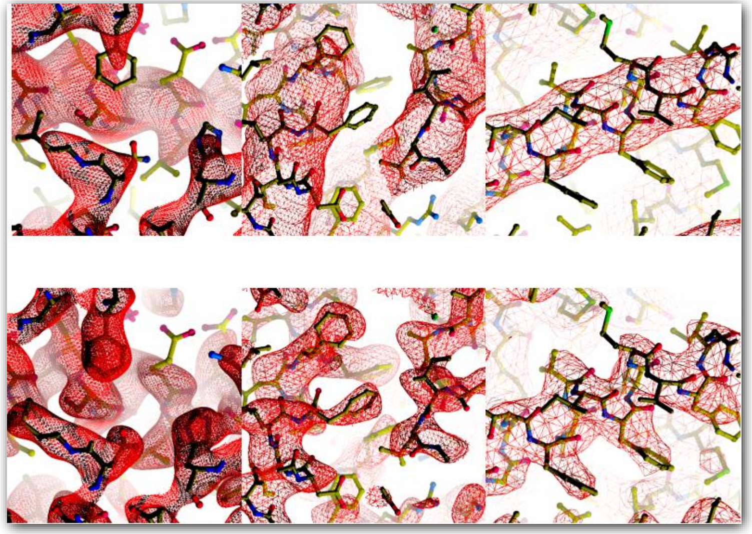
The solution:
apply negative B-factor to "sharpen" the map

What B-factor to use?
- Calculate from Guinier plot
- ad hoc - increase until noise becomes problematic

rotationally averaged spectrum



Rosenthal and Henderson JMB 2003



Unsharpened

Sharpened

Terwilliger et al. bioRxiv doi.org/10.1101/247049

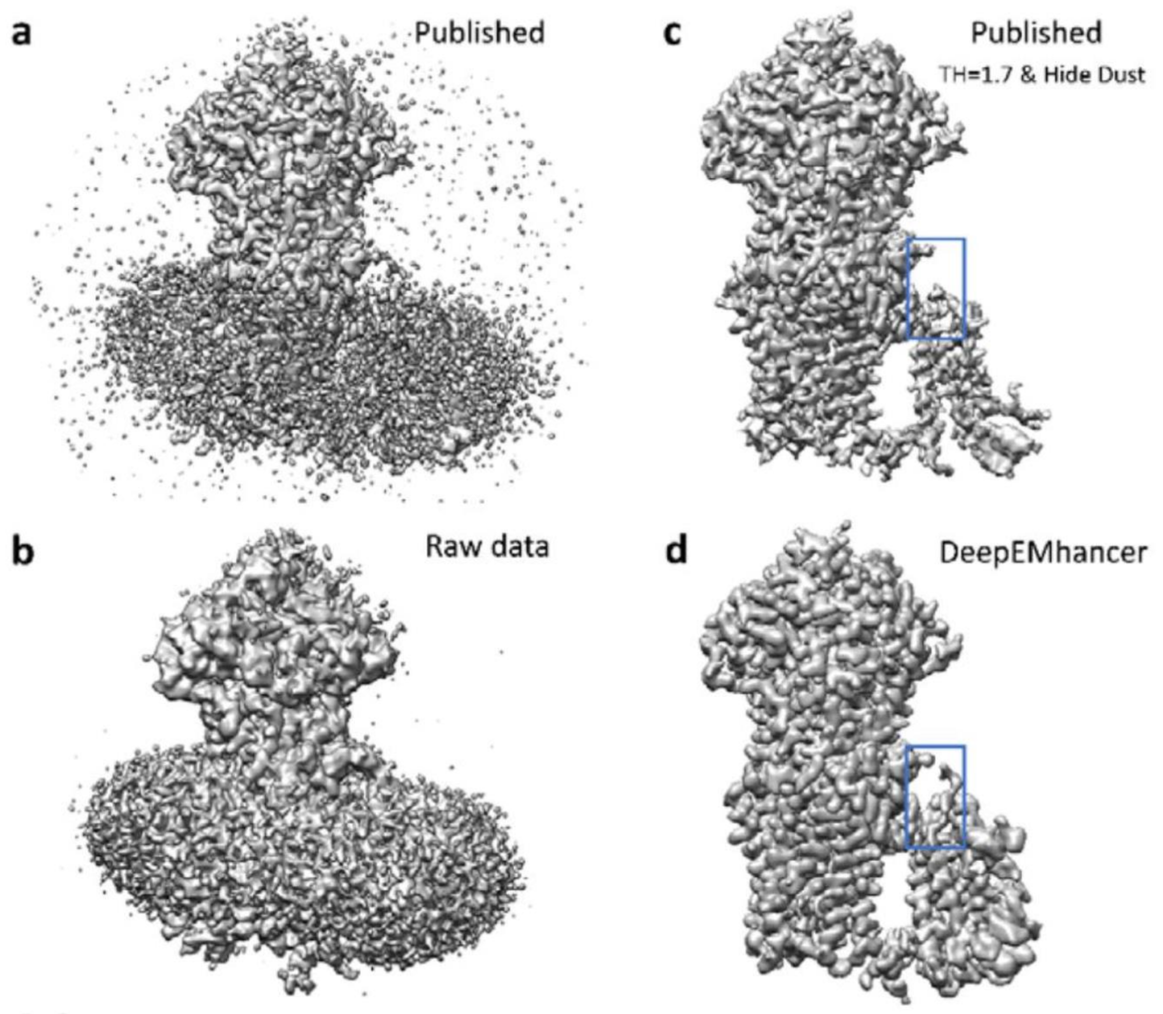
output: map with enhanced high resolution features

B-factor to balance attenuation of amplitudes at high resolution

- frame alignment
- CTF estimation
- particle picking
- 2-D align/classify
- initial model
- 3D align/classify
- refinement
- particle polishing
- resolution estimation
- **map sharpening**
- model building

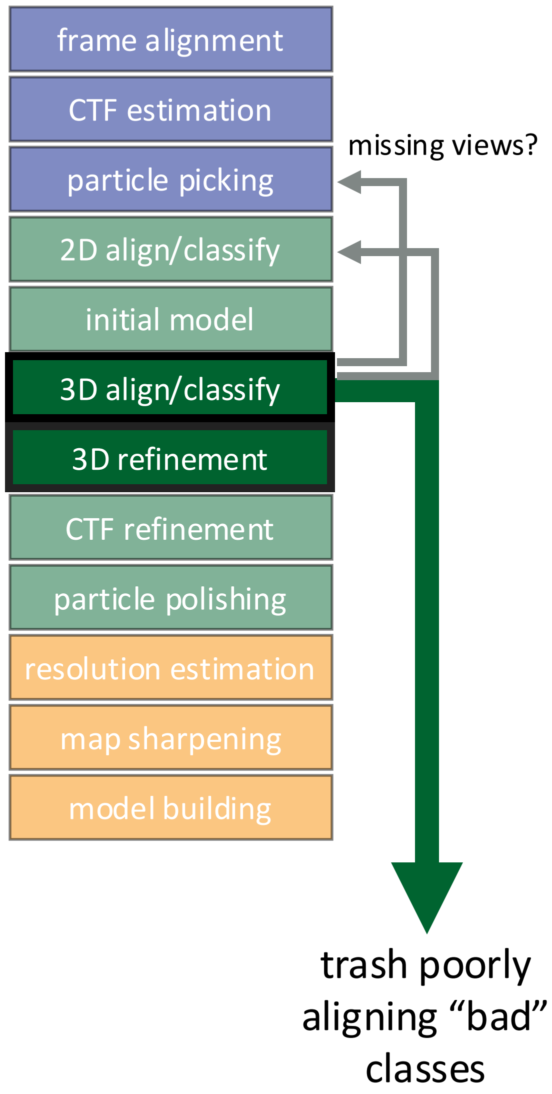
What B-factor to use?

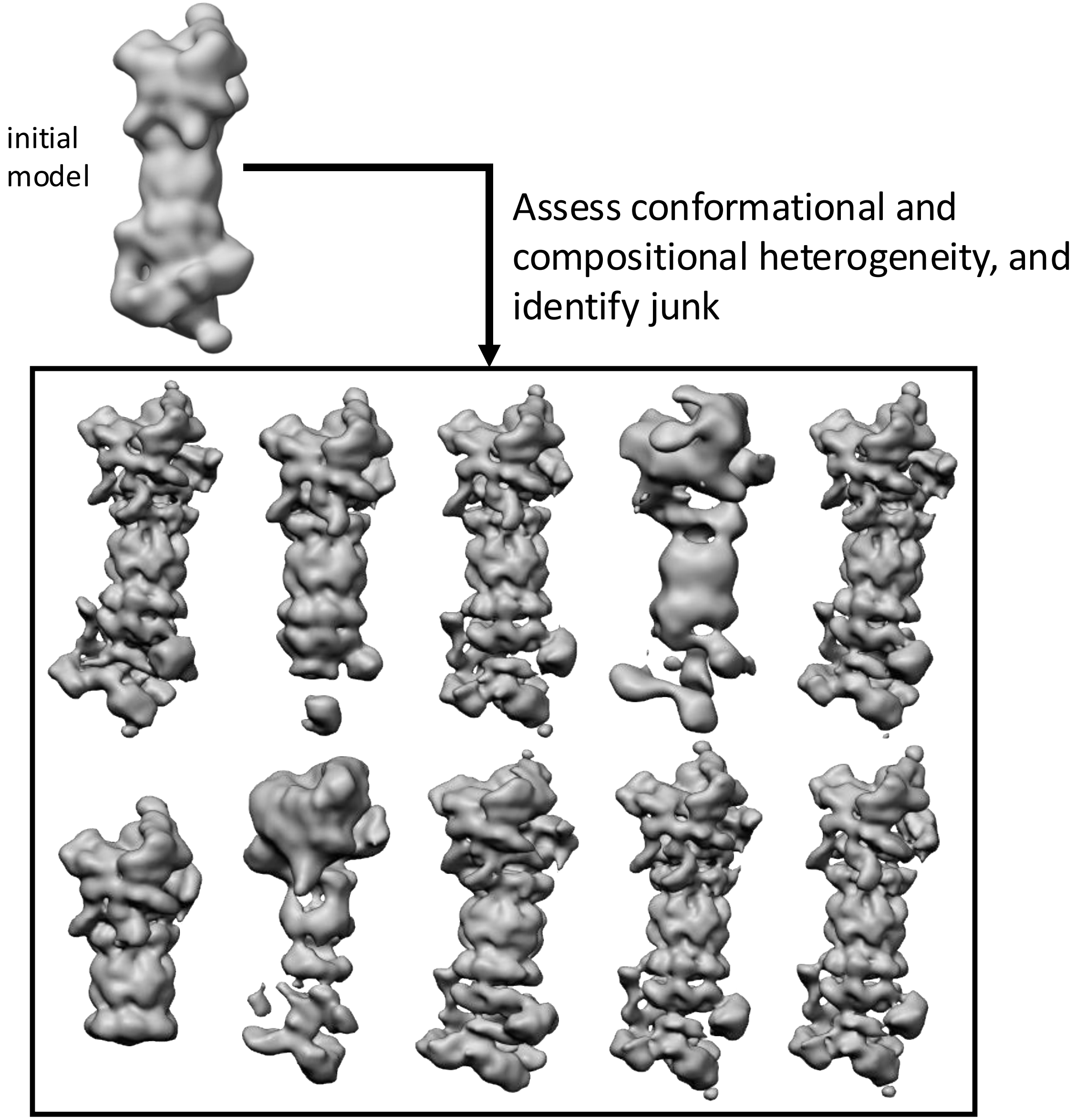
- Advanced processing software assign local B-factor



output: map with enhanced high resolution features

3D Classification – Resolving diversity across particles

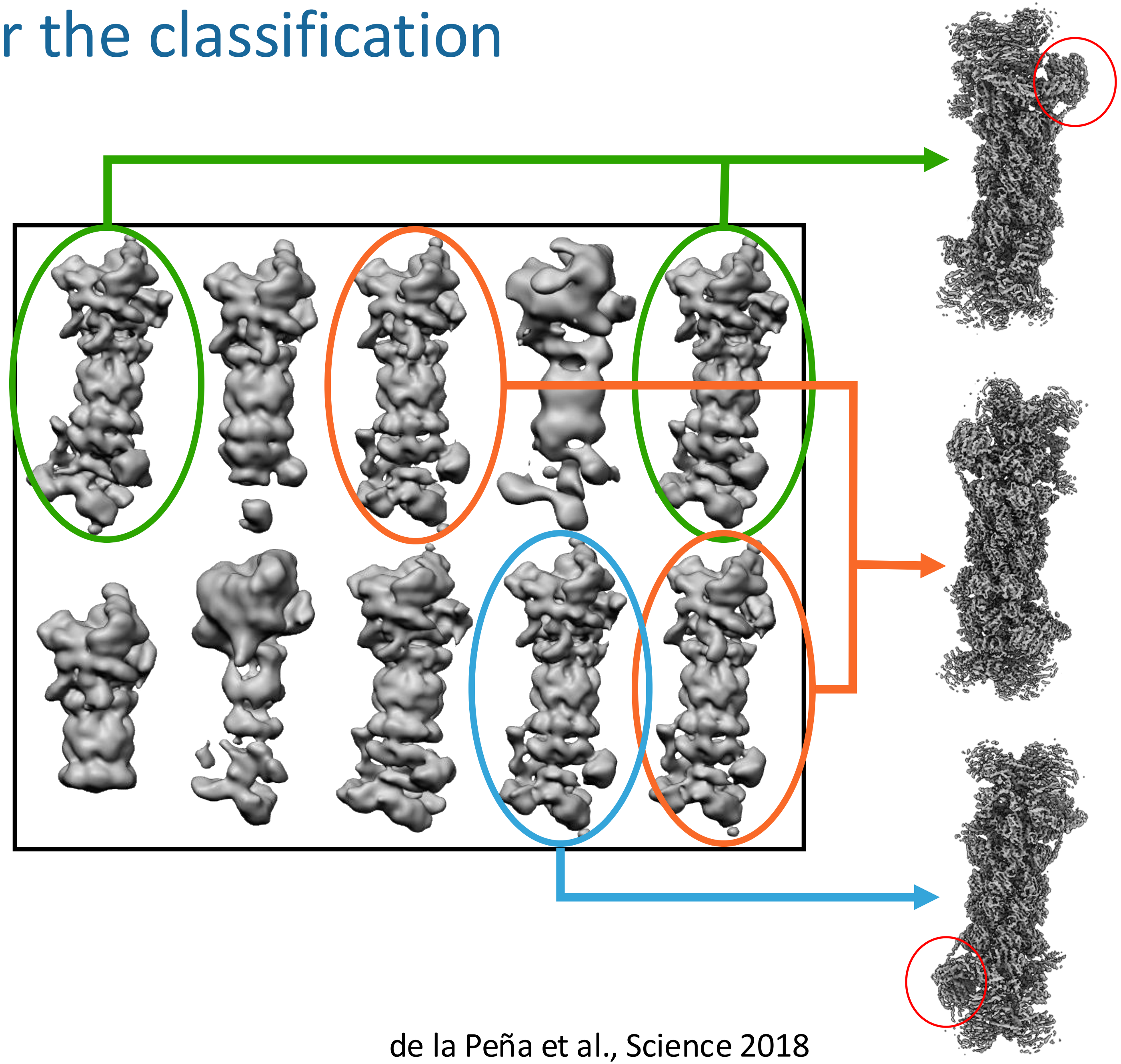




de la Peña et al., Science 2018

3D Refinement after the classification

- frame alignment
- CTF estimation
- particle picking
- 2D align/classify
- initial model
- **3D align/classify**
- **3D refinement**
- particle polishing
- resolution estimation
- map sharpening
- model building



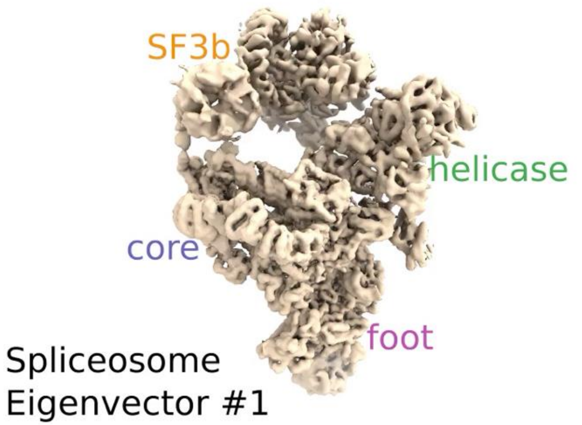
de la Peña et al., Science 2018

Methodologies for more complete descriptions of conformational variability

- frame alignment
- CTF estimation
- particle picking
- 2D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- resolution estimation
- map sharpening
- model building

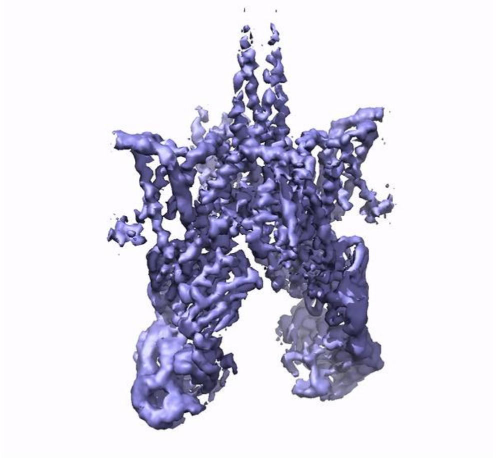
3D landscape characterization (3D align/classify, 3D refinement)

Multi-body refinement (RELION)



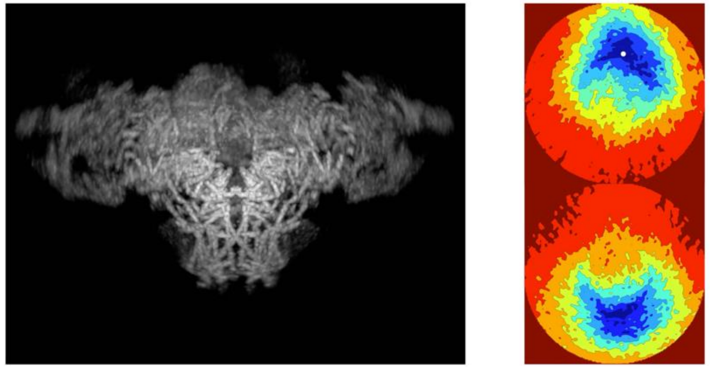
Nakane et al. eLife 2018

Variability Analysis (CryoSPARC)



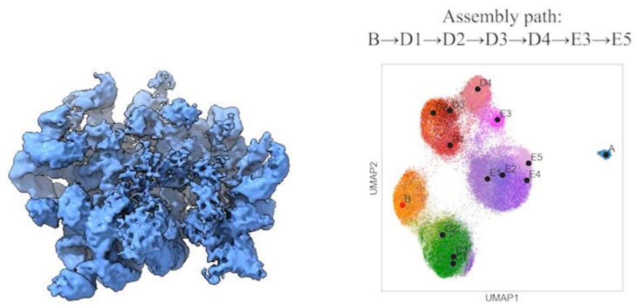
Punjani & Fleet JSB 2021

ManifoldEM



Dashti et al. Nat Comm 2020

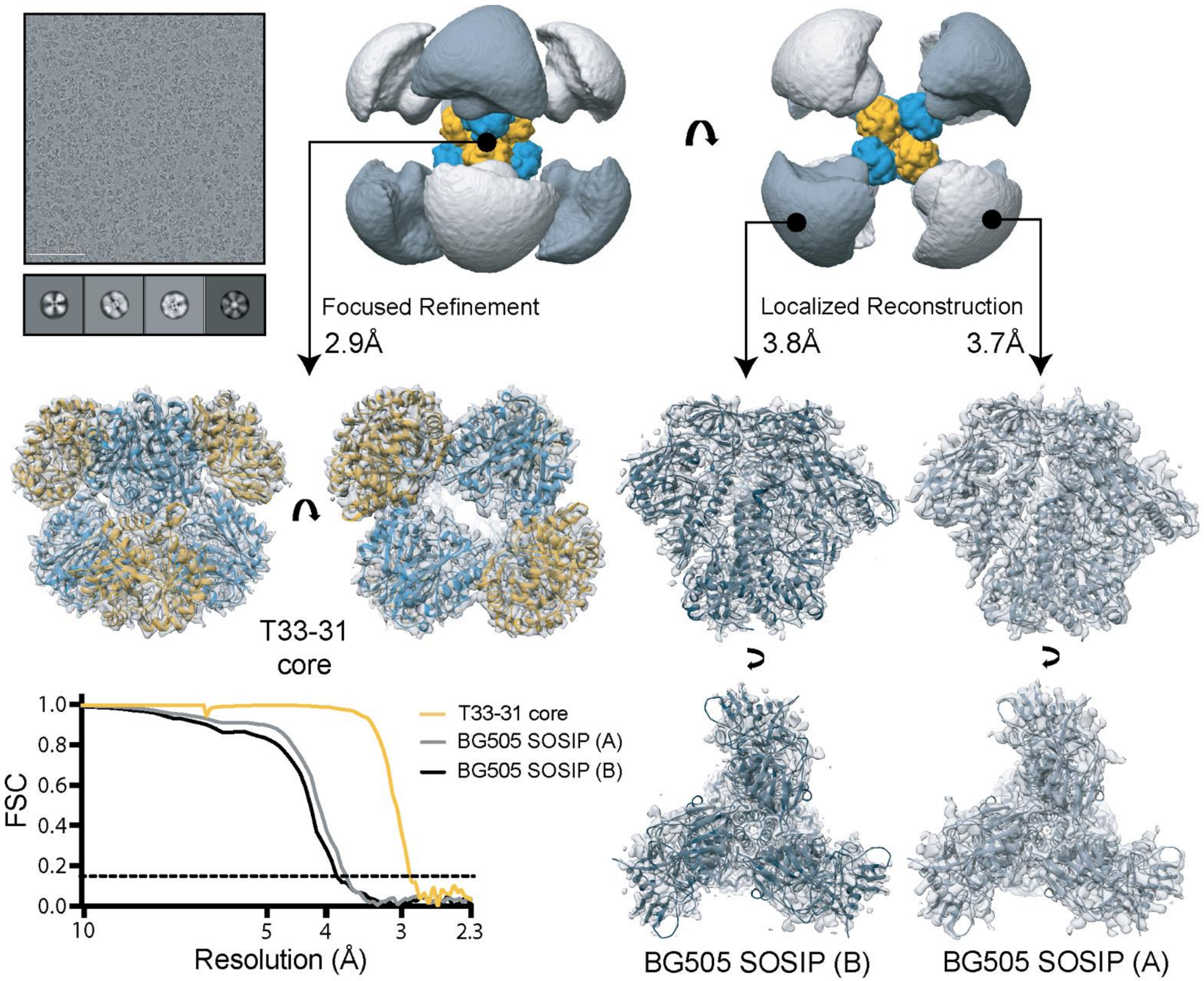
CryoDRGN



Zhong et al. Nat Methods 2021

3D refinement with partial signal subtraction

- frame alignment
- CTF estimation
- particle picking
- 2D align/classify
- initial model
- 3D align/classify
- 3D refinement
- particle polishing
- resolution estimation
- map sharpening
- model building



Protein nanoparticles (blue and gold) carrying 8 flexibly linked HIV glycoproteins (different shades of gray)

Antanasijevic et al., 2021

The main topics/questions from today's lecture

- What is the difference between amplitude contrast and phase contrast?
- How is image created in electron microscopy?
- What is a Contrast Transfer Function?
- What are the different components of the data processing workflow?
- What is the purpose of 2D classification and how does it work?
- What are Euler angles?
- How to generate an initial model of your particles?
- How does projection matching work?
- How is resolution calculated in EM?